Measuring the Value of Targeted Television Advertising

Lu Liao
University of Wisconsin
lliao5@wisc.edu

Alan Sorensen
University of Wisconsin & NBER
sorensen@ssc.wisc.edu

Andrey Zubanov
University of Wisconsin
zubanov@wisc.edu

June 2020

PRELIMINARY DRAFT: PLEASE DO NOT CIRCULATE OR CITE

Abstract

We propose a modeling framework for measuring the value of targeted advertising, and apply that framework to detailed data on television advertising. Brands can concentrate their advertising in a particular demographic group only insofar as they can find television programs whose viewers are predominantly in that group. We measure the degree to which this is possible with the current menu of programs, and then analyze the advertising profiles—i.e. the demographic composition of the purchased advertising impressions—of over 13,000 brands to estimate their demographic group-specific returns to advertising. With these estimates we can simulate how changes in the ability to target affect brands' overall returns to advertising.

1 Introduction

Over \$60 billion is spent on television advertising in the United States each year, but traditional televison's share of the advertising market has fallen steadily over the past decade. Meanwhile, the share of internet advertising has steeply increased.[1] One explanation for this trend is that the internet enables finer targeting, allowing brands to more easily direct their advertisements to individuals with certain demographic characteristics or past purchase behaviors. Television reaches many viewers but allows comparatively less control over who those viewers are.

This paper analyzes detailed data on advertising patterns to measure the degree to which targeting is possible on traditional television and to quantify the value of this targeting. Our framework begins with the observation that for television advertising, an advertiser's ability to target is limited by the degree to which television shows' viewers are concentrated in particular demographic categories. If a firm wishes to advertise its product primarily to women, for example, it can do so most efficiently if it can find programs that are watched exclusively by women. If programs intended for women are still watched by at least some men, then since ad slots' prices reflect *all* the impressions generated by the ads—not just the impressions in the advertiser's desired demographic—the advertiser targeting women must necessarily also purchase some spillover impressions for men.

We use data on national television programs' demographic viewership breakdowns along three dimensions (age, sex, and income) to determine how narrowly an advertiser could target its advertisements—or, put differently, to determine how many spillover impressions it would need to buy in order to reach a given level of impressions in a target demographic. We then examine detailed data on the advertising purchases of over 13,000 brands and develop a framework for estimating the distribution of brand-specific returns to advertising for different demographic groups. With the estimates from our model we can then simulate the value to advertisers of improved targeting—i.e., the value of having a (counterfactual) menu of television programs that allowed for more narrow targeting based on age, sex, and/or income.

While this research is still a work in progress, we can summarize our preliminary findings as follows. First, while the variety of available television programs allows for some degree of demographic targeting, such targeting can be at best imperfect. For instance, even the shows most heavily skewed to female viewers still have around 15 percent male viewers. Moreover, some demographic groups are much easier to target than others. For example, there are many TV programs that are watched almost exclusively by viewers over the age of 50. Second, brands vary considerably

[1] One study by PwC estimated that global spending on internet advertising surpassed spending on television advertising in 2015, and based on current trends internet spending will be more than double television spending by 2022.

in how (and how much) they target different demographic groups, but even the brands with the most targeted campaigns are not completely concentrated in one group. Third, when we use our estimates of brands "preferences" for demographic groups to simulate how changes in targeting ability would affect the returns to advertising, we find the effects to be small. This is primarily because our estimates imply that the majority of firms are unconstrained—i.e., the existing menu of programs allows them to allocate their advertising investments across demographic groups in a way that fully maximizes their returns.

Most prior work on the economics of targeted advertising has focused on theoretical aspects of the highly targeted advertising that can be achieved on internet platforms. For example, Bergemann and Bonatti (2011) describe a model of competing advertising markets with differential targeting ability, motivated by the distinction between internet and traditional media (like television and newspapers). They show that an increase in targeting capabilities will raise the value of advertising by enabling higher quality consumer-product matches, but also have the offsetting effect of increasing the concentration of firms advertising in each market. Their focus is on what happens to the equilibrium price of advertising as targeting capacity increases. In a subsequent theoretical paper (Bergemann and Bonatti (2015)) they also explore a setting in which advertisers' targeting can be extremely granular, such as when advertisements can be triggered by online search queries.

Gal-Or and Dukes (2003) analyze a model of locational choice among competing media outlets (e.g. television stations) and describe circumstances under which outlets have incentives to minimize differentiation. Their model is most applicable to broadcast television networks, which tend to provide content aimed at broad audiences. They note that cable channels actually tend to maximize differentiation, tailoring their content to narrower audiences in order to facilitate more targeted advertising. This idea is echoed by Athey and Gans (2010), who also note that a key feature of newer internet technologies is that they enable targeting without having to tailor content. Unlike television channels, where viewer demographics are directly a function of the types of programs the channel chooses to provide, internet sites can offer general-interest content but still serve different advertisements to different users depending on their characteristics, thus allocating ad space more efficiently.

Some work on targeted advertising has emphasized the interaction with privacy concerns. Johnson (2013) describes a model in which targeted advertising is valuable to firms but annoying to consumers, perhaps because they find such advertising to be an intrusive violation of their privacy. In the equilibrium of the model, improved targeting increases the amount of advertising, increasing firms' profits but decreasing consumers' welfare, because consumers get disutility from marginal advertisements even though they may benefit from the inframarginal ones. In an empirical paper analyzing data from a large field experiment, Goldfarb and Tucker (2011a) show that online banner ads are more effective when they are matched to the website's content (a form of targeting) or when they are made more obtrusive, but when ads are *both* more targeted and more obtrusive they are actually less effective. They interpret this effect as indicative of consumers' negative reactions to what they may see as a violation of privacy. In a related paper, Goldfarb and Tucker (2011b) show that privacy regulations enacted in the European Union in 2003-2004 appear to have reduced the effectiveness of online banner ads. These papers highlight a potentially important difference between targeted advertising on television vs. the internet. Because traditional television advertising is inherently less targeted, it is less vulnerable to the complications that result from privacy concerns.

We outline our framework for thinking about targeted television advertising in the next section. We then explain our data sources in Section 3 and describe several patterns in the data that provide both motivation and context for our empirical approach. In Section 4 we explain the details of our empirical analysis, which proceeds in four steps. First, we use the data to estimate advertising prices (cost per thousand impressions) for different demographic subgroups. Second, we estimate the distribution of brands' "preferences" over demographic groups. Third, we quantify spillover impressions—i.e., for each impression purchased in demographic subgroup g, how many impressions are collaterally purchased in subgroup g'. Finally, given the estimates of brands' preferences and of the spillover impressions, we conduct counterfactual simulations in which we relax or tighten the targeting constraints (by reducing or increasing the spillover rates) and measure how much these changes would affect the brands' overall returns to advertising.

We view our study as a useful step in the empirical analysis of targeted advertising, but there are a few important limitations we must acknowledge from the start. Our counterfactual analyses hold advertising prices fixed, so they do not incorporate any general equilibrium effects. Also, we conduct our analysis for a given categorization of demographic subgroups; what we are measuring is the value of improved targeting for these subgroups, rather than the ability to target a more finely subdivided set of demographic groups. Last, and perhaps most obviously, our analysis yields an estimate of the value of targeted advertising on television, which may not neatly extrapolate to the value of targeted advertising in other media, including internet. However, the framework we develop can be readily applied to other media if and when more detailed data become available.

2 Theoretical framework

In this section we briefly outline our modeling framework for quantifying the benefits of targeting. Details about how we estimate prices and parameterize brands' revenue functions are postponed until Section 4.

The supply side of the market consists of N advertising slots (ads), each of which can be described by a $G \times 1$ vector a_n whose elements a_{ng} indicate the number of impressions ad n will generate in demographic group g for $g = 1, \ldots, G$. These categories are demographic buckets like "females with high incomes between the ages of 18 and 49."

On the demand side are J brands that purchase the ads. Each brand purchases many ads, but what we will focus on is the total advertising by demographic group. Firm j's ad purchases are summarized by the $G \times 1$ vector e_j. Its elements e_{jg} indicate the total number of impressions purchased by firm j in demographic group g, normalized by the total number of persons in group g.[2]

Let p_g be the market price for an impression in demographic group g. These prices represent the expected marginal cost to a firm of purchasing additional impressions in group g. These prices are not directly posted or observed, but they can be estimated from the data as we discuss in more detail below. Absent any constraints, each firm would choose its ad purchases to solve

$$\underset{\{e_1,\ldots,e_G\}}{\text{Max}} \; R(e_1, \ldots, e_G) - \sum_{g=1}^{G} p_g e_g \tag{1}$$

where $R(\cdot)$ is the advertising revenue function, meaning the function that translates advertising impressions into additional sales revenues. The first-order conditions would then be

$$\frac{\partial R}{\partial e_g} = p_g \;\; \text{for } g = 1, \ldots, G \tag{2}$$

In other words, in the absence of constraints, firms target their advertisements to the demographic groups for which they expect the biggest increases in revenues, spending on each demographic group up to the point where marginal revenue equals marginal cost.

Within this framework, we can describe limits on a brand's ability to target as constraints on its ability to purchase impressions solely within one demographic group. Suppose, for example, a firm sells a product that is only relevant to consumers in group g. Then the firm would like to spend all its advertising dollars on group g. However, it is impossible to do this, because even television programs that are watched primarily by consumers in group g are also watched by at least some

[2] In other words, e_{jg} can be thought of as the gross rating points (GRP) purchased by firm j in demographic group g, a measure that is sometimes called target rating points (TRP).

consumers in other demographic groups. Put differently, there is no supply (or at least not a sufficient supply) of ads for which all the purchased impressions would be in group g.

To formalize this notion, we assume that for every impression purchased in group g, another γ_{gh} impressions must be purchased in group h. The γ_{gh} parameters thus represent how targeted a firm's advertising can be: if the γ_{gh}'s are all zero, then perfect targeting (within the G groups) is possible; if $\sum_h \gamma_{gh} = G - 1$, then it is only possible to buy completely untargeted advertising, meaning an equal number of impressions in all groups.

In this setting, a firm's desired demographic targets may not be what we observe in the data. Let $e = (e_1, \ldots, e_G)$ be the firm's chosen targeting vector, and let $\tilde{e}$ be the vector of impressions we observe in the data, so

$$\tilde{e}_g = e_g + \sum_{h \neq g} \gamma_{gh} e_h \tag{3}$$

The cost associated with e_g is the direct cost along with all of the additional spillover impressions that must be purchased, and the firm's profit function becomes

$$\pi(e) = R(\tilde{e}_1, \ldots, \tilde{e}_G) - \sum_{g=1}^{G} e_g \left(p_g + \sum_{h \neq g} \gamma_{gh} p_h \right) \tag{4}$$

Firms, knowing the spillover rates γ_{gh}, choose their target vectors e to generate actual impressions $\tilde{e}$ to maximize profits. Importantly, firms cannot choose negative targets: when maximizing profits, the firm must set each $e_g \geq 0$. This reflects the simple reality that purchasing negative impressions isn't possible, but also ensures that the spillovers meaningfully change the firm's optimization problem. If negative e_g's were possible, firms could always find a combination of e_g's that generates exactly the desired vector of impressions.

The system of first order conditions for profit maximization is then

$$\frac{\partial R}{\partial e_g} + \sum_{h \neq g} \gamma_{gh} \frac{\partial R}{\partial e_h} = p_g + \sum_{h \neq g} \gamma_{gh} p_h + \lambda_g \text{ for } g = 1, \ldots, G \tag{5}$$

along with the complementary slackness conditions

$$\lambda_g e_g = 0 \text{ for } g = 1, \ldots, G$$

where the λ_g's are the Lagrangian multipliers. The extra terms on the left hand side of the first-order conditions reflect the assumption that buying an additional impression in group g requires buying γ_{gh} additional impressions in group h, and the extra terms on the right hand side reflect the additional cost of those extra impressions.

Though somewhat simplistic, this framework is useful because the γ_{gh}'s can be estimated directly from demographic data on television programs' viewers. Along with estimated group-specific prices p_g, this makes it possible to back out firm-specific estimates of the marginal revenues $\partial R / \partial e_g$, or an estimate of the distribution of firms' heterogeneous "preferences" over demographic groups. These estimates can then be used to evaluate the marginal profit gain that would result from a decrease in the γ_{gh}'s—i.e., improved targeting in the form of a relaxation of the spillover constraints.

3 Industry Background and Data

Before describing our data, we first provide a very brief overview of the market for television ads. The main participants in this market are content producers (television channels), advertisers, and viewers. Television channels supply content (television programs) that attracts viewers, and advertisers pay the television channels for the ability to show advertisements to those viewers. A channel's content may be targeted at broad audiences (e.g. *American Idol*) or at narrower audiences (e.g. a children's cartoon), and the market price for an ad slot during a particular program will depend on both the number and composition of that program's viewers.

In order to reach the right audiences for their products, firms try to carefully select the channels, programs, and times of day in which to run their advertisements. They often employ third-party agencies, called media buyers or media planners, to help ensure their advertising budgets are well spent. With this in mind, our framework treats advertisers as optimizers, allocating their expenditures in a way that maximizes the returns. However, there are some limits to how precisely firms can optimize their spending in practice. For instance, most ad slots are purchased months in advance, so at the time of purchase firms cannot know exactly what the audience sizes and compositions will be. In spite of these frictions, we believe it is reasonable to interpret brands' overall expenditure profiles as indicative of the relative importance they place on advertising to different groups of viewers.

3.1 Data

Our advertising data come from Nielsen Ad Intel and are provided by the Kilts Center for Marketing at the Chicago Booth School of Business. We focus on nationally televised advertisements in the year 2015. For each of 13,905 brands, we observe every advertising *occurrence* on national television,[3] along with an estimate of the *impressions* associated with that occurrence. An occurrence is an advertisement on a particular channel at a particular time—e.g., a 30-second slot on CNN at 7:48 PM on 20 Apr 2015. There are a total of 24,057,831 occurrences in the data. Impressions are the number of viewers Nielsen estimates were watching the advertisement, and are broken down into various demographic categories. In the analyses below we focus on the 12 demographic categories defined by age (under 18, 18 to 49, and 50 and over), sex, and income (below or above $60,000).

Data on occurrences are captured by Nielsen using proprietary technologies, which recognize the commercials on TV channels. The data on impressions are gathered using Nielsen's People Meter—a panel of people who report their TV viewership habits. The data for the price per ad slot are collected by Nielsen by contacting the monitoring networks, whenever available. If unavailable, other industry sources, including proprietary models are used.

Table 1 shows summary statistics for the advertising occurrences and the resulting impressions. Most advertisements are either 15 or 30 seconds long, with an average cost to the advertiser of $2,388 per 30 seconds. As we discuss below, however, the more relevant cost is the cost per thousand impressions, which for our sample is on average $3.26. The number of impressions per ad varies widely across ads, as does the distribution of impressions across demographic groups. For some ads the percentage of female viewers is over 80 percent, for example.

Table 2 lists some examples of television programs for which viewership is highly concentrated in a demographic group. For instance, *What Not to Wear* on TLC is mostly watched by women, and *The Herd with Colin Cowherd* on Fox Sports 1 is mostly watched by men. Given that each of the listed examples is among the most extreme in the data, one of the lessons from this table is that even if an advertiser purchases ad slots on the programs with the narrowest viewerships, it still cannot perfectly target a particular group. It also highlights that some demographic groups are easier to target than others. There are shows that are watched almost entirely by older viewers or by low-income viewers, but in other categories it is harder to find shows that are so focused. For example, the *U.S. Open Golf Post-match* show was among the programs most heavily skewed toward high-income viewers, but it still had over two thirds of its viewers in the lower income category.

[3] National channels include those on major networks like NBC and Fox, plus cable channels like CNN and TBS.

Table 1: Summary statistics: occurrences ($N = 24,057,831$)

	Mean	Std. Dev.	Percentile .10	.50	.90
Duration (seconds)	28	19	15	30	60
Cost ($ per 30 secs.)	2388	18591	69	532	3602
Cost ($ per 1000 impressions)	3.26	4.72	0.68	2.10	6.83
Impressions:					
Total (1000)	391	874	39	220	794
% female	52.5	17.7	26.6	53.3	75.3
% high inc.	40.0	12.5	24.7	39.8	55.9
% age 18-49	39.5	14.4	20.2	39.7	58.0
% age 50+	48.0	20.8	17.4	49.1	75.4

Table 2: Examples of TV programs with concentrated viewership

		Percentages of viewers			
Show	Network	Female	High income	Age < 18	Age ≥ 50
What Not to Wear	TLC	84.7	18.6	0.1	66.5
The Herd w/ Colin Cowherd	FS1	14.8	21.0	4.5	45.2
U.S. Open Golf Post-match	Fox	35.8	32.8	5.8	66.9
Mann & Wife	Bounce	56.0	1.4	10.1	47.6
Kirby Buckets	Disney	40.9	15.6	73.2	5.9
Joey & Rory Show	RFD	52.1	4.4	1.6	95.1

Summary statistics for brands' advertising expenditures are shown in Table 3. We follow the structure of the data and treat brands as the advertising unit, rather than the firm that owns the brands. So, for example, Coke and Sprite are separate advertisers, even though both brands are owned by Coca-Cola. As the table shows, there is enormous heterogeneity across brands in the overall level of spending. The largest advertisers in our sample, like major auto insurance companies, spend over $450 million per year, while the smallest spend less than $10.

Brands are also heterogeneous in how they allocate their spending across demographic groups. For many brands, impressions are spread more or less evenly across demographic groups. But some brands clearly target demographic subgroups. Figure 1 shows four of the most extreme examples of demographically targeted advertising. Each plot shows the brand's impression shares across six categories defined by sex and age groups.[4] The mens shaving cream brand targets its advertising disproportionately to adult men, and the bra brand's advertisements are unsurprisingly targeted at

[4]To keep the plots visually uncomplicated, we left out income categories. But there are similar examples of brands that focus their advertising on high or low income viewers, such as Charles Schwab (71.8% of impressions are high-income viewers) and McDonalds (78.2% of impressions are low-income viewers).

Table 3: Summary statistics: brands ($N = 13,905$)

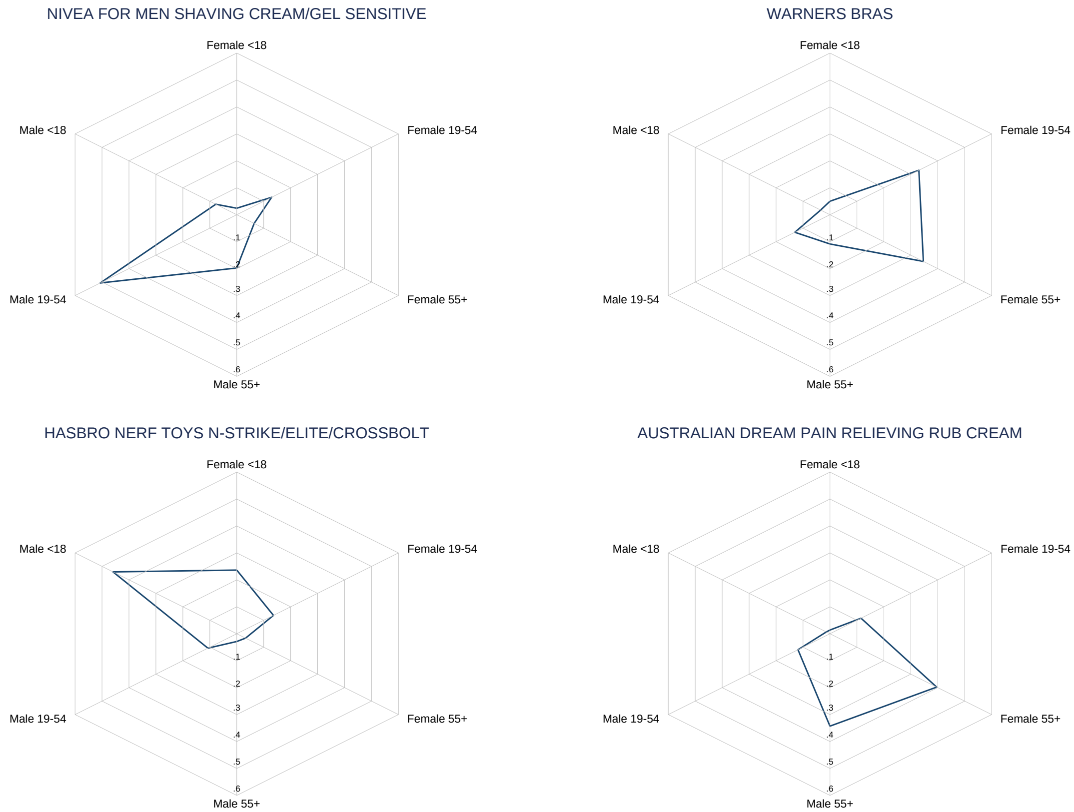
	Mean	Std. Dev.	Percentile .10	Percentile .50	Percentile .90
Number of ads	1730	5665	3	97	4570
Total spending ($1000)	3534	13106	1.4	101	8272
Impressions:					
Total (1000)	525	1854	40	297	798
% female	46.6	15.3	24.5	47.5	66.2
% high income	41.3	11.9	25.1	41.1	56.1
% age 18-49	34.0	12.7	15.7	35.5	48.9
% age $\geq$ 50	54.0	20.5	27.9	54.2	81.7

adult women. What is notable about these examples is that even though they are among the most extreme in the data, they are still quite far from being fully targeted. For example, women account for roughly a fifth of the impressions for the mens shaving cream, and men make up roughly a quarter of the impressions for the bra brand. This highlights how difficult it is to make television advertisements narrowly targeted.

4 Empirical analysis and results

Our empirical analysis proceeds in four steps. First, we use the data on advertising occurrences to estimate price per impression (or CPM) separately for demographic groups. Throughout the analysis we consider the 12 demographic groups defined by the interactions of sex (male or female), income (low or high), and age (under 18, 18-49, or 50 and above). Second, we specify a parametric model of brands' net returns to advertising, and—based on the assumption that brands allocate their advertising expenditures optimally across demographic groups—use the data to recover estimates of the distribution of brands' group-specific returns. Third, we use the data on TV programs' viewerships to measure the spillover impressions γ_{gh} described in Section 2. Fourth, we simulate how changes in the ability to target—which we capture with changes in the spillover rates γ_{gh}—affect brands' returns to advertising.

Figure 1: Examples of brands with demographically targeted advertising



4.1 Estimating cost per impression by demographic group

The conventional measure of an ad slot's price is the cost per thousand impressions,[5] since the value of an ad slot lies in the number of viewers it will reach. For our purposes, we need to measure prices specific to demographic groups—for example, the cost per thousand female/high-income/middle-aged impressions. We do this by running the following regression:

$$\log(\text{spend}_i) = \log(V_i'\theta) + Z_i'\delta + \epsilon_i \tag{6}$$

Here i indexes advertising occurrences, and spend_i is the amount that was paid for occurrence i. V_i is a 12×1 vector indicating the number of impressions (in thousands) in each of the twelve demographic subgroups, Z_i is a vector of controls including dummies for channel, month, day of week, and prime time, and ϵ_i is a mean-zero error. The estimated θ vector thus captures

[5]The cost per thousand impressions is typically labeled CPM (cost per mille).

differences in the costs of advertising to the different demographic groups, while the elements of δ multiplicatively scale those costs to control for aspects of the particular ad like whether it was aired during prime time, on a weekend vs. a weekday, etc.

The predicted marginal cost per thousand impressions in demographic group g is $\theta_g \cdot exp\{Z_i'\delta\}$, which we calculate for our estimated values of θ and δ at the average value of Z_i and report in the table below.

Table 4: Estimated costs per thousand impressions, by demographic subgroup

Sex	Age	Income	CPM
Female	< 18	$<$ \$60k	1.17
Female	18-49	$<$ \$60k	2.38
Female	≥ 50	$<$ \$60k	2.03
Female	< 18	$\geq$ \$60k	2.02
Female	18-49	$\geq$ \$60k	2.57
Female	≥ 50	$\geq$ \$60k	2.04
Male	< 18	$<$ \$60k	1.89
Male	18-49	$<$ \$60k	3.04
Male	≥ 50	$<$ \$60k	2.81
Male	< 18	$\geq$ \$60k	4.21
Male	18-49	$\geq$ \$60k	3.88
Male	≥ 50	$\geq$ \$60k	4.37

The CPM estimates in Table 4 may appear low relative to some industry reports, which tend to put average CPM for television advertising in the \$10-\$25 range. One reason for this apparent discrepancy is that industry reports tend to focus on prime time advertising to adult populations, and may not distinguish between broadcast and cable channels. The prices in Table 4 are reported for an average value of the controls Z; if instead we calculate predicted CPM for prime-time slots on broadcast networks, the numbers are mostly in the \$7-\$17 range. Industry studies that average over all channels, days of week, and times of day report numbers closer to our estimates.[6]

4.2 Recovering brands' marginal returns to advertising

Having estimated demographic group-specific prices, we use brands' observed advertising expenditures to estimate the distribution of brand "preferences" for different demographic groups. The central assumption is that brands allocate their ad spending optimally in a way that equates marginal returns across demographic groups. We parameterize a brand's net revenue function as

[6] For example, a 2018 article on MediaPost puts the overall average CPM for television ads at \$2.26.

$$R(e;\alpha,\beta) = R_0 + \beta \sum_g \alpha_g N_g \frac{1}{\phi} \left(\frac{e_g}{N_g} \right)^{\phi} - \sum_g p_g e_g \tag{7}$$

where e is the vector of impressions in demographic groups $g = 1, \ldots, G$, N_g is the size of group g, and p_g is the price per impression for group g. This function represents advertising's contribution to profits—i.e. the technology with which the brand can convert advertising expenditures into net revenues. The parameter β scales the returns to advertising across all demographic groups (brands that have high β's will do a lot of advertising) and the α_g parameters determine the returns to advertising specific to a demographic group (brands with higher α_g's for females will want to target their advertising toward females). With β as the scaling factor, we impose without loss of generality the normalization that $\sum_g \alpha_g = 1$, so the α_g's can be interpreted as relative weights on the different demographic groups. A brand's advertising "preferences" are thus characterized by its β and its α_g's, and our primary purpose is to estimate the distribution of these preferences.

The parameter ϕ is an important parameter that governs the degree to which advertising exhibits diminishing returns. If $\phi = 1$, each brand would prefer to advertise exclusively to the one group for which it has the highest α_g. We assume ϕ is common across brands.

Normalizing impressions e_g by group size N_g is important to ensure that marginal returns to advertising decrease equally fast for large groups and small groups. Without this normalization, if there were twice as many people in group h as in group g then brands would advertise less intensively in group h simply due to diminishing returns. Also, normalizing in this way allows for a natural link between the equations and real-world measurements: $100 \cdot \frac{e_g}{N_g}$ represents the gross rating points (GRP) purchased by the brand in demographic group g, which is the advertising industry's conventional measure of advertising intensity.[7] GRPs can be understood as measuring advertising intensity because the ratio $\frac{e_g}{N_g}$ can be interpreted as the number of times a representative viewer from group g has watched the brand's ads.[8]

For purposes of estimation, we assume that each brand's expenditure vector e results from an *unconstrained* maximization problem. This may seem at odds with this paper's central premise, which is that the menu of available programs may not permit a brand to target its advertising as narrowly as it would like. However, there are several reasons for imposing this simplification. First, it makes estimating the key parameters more straightforward. As shown below, for a fixed value of ϕ the optimality conditions of the unconstrained problem yield a system of G nonlinear equations that can be solved analytically to get brand-specific values of β and the α_g's. Second, while it would

[7] Demographic group-specific gross rating points are sometimes called target rating points (TRP).

[8] This does not distinguish if one person watched the ads 5 times or 5 people watched once each.

be possible (though more computationally burdensome) to estimate the distribution of β's and α_g's under the assumption that brands' advertising expenditures are the outcomes of *constrained* maximization problems—e.g. an analogue to the problem described in Section 2, with spillover impressions—it is very difficult to infer from the data which brands are constrained and which are not. In principle constrained firms' advertising purchases would line up along discernible boundaries, but in practice there is enough randomness in firms' advertising purchases that any bunching near boundaries is quite noisy. Finally, for the purposes of our counterfactual simulations below, we believe this approach to estimation is conservative, in the following sense. If some of the brands in our sample are in fact constrained, our approach—which assumes they are unconstrained—will lead to a slight downward bias in the estimates of how variable are the α_g's. Loosely speaking, if a brand spends 90% of its advertising on female impressions, our approach assumes this is precisely where the brand wanted to be—when in reality it may have been constrained to be at 90% instead of 95%. This will lead our estimator to infer brand "preferences" that are more concentrated in the middle of the distribution, i.e. a distribution of the α_g's that has most brands wanting to advertise evenly across groups. So when we simulate the value of relaxing the constraints by reducing spillover impressions, we will tend to underestimate how many brands are in the extremes of the distribution, and thus underestimate how many (and how much) brands will benefit from the relaxed constraints.

The solution to the brand's unconstrained maximization problem has the G first order conditions

$$\beta\alpha_g \left(\frac{e_g}{N_g}\right)^{\phi-1} = p_g \text{ for } g = 1, \ldots, G \tag{8}$$

Since the α_g's sum to one, these first order conditions imply that

$$\beta = \sum_{g=1}^{G} p_g \left(\frac{e_g}{N_g}\right)^{1-\phi} \tag{9}$$

and

$$\alpha_g = \frac{p_g \left(\frac{e_g}{N_g}\right)^{1-\phi}}{\sum_{h=1}^{G} p_h \left(\frac{e_h}{N_h}\right)^{1-\phi}} \tag{10}$$

Since we have estimates of group-specific prices p_g (as explained in Section 4.1), and group-specific GRPs $\frac{e_g}{N_g}$ can be computed from the data, equations (9) and (10) allow us to back out each brand's

β and α_g's for a given value of ϕ.

We assume the brand-specific β's are draws from an exponential distribution, and estimate its mean μ_β. We assume the brand-specific α vectors are draws from the logistic-normal distribution:

$$\alpha_g = \frac{\exp\{a_g\}}{1 + \sum_{h=2}^{G} \exp\{a_h\}}$$

with a_1 normalized to zero and the vector $[a_2 \cdots a_G]$ distributed multivariate normal,[9] and we estimate the corresponding mean μ_a and covariance matrix Σ_a. Since we can back out brand-specific values of β and α using equations (9) and (10), estimating the parameters of these distributions is straightforward for any given value of ϕ. We simply estimate μ_β to be the sample mean of the recovered β's, and μ_a and Σ_a to be the sample mean and sample covariance of the a vectors implied by the recovered α's.[10]

Identification of the diminishing returns parameter ϕ is less straightforward. Price variation would be ideal for identifying this parameter, but we are estimating the prices p_g from a single cross section of annual data. These prices are unlikely to change at a high frequency, and even if we used multiple years of data it would be difficult to control for other concurrent changes in brands' advertising incentives over such a long time period. So we instead estimate ϕ by choosing the value that best matches the mean and standard deviation of total advertising expenditures across brands. To explain this precisely, let $\hat{\mu}_\beta(\phi)$, $\hat{\mu}_a(\phi)$, and $\hat{\Sigma}_a(\phi)$ denote the parameter estimates for the distributions of β and α corresponding to a given value of ϕ. We can then draw S simulated brands from these distributions and compute their optimal advertising expenditures. Letting $\tilde{E}_{s,\phi}$ denote the total advertising expenditures of simulated brand s for the given value of ϕ, we compute the sample mean and sample standard deviation of $\tilde{E}_{s,\phi}$, and form the moment vector

$$m(\phi) = \begin{bmatrix} \text{mean}(E_k) - \text{mean}(\tilde{E}_{s,\phi}) \\ \\ \text{sd}(E_k) - \text{sd}(\tilde{E}_{s,\phi}) \end{bmatrix} \tag{11}$$

where E_k denotes the total advertising expenditures of brand k observed in the data. We choose ϕ to minimize $m(\phi)'Wm(\phi)$, with the weight matrix W chosen as suggested by Hansen (1982).[11]

[9] See Aitcheson and Shen (1980) for a description of this distribution and its properties.

[10] The a's corresponding to the α's can be computed from the relation $a_g = \log(\alpha_g) - \log(\alpha_1)$.

[11] Note: for the numbers reported in this draft, we used the identity matrix as the weight matrix.

This procedure yields an estimate of $\hat{\phi} = 0.572$ and an estimate of $\mu_\beta = 0.0247$.[12] Figure 2 plots a histogram of the brand-specific β's corresponding to the estimated value of ϕ, with the exponential density superimposed. While not a bad approximation, the exponential distribution is not a perfect fit, suggesting it may be useful to consider more flexible specifications for the distribution of β.

Figure 2: Distribution of β across brands

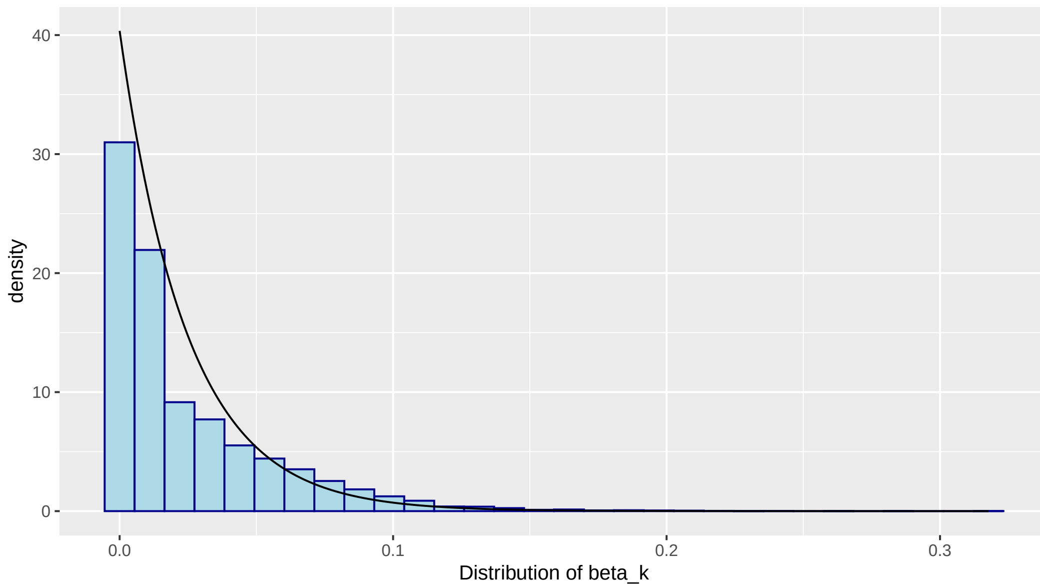


Table 5 reports the means μ_a and correlations from Σ_a across demographic groups. Unsurprisingly, correlations are highest for similar groups (e.g. 0.95 for low-income males over the age of 50 and high-income males over the age of 50) and lowest for dissimilar ones (e.g. 0.02 for low-income females under the age of 18 and high-income males over the age of 50).

4.3 Measuring spillover impressions

As described above in Section 2, our framework for understanding a brand's ability to target its advertising is based on the notion of spillover impressions: the additional impressions that must be purchased in demographic groups $h \neq g$ in order to reach a certain level of impressions in group g. To measure these spillovers in the data, we consider the problem of a sophisticated firm that has full knowledge of the demographic composition of every television program's viewers, and places its advertisements on programs in a way to target a specific demographic group as narrowly as possible. For every ad slot i in the data (which we define as an advertisement on a particular TV program on a particular day), the data indicate the number of impressions n_{ig} it generated in each demographic group g. To estimate the spillover rate between two groups g and h, we consider a

[12]The values of β are low because they scale the returns to the level of price per impression.

Table 5: Estimated means and correlations for the a parameters

		Correlations										
	Means	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
(1)		1.00	0.81	0.50	0.90	0.86	0.15	0.67	0.68	0.36	0.78	0.76
(2)		0.81	1.00	0.28	0.79	0.95	0.07	0.56	0.89	0.22	0.69	0.87
(3)		0.50	0.28	1.00	0.61	0.43	0.11	0.22	0.15	0.61	0.51	0.35
(4)		0.90	0.79	0.61	1.00	0.90	0.02	0.53	0.62	0.45	0.82	0.78
(5)		0.86	0.95	0.43	0.90	1.00	0.02	0.54	0.83	0.33	0.77	0.90
(6)		0.15	0.07	0.11	0.02	0.02	1.00	0.53	0.27	0.40	0.31	0.21
(7)		0.67	0.56	0.22	0.53	0.54	0.53	1.00	0.73	0.43	0.78	0.70
(8)		0.68	0.89	0.15	0.62	0.83	0.27	0.73	1.00	0.30	0.75	0.92
(9)		0.36	0.22	0.61	0.45	0.33	0.40	0.43	0.30	1.00	0.60	0.44
(10)		0.78	0.69	0.51	0.82	0.77	0.31	0.78	0.75	0.60	1.00	0.87
(11)		0.76	0.87	0.35	0.78	0.90	0.21	0.70	0.92	0.44	0.87	1.00

The reference group (for which a_1 is normalized to 0) is low-income males under the age of 18. The remaining groups are as follows:

(1) Low-income / Male / Age 18-49
(2) Low-income / Male / Age 50 and above
(3) High-income / Male / Age under 18
(4) High-income / Male / Age 18-49
(5) High-income / Male / Age 50 and above
(6) Low-income / Female / Age under 18
(7) Low-income / Female / Age 18-49
(8) Low-income / Female / Age 50 and above
(9) High-income / Female / Age under 18
(10) High-income / Female / Age 18-49
(11) High-income / Female / Age 50 and above

hypothetical firm that purchases slots sequentially in descending order of n_{ig}/n_{ih} until it reaches its target impressions for group g, which we will denote $\bar{n}_g$. In other words, the hypothetical firm purchases ad slots in a way that achieves its target in group g while minimizing impressions in group h. Letting $\mathcal{I}(\bar{n}_g)$ denote the set of ad slots purchased by a firm targeting $\bar{n}_g$ impressions in group g, we then calculate the spillover rate as

$$\gamma_{gh} = \frac{\sum\limits_{i \in \mathcal{I}(\bar{n}_g)} n_{ih}}{\sum\limits_{i \in \mathcal{I}(\bar{n}_g)} n_{ig}} \tag{12}$$

Since ad slots become less demographically concentrated as the firm moves down the list, the spillover rates γ_{gh} are larger when the target number of impressions $\bar{n}_g$ is large. Firms that intend to do very little advertising can purchase all they need by cherry-picking the ad slots that are most focused on the desired demographic; but the more a firm advertises, the tighter the spillover constraints become. Because of this, we calculate separate $\{\gamma_{gh}\}$ matrices for ten different levels of this target. In the counterfactual exercises below, each brand faces the $\{\gamma_{gh}\}$ matrix that most closely matches its level of advertising.

Table 6 reports the spillover rates when the target level of impressions $\bar{n}_g$ is 10 million (which would be representative of a brand roughly in the middle of the distribution). As one would expect, the largest spillovers are between similar groups—for example, 0.12 for low-income females under 18 and low-income females aged 18-50. Overall, however, the measured spillovers tend to be quite small. We believe this mostly reflects our conservative approach to calculating them: our hypothetical firm is extremely efficient about getting impressions in group g without getting impressions in group h. In practice, real world advertisers (and ad buying agencies) are probably not this sophisticated. However, the low spillover rates in Table 6 reveal that while a lot of television programming is watched by broad audiences, there are some programs that have very narrow audiences. If the target audience is children, for example, there are channels like Disney and Nickelodeon that are watched almost exclusively by children.

4.4 Quantifying the value of targeting

With estimates of our model's parameters and measures of the spillover constraints, we can estimate the value of targeted advertising by calculating the changes in returns to advertising that would result from a counterfactual relaxing or tightening of the spillover constraints. We do this in two ways. First we consider the brands observed in the data, along with the recovered values

Table 6: Spillover rates

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
(1)	1.000	0.038	0.006	0.000	0.012	0.001	0.005	0.071	0.010	0.001	0.015	0.001
(2)	0.001	1.000	0.025	0.000	0.009	0.003	0.000	0.018	0.014	0.000	0.008	0.002
(3)	0.000	0.002	1.000	0.000	0.001	0.005	0.000	0.001	0.030	0.000	0.000	0.002
(4)	0.008	0.060	0.008	1.000	0.032	0.002	0.003	0.059	0.014	0.005	0.044	0.002
(5)	0.000	0.038	0.027	0.001	1.000	0.008	0.000	0.006	0.011	0.000	0.016	0.004
(6)	0.000	0.006	0.116	0.000	0.011	1.000	0.000	0.001	0.023	0.000	0.005	0.017
(7)	0.009	0.045	0.005	0.001	0.017	0.001	1.000	0.121	0.018	0.001	0.026	0.002
(8)	0.001	0.027	0.008	0.000	0.004	0.001	0.001	1.000	0.026	0.000	0.006	0.001
(9)	0.000	0.001	0.022	0.000	0.001	0.002	0.000	0.001	1.000	0.000	0.001	0.002
(10)	0.015	0.061	0.009	0.009	0.039	0.003	0.010	0.141	0.039	1.000	0.083	0.005
(11)	0.001	0.019	0.010	0.000	0.022	0.003	0.001	0.032	0.037	0.000	1.000	0.004
(12)	0.000	0.010	0.043	0.000	0.010	0.031	0.000	0.004	0.119	0.000	0.021	1.000

Cells report the measured spillover impressions in the column group resulting from an impression purchased in the row group, for a firm targeting 10 million impressions in the row group. Groups are numbered as follows:

(1) Low-income / Male / Age under 18
(2) Low-income / Male / Age 18-49
(3) Low-income / Male / Age 50 and above
(4) High-income / Male / Age under 18
(5) High-income / Male / Age 18-49
(6) High-income / Male / Age 50 and above
(7) Low-income / Female / Age under 18
(8) Low-income / Female / Age 18-49
(9) Low-income / Female / Age 50 and above
(10) High-income / Female / Age under 18
(11) High-income / Female / Age 18-49
(12) High-income / Female / Age 50 and above

of β and α for each of those brands.[13] If we then tighten the spillover constraints by counterfactually increasing the γ_{gh} values, we can solve each brand's optimization problem under the new constraints and calculate the resulting change in its net revenue function (7). The specific change we consider is a having all brands face the $\{\gamma_{gh}\}$ matrix corresponding to the highest target number of impressions—i.e., every brand faces the same spillovers as the largest advertisers in the data, which are substantially higher because these advertisers cannot simply cherry pick a few shows with highly concentrated viewerships. Under this scenario, we find that 7 percent of brands then face a binding constraint, meaning they would purchase negative impressions in one or more demographic groups if they could. However, these newly constrained brands' returns to advertising are hardly affected, falling by an average of less than 1 percent.

Because our estimation approach assumes that observed brands' advertising profiles reflect *unconstrained* decisions, for observed brands the value of relaxing the spillover constraints is zero by assumption. We therefore consider a second approach, which is to draw a simulated set of brands from the estimated distributions of β and α, and then evaluate the profit consequences of changes to the spillovers for these simulated brands. Among 15,000 brands that we simulated, we found that over a third faced at least one binding constraint under the estimated spillovers $\{\gamma_{gh}\}$. This likely reflects two issues. One is simply that our functional forms for the distributions of β and α allow for some extreme values to be drawn. The second is that our current model treats β and α as independent, so some simulated brands have high β's along with very skewed α vectors, even though this is rarely true of the brands observed in the data. Since it is more difficult for large brands to concentrate their spending—they have to move farther down the list of concentrated TV programs to achieve the desired level of impressions in their target group—this aspect of our simulation likely overstates how common it is for the spillover constraints to be binding. In future drafts we therefore intend to allow for dependence between β and α.

With these caveats, we can report that for simulated brands a complete removal of spillover constraints—that is, a counterfactual world in which a brand could purchase 100% of its impressions in one demographic group if it wanted to—has only a very small effect on returns to advertising, increasing them by an average of less than half a percent.

These estimates are quite preliminary, and we suspect they will get larger as we refine various aspects of our estimation procedure. For now, however, we note that small effects of targeting are plausible in this context. Brands that choose to advertise on television may do so precisely because they want to target relatively broad audiences. Alternatively, diminishing returns to advertising could make it so that targeting all ads to a particular group is not very profitable even when it is

[13] Recall that, given our estimated value of ϕ, equations (9) and (10) allow us to directly solve for a given brand's β and α_g's.

possible.

5 Conclusion

This is a preliminary draft, and it is too early to draw strong conclusions. However, the key takeaways from the analysis conducted thus far include:

- Brands' advertising profiles indicate substantial differences in their target audiences.
- The composition of TV program viewership varies substantially across programs, but in general it is difficult for a brand to purchase advertising impressions entirely in one narrow demographic. For instance, even a brand that only wants to reach females will still inevitably spend a meaningful share of its ad budget on male impressions.
- Advertising costs vary across demographic groups. For example, male impressions are on average more expensive than female impressions, and high-income impressions are more expensive than low-income impressions.
- Using a framework that assumes brands allocate their advertising expenditures to maximize returns, we can use estimated group-specific prices combined with data on brands' purchased impressions to estimate brands' "preferences" over different groups. These estimates can then be used to infer how brands' returns to advertising would be affected by counterfactual changes in the ability to target.

References

Atchison, J., and Sheng M. Shen. "Logistic-normal distributions: Some properties and uses." *Biometrika* 67, no. 2 (1980): 261-272.

Athey, Susan, and Joshua S. Gans. "The impact of targeting technology on advertising markets and media competition." *American Economic Review* 100, no. 2 (2010): 608-13.

Bergemann, Dirk, and Alessandro Bonatti. "Targeting in advertising markets: implications for offline versus online media." *The RAND Journal of Economics* 42, no. 3 (2011): 417-443.

Bergemann, Dirk, and Alessandro Bonatti. "Selling cookies." *American Economic Journal: Microeconomics* 7, no. 3 (2015): 259-94.

Gal-Or, Esther, and Anthony Dukes. "Minimum differentiation in commercial media markets." *Journal of Economics & Management Strategy* 12, no. 3 (2003): 291-325.

Goldfarb, Avi, and Catherine Tucker. "Online display advertising: Targeting and obtrusiveness." *Marketing Science* 30, no. 3 (2011): 389-404.

Goldfarb, Avi, and Catherine E. Tucker. "Privacy regulation and online advertising." *Management science* 57, no. 1 (2011): 57-71.

Hansen, Lars Peter. "Large sample properties of generalized method of moments estimators." *Econometrica* (1982): 1029-1054.

Johnson, Justin P. "Targeted advertising and advertising avoidance." *The RAND Journal of Economics* 44, no. 1 (2013): 128-144.