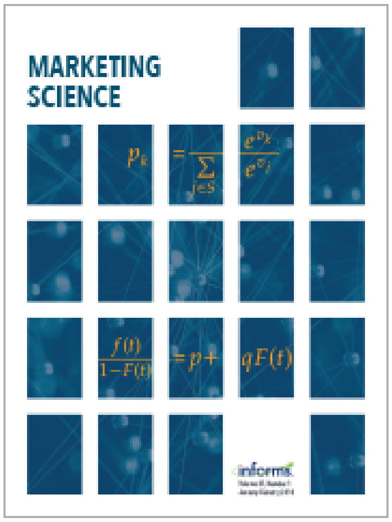


Marketing Science

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

Informational and Noninformational Advertising Content

Yi-Lin Tsai, Elisabeth Honka

To cite this article:
Yi-Lin Tsai, Elisabeth Honka (2021) Informational and Noninformational Advertising Content. Marketing Science

Published online in Articles in Advance 28 Sep 2021

. https://doi.org/10.1287/mksc.2021.1291 http://pubsonline.informs.org/journal/mksc

MARKETING SCIENCE
Articles in Advance, pp. 1–29
ISSN 0732-2399 (print), ISSN 1526-548X (online)

Informational and Noninformational Advertising Content

Yi-Lin Tsai,[a] Elisabeth Honka[b]

[a] Alfred Lerner College of Business and Economics, University of Delaware, Newark, Delaware 19716; [b] Anderson School of Management, University of California, Los Angeles, Los Angeles, California 90024

Contact: yilint@udel.edu, https://orcid.org/0000-0001-6834-9004 (Y-LT); elisabeth.honka@anderson.ucla.edu, https://orcid.org/0000-0003-0435-4220 (EH)

Received: August 15, 2018
Revised: October 11, 2019; August 3, 2020; November 4, 2020
Accepted: November 23, 2020
Published Online in Articles in Advance: September 28, 2021

https://doi.org/10.1287/mksc.2021.1291

Abstract. We investigate the relationship between both advertising content and quantity and several stages of consumers' decision making, namely, unaided and aided awareness, consideration, and purchase. Understanding how the amount and content of advertisements affect consumers' decision making is crucial for companies to effectively and efficiently use their advertising budgets. Spanning a time period from 2010 to 2016, we combine a unique data set on TV advertising content and quantities with individual-level data containing information on purchases, consideration and awareness sets, demographic variables, and perceived prices. Our results reveal that advertising quantity significantly increases consumer (unaided and aided) awareness but has no effect on conditional consideration and conditional purchase. However, when investigating the relationship between different types of advertising content and purchase stages, we find a more nuanced set of results: advertising only containing noninformational content increases unaided awareness, whereas advertising only containing informational content increases aided awareness. Advertising with both informational and noninformational content affects shoppers' but not nonshoppers' awareness and the awareness of other groups of involved consumers.

History: Avi Goldfarb served as the senior editor for this article. Anja Lambrecht, Associate Editor.
Funding: This work was supported by the Marketing Science Institute [Grant 4000161] and University of Delaware General University Research [Grant 15A00687].
Supplemental Material: Data and the web appendix are available at https://doi.org/10.1287/mksc.2021.1291.

Keywords: advertising • advertising content • purchase funnel • auto insurance industry

1. Introduction

Companies spend millions, and in some cases even billions, of dollars to advertise their products to consumers. Given the large amounts spent on advertising, the question of how well the money is spent, that is, how effective advertising spending is in influencing consumers' purchase behavior, is obviously a very important one to marketing managers. Numerous studies in both marketing and economics have provided answers to this question for various products and industries.[1] However, another crucial but much less investigated question is whether the content of ads matters, that is, whether the effects of advertising vary across different types of content or whether it only matters *that* companies advertise? Little is known about the answer to this key question.

Marketing managers (often together with an advertising agency) must decide what to communicate to consumers in their ads. On the one hand, product-related content might be more useful to consumers during the purchase process as it informs them about relevant product attributes. On the other hand, given the abundance of ads consumers encounter every day, ads must catch consumers' attention and be memorable to be effective. It is an empirical question which advertising content (or a combination of different advertising contents) affects consumers. Furthermore, the effects of advertising content might vary depending on consumers' stage in the purchase process. For example, early on, attention-catching and memorable content might be more important, whereas later, closer to the actual purchase, product-related information might be more relevant. It is critical for marketing managers to know how advertising content affects each stage of consumers' purchase process so that they can choose the advertising content that is most effective in achieving their marketing goal.

Much of the difficulty in quantifying the effects of advertising content stems from a lack of data on advertising content especially for nondigital advertising, for example, TV advertising or direct mail. Previous research on nondigital advertising content has therefore mostly investigated a small number of easily identifiable or manipulatable aspects of content, such as mention of a competitor, photo of an attractive woman, or a call for action (Bertrand et al. 2010, Liaukonyte et al. 2015, Anderson et al. 2016). In this paper, we set out to conduct a systematic and comprehensive empirical analysis of the effects of the main aspects of advertising content on consumers' purchase process.

Because of the unavailability of canned data on advertising content, we acquired data on TV advertising quantities and the corresponding creatives, that is, files containing the TV commercials, for the U.S. auto insurance industry between 2010 and 2016 from Kantar. We then hired a team of research assistants who watched the ads and recorded the content. More specifically, research assistants recorded the presence or absence of content on prices, nonprice product features, brand name focus,[2] and content with emotional appeal. Following previous literature (e.g., Resnik and Stern 1977, Lee et al. 2018), we classify price-related and (nonprice) product feature-related content as "informational," that is, transmitting information about the product and its characteristics to consumers, and brand name focused and emotionally appealing content as "noninformational," that is, not containing information about the product and its characteristics. Next, we classify all ads into one of three types: advertising only containing noninformational content, advertising only containing informational content, and advertising containing both informational and noninformational content.

We then use these data on advertising quantities and advertising content to investigate how both factors influence each stage of consumers' purchase process, that is, awareness, consideration, and purchase.[3] Our data on consumers' purchase process come from J.D. Power and Associates' annual screener surveys and annual "Insurance Shopping Studies" conducted between 2010 and 2016. These surveys provide us with information on shoppers' and nonshoppers' unaided and aided awareness sets, consideration sets, and purchase decisions.[4] Additionally, we also know the survey and shopping months, have location and demographic information, information on the identity of the previous insurance provider, and categorical information on insurance premia.[5]

We employ a set of linear probability models in our empirical analysis utilizing two approaches: a basic specification that consists of fixed effects regressions, controls for observables, and only uses within-variation and the border strategy suggested by Shapiro (2018). The basic specification addresses endogeneity concerns due to targeting, global unobservables, and time-invariant, brand-specific local unobservables.[6] The border strategy addresses endogeneity concerns due to targeting, global unobservables, and time-varying, brand-specific local unobservables, that is, a broader range of unobservables. However, the border strategy can only be empirically implemented for consumers living at designated market area (DMA) borders, that is, a subsample of the available data; it only uses local variation to identify the effects of advertising, whereas the basic specification can be estimated using data on consumers living anywhere in DMAs and uses total, that is, local and national, variation in advertising.

We discuss, in depth, the assumptions that are necessary for a causal interpretation of the advertising quantity and advertising content results and provide empirical evidence for their validity where possible. Because the analysis of advertising content requires an additional assumption for causal interpretation and this additional assumption is not empirically verifiable in our data, we interpret our results for advertising content as correlational. Although our data do not allow for a fully conclusive result on causality, our results are in line with causal patterns found by previous literature. We also compare the results from the basic specification and the border strategy and present findings supporting the conclusion that it is important to address endogeneity concerns due to time-varying, brand-specific local unobservables in our empirical context. And lastly, we show evidence for the interpretation that the advertising effects found using the border strategy and only consumers who live at DMA borders can be generalized to the whole population for the auto insurance industry.

Our results reveal that advertising intensity affects consumers' (unaided and aided) awareness, but has insignificant effects on conditional consideration and conditional purchase. These findings are consistent with prior literature (e.g., Honka et al. 2017). However, when estimating the separate effects of different types of advertising content, that is, ads with only informational content, ads with only noninformational content, and ads with both informational and noninformational content, we find a more nuanced set of results: advertising only containing noninformational content increases unaided awareness, whereas advertising only containing informational content increases aided awareness. We do not find significant effects of advertising containing both informational and noninformational content. Next, because many companies spend a significant portion of their budgets on advertising with both informational and noninformational content, we investigate whether this type of advertising significantly affects certain groups of consumers. We find it to increase shoppers' unaided awareness and the awareness of other groups of relatively involved consumers, such high-risk consumers or consumers with a change in their family or policy circumstances.

The contribution of this paper is two-fold: First, we provide new insights on advertising content. Because of very limited data availability, systematic, large-scale research especially on nondigital advertising content is scarce. We overcome this challenge by creating our own data set containing information on the main aspects of advertising content for the auto insurance industry for a time period of seven years. And second, we estimate the effects of advertising on each stage of the consumers' purchase process.

Understanding how both the amount and content of advertisements affect each stage of consumers' decision making is crucial for companies to effectively and efficiently use their advertising budgets. We show that advertising primarily affects consumer awareness and that advertising content matters. Our results contribute to managers' and researchers' knowledge of how advertising influences consumers.

The remainder of this paper is organized as follows: In the next section, we review the relevant literature. In Section 3, we discuss advertising endogeneity concerns and our identification approach. We describe our data in the following section. In Section 5, we introduce our models and estimation approach. In the following two sections, we discuss our results and present robustness checks. In Section 8, we examine limitations of our work and opportunities for future research. And, finally, we conclude in Section 9.

2. Relevant Literature

Our paper is related to three streams of literature on advertising, consumers' limited information, and demand for financial services. In the following, we review the relevant literature and delineate the positioning of our research vis-à-vis the findings from extant research.

Empirical researchers have long tried to determine the role(s) advertising plays in consumers' decision making. Most work has focused on finding empirical evidence for the informative or persuasive view of advertising first developed by Chamberlin (1933) (e.g., Ackerberg 2001, Ackerberg 2003, Narayanan et al. 2005, Ching and Ishihara 2012, Chan et al. 2013, Lovett and Staelin 2016).[7] Our focus is on financial services and, more specifically, on auto insurance. There is little academic research that investigates the precise way through which advertising affects consumer demand for financial products. Gurun et al. (2016) and Hastings et al. (2017) explore the effects of advertising in the subprime mortgage and social security markets, respectively; but neither of these studies can distinguish whether advertising affects awareness and/or consideration/purchase because of data limitations.

Most closely related to our paper is Honka et al. (2017) who investigate the role of advertising in the retail banking industry. However, our paper differs from theirs in several respects: First, the questions both papers can and do answer are different. Honka et al. (2017) find that advertising plays a primarily informative role by informing consumers about the existence of banks. They use their results to quantify branch-advertising substitutability and to analyze the competitive effects of advertising in the retail banking industry. Although we also study whether advertising affects awareness and/or consideration/purchase in the auto insurance industry, we focus on investigating the relationship between different types of advertising content and the stages of consumers' purchase process. Further, we study whether advertising content has heterogeneous effects across different consumer groups. And second, to answer the respective research questions, the empirical approaches are different. Whereas Honka et al. (2017) develop a structural model and address the issue of advertising endogeneity using the control function approach, we use reduced-form modeling and the regression discontinuity approach to address advertising endogeneity.[8]

The majority of the empirical literature on advertising content investigates the effects of specific informational cues on consumers' purchase decisions (e.g., Bertrand et al. 2010, Liaukonyte et al. 2015, Tucker 2015, Anderson et al. 2016, Sahni et al. 2018). For example, Bertrand et al. (2010) conduct a direct mail field experiment and find that showing fewer sample loans or including a photo of an attractive woman increases the demand for loans. They conclude that advertising content persuades by appealing "peripherally" to intuition rather than to reason. Liaukonyte et al. (2015) are closest to our paper in that they also investigate four content pieces, albeit different ones (action focus, information focus, emotion focus, imagery focus).

Zooming in on financial services, there is a handful of papers that investigate how different types of advertising content (together with advertising quantity) affect consumer demand in this industry. Using data from Sweden, Cronqvist (2006) finds that only a small fraction of advertisements for funds is informational in the sense that the ads contain information on relevant product characteristics. Nevertheless, he finds that advertising affects investors' choices even though it provides little information. Agarwal and Ambrose (2018) use data on home equity credit choices from direct mail and walk-in customers and find noninformational content to influence consumer choices. Gurun et al. (2016) analyze consumers' borrowing behavior in the context of subprime mortgages. They find that initial/introductory rates are frequently and prominently advertised, whereas reset rates and other characteristics of mortgages or lenders are rarely advertised. Further, Gurun et al. (2016) show that expensiveness and advertising intensity of a lender within a market are positively correlated and conclude that their results are consistent with the persuasive view of advertising. And lastly, Mullainathan et al. (2008) investigate whether predictions from their theoretical model of the role of advertising in the mutual funds industry are consistent with empirical patterns. They analyze the content of ads from two business magazines and find that the inclusion of past returns data are used to frame mutual fund investing as grabbing an opportunity rather than as hiring advice. The results from these four papers are broadly consistent with a persuasive role of advertising, that is, advertising mostly not containing information on product characteristics but nevertheless affecting consumer choice. However, what these four papers implicitly assume is that consumers have full information in the sense that they know that all these financial institutions operate in the marketplace. Whereas we use data on advertising content and quantity as do the previous papers, what distinguishes our paper from theirs is that we have information on consumers' awareness and consideration sets allowing us to relax the full information assumption made by previous literature.

3. Advertising Endogeneity and Identification

Advertising endogeneity is a well-known concern when estimating the effects of advertising on demand (see, e.g., Allenby and Rossi 2019). Its cause is omitted variables, that is, variables that are not observed in the data but are correlated with brands' advertising decisions (see, e.g., Angrist and Pischke 2009). Advertising endogeneity is a concern for the estimation of effects of both advertising quantity and advertising content. In this paper, our empirical strategy to tackle endogeneity concerns makes use of our unique and rich data: we estimate linear probability models with a large number of fixed effects (see, e.g., Wooldridge 2001) and also utilize the border strategy (Shapiro 2018). The idea behind the fixed effects approach is to partition the variation in advertising into that which is "clean" and that which is not and only use the clean portion of variation to estimate the effect of advertising (see Allenby and Rossi 2019).

Before we delve into the examination of endogeneity, we quickly introduce our four advertising measures to simplify the subsequent discussion: advertising quantity is operationalized as the logarithm of total, that is, national and DMA-level, TV advertising spending per household by brand, DMA, and month. This operationalization serves as a gross rating point approximation (Shapiro 2018). Our three measures of advertising content, that is, ad types, are the logarithms of total TV advertising spending per household by brand, DMA, and month on (i) ads with only informational content, (ii) ads with only noninformational content, and (iii) ads with both informational and noninformational content.

We discuss two specifications in the following. The first specification (basic specification) tackles the following types of endogeneity: targeting (based on observables); global unobservables; and time-invariant, brand-specific local unobservables. The second specification (border strategy) additionally addresses time-varying, brand-specific local unobservables.

3.1. Basic Specification

3.1.1. Targeting Based on Observables. Many brands target specific groups of consumers with their advertising. For example, a brand might want to advertise more to younger consumers and communicate to them that the brand (product) is priced inexpensively. That is, targeting influences brands' advertising decisions regarding both advertising quantity and advertising content. Brands' targeting rules are a classic example of unobservables in marketing—we do not observe them in our data either. If we were to observe brands' targeting rules and condition on them in the empirical analysis, the endogeneity concerns related to targeting would be resolved given the conditional independence assumption (e.g., Cameron and Trivedi 2005).

As mentioned, although we do not observe brands' targeting rules in our data, we observe a large number of consumer demographics (age, gender, education, income, past accidents, shopping for other insurance products, time spent online). We define different groups of consumers based on demographics and estimate fixed effects for these groups.[9] As long as brands' targeting rules are based on these demographics, the demographic fixed effects address endogeneity concerns related to targeting (Wooldridge 2001). However, the demographic fixed effects cannot address all endogeneity concerns related to targeting. For example, they cannot address endogeneity concerns due to targeting based on expected higher lift.

More specifically, in our basic specification, we estimate brand-demographics-year and online-brand-demographics-year fixed effects. *Online* is a dummy variable that indicates whether a consumer spent an above-median amount of time online (compared with other individuals in our data). First, note that we estimate a different set of demographic fixed effects for each brand, that is, our estimation approach allows for brands to target different demographic groups of consumers. Second, note that we estimate a different set of fixed effects for each calendar year, that is, our estimation approach allows for brands to adjust the demographic groups they target on an annual basis. Third, we also estimate separate sets of fixed effects for consumers who spend a lot of time online, that is, our estimation approach allows for consumers who spend a lot of time online to be exposed more to online advertising (and potentially differently affected by it) compared with consumers who do not spend a lot of time online. We conclude that by including brand-demographics-year and online-brand-demographics-year fixed effects, our base model is immune to targeting based on observable demographics.

3.1.2. Time-Invariant, Brand-Specific Local Unobservables. Brands might also make advertising decisions based on unobservables at the brand-market level

Table 1. Summary of Combinations of Advertising Types

Brand	A	B	C	A and B	A and C	B and C	A, B, and C
21st Century			✓				
AAA						✓	
Allstate							✓
American Family		✓				✓	
Amica Mutual			✓		✓	✓	✓
Auto Owners		✓				✓	
Erie	✓						
Esurance					✓	✓	✓
Farmers		✓				✓	✓
Geico			✓		✓	✓	
Hartford			✓			✓	✓
Liberty Mutual			✓			✓	✓
Mercury			✓		✓	✓	
MetLife			✓			✓	
Nationwide		✓	✓		✓		✓
Progressive			✓			✓	✓
Safeco		✓					
State Farm						✓	✓
Travelers		✓	✓			✓	
USAA		✓					

Note. ✓, Combination of advertising types employed in at least one brand-year; A, advertising with only informational content; B, advertising with only noninformational content; C, advertising with both informational and noninformational content.

leading to endogeneity concerns. For example, a brand might advertise more in DMAs close to its headquarters and also emphasize its local roots in the ads. We can address endogeneity concerns stemming from such time-invariant, brand-specific local unobservables by incorporating brand-DMA fixed effects (see, e.g., Cunningham 2020).

3.1.3. Global Unobservables. Note that brand fixed effects, year fixed effects, and brand-year fixed effects are subsumed in the brand-demographics-year and online-brand-demographics-year fixed effects. That is, any omitted variables that are global, that is, specific to the brand, to the year, and to the brand-year, are conditioned on through the fixed effects. To keep the discussion simple, in the remainder of this subsection, we only refer to brand fixed effects even though we technically estimate brand-demographics-year and online-brand-demographics-year fixed effects in the empirical analysis.

Suppose brand X always spends, on average, three times more on advertising than brand Y during the study period. Brand fixed effects control for that difference in average advertising quantity levels between the two brands and only within-brand variation is used to estimate the effect of advertising quantity. To put it differently, the within-brand variation in advertising quantity is assumed to be clean allowing for a causal interpretation of the coefficient estimate as the average treatment effect (ATE; see, e.g., Allenby and Rossi 2019).

For the analysis of advertising content, we similarly assume that the within-brand variation in spending on a type of advertising is clean and use it in the estimation. However, for the analysis of advertising content, a second issue arises: not all brands employ all three types of advertising in their ads during the study period. That is, some brands use all three types of advertising, whereas others only use one or two types of advertising and the employed types of advertising in some cases also vary over time. We show the combinations of advertising types each brand utilizes in Table 1. The letters A, B, and C abbreviate each of the three advertising types (A, only informational; B, only noninformational; C, both informational and noninformational) and a checkmark symbolizes that a specific combination of advertising types was used by a brand during at least one brand-year, that is, a brand can have more than one checkmark if it utilized different combinations of advertising types in different years.[10]

Ideally, we would like all brands to only have a checkmark in the column "A, B, and C" (or the utilized advertising type(s) be randomly distributed). The reality presented in Table 1 is that nine brands employ all three advertising types simultaneously in at least one year, however, not necessarily in all years. A tenth brand employs all three advertising types during the study period but not simultaneously in one year. On the other end of the spectrum, four brands (21st Century, Erie, Safeco, and USAA) only employ one type of advertising throughout the whole study period. Although there are more brands that employ one type of advertising during at least one year, these brands also employ at least one different advertising type in other years. Note that 21st Century, Erie, Safeco, and USAA are relatively small insurance brands. Their joint spending on advertising during the study period only represents 1% of advertising spending among the 21 brands under study.

It is an open question why some brands only use one or two types of advertising, whereas others use all three types of advertising and why the employed advertising types vary over time in some cases. In a recent working paper, Honka and Tsai (2019) examine 17 types of advertising content from the auto insurance industry over a time period of 15 years. Consistent with the picture presented in Table 1, using multivariate analysis of variance (MANOVA), they find that most of the variation in advertising content is driven by brand followed by time and media channel. Furthermore, Honka and Tsai (2019) also find that brands react to environmental conditions. For example, brands included more price-related information in ads during the Great Recession. During the 2000s, brands employed relatively more fear-inducing content (warning consumers not to buy auto insurance over the internet), whereas during the 2010s brands employed more funny/entertaining content.

We also investigated whether changes in advertising content over time are related to changes in advertising agencies (lead creative agencies) and found little evidence for it. Table A-1 in Web Appendix A displays information on insurance brands' lead creative agencies (and changes thereof) during the study period for the brands for which we were able to find this information. We mostly miss information on creative agencies for smaller insurance brands that do not spend a lot on TV advertising. Overall, relationships between insurance brands and creative agencies are long lived. In Table 1, we see that 14 out of the 20 brands changed the advertising type or combination of advertising types at least once during the study period. However, we only observe three changes in lead creative agencies: American Family changed its agency in 2014, and there were no changes in employed advertising types following the switch. Liberty Mutual changed its agency in 2014, and the mix of employed advertising types changed starting in 2015. Travelers changed its agency in 2015. However, starting in 2014, Travelers had stopped advertising auto insurance on TV. Thus, we conclude that we do not find evidence that changes in advertising content are systematically related to changes in ad agencies. Note that this does not imply that brands do not change advertising campaigns. Rather, it implies that that the communicated content does not significantly change even if the delivery, that is, creative design, changes.

The concern with brands not using all types of advertising is selection, that is, brands might choose to utilize the advertising type(s) that are most effective for them. In principle, brand fixed effects control for differences in employed advertising types across brands. However, we are interested in measuring the *average* treatment effect for the auto insurance industry, that is, across *all* brands. If a brand does not use a type of advertising, there is no within-brand variation in spending on that type of advertising and the brand does not contribute to the estimation of the average effect of that type of advertising. To put it differently, the average is calculated only across brands that actually employ an advertising type. To be able to make causal claims for the effects of advertising content for *all* brands, that is, interpret the coefficient estimates as ATEs, the following additional assumption is required:

Assumption 1. *The within-brand variation in a type of advertising is the same for a brand that employs this type of advertising and a brand that does not (if it were to use it).*

That is, we do not employ an econometric technique to address this potential selection concern. Assumption 1 holds if brands do not base their decisions which advertising content to communicate on its effectiveness in affecting consumers' immediate purchase decisions. This might be the case if advertising content is chosen based on other considerations, such as (long-term) brand image, design updates, or production costs. Although we provide suggestive empirical evidence for the validity of Assumption 1 in Section 6.2, we are not able to provide conclusive evidence with our data.[11] To be cautious, we therefore interpret our results for advertising content as correlational. Finally, in Section 6.6, we discuss that our results for advertising content are in line with causal patterns found by previous literature.

3.1.4. Identification and Interpretation of Basic Specification. If the fixed effects and controls discussed in Sections 3.1.1–3.1.3 are included in the empirical analysis, the estimation is immune to endogeneity stemming from targeting; global unobservables; and time-invariant, brand-specific local unobservables. If these are the only sources of endogeneity the researcher is concerned about, the estimated effect of advertising quantity is causal (see, e.g., Angrist and Pischke 2009). For a causal interpretation of the advertising content results, Assumption 1 must additionally hold. If Assumption 1 does not hold for an analysis of advertising content, the estimated effects should be interpreted as correlational. Because we are not able to provide conclusive evidence for the validity of Assumption 1 in our data, we will interpret the empirical results for advertising content as correlational.

The effects of advertising are identified by variation in total advertising, that is, both variation in national and variation in local advertising contribute to the identification of the advertising effects. In addition, through its functional form, the logarithmic operationalization of the advertising variables also contributes to the identification of the advertising effects.

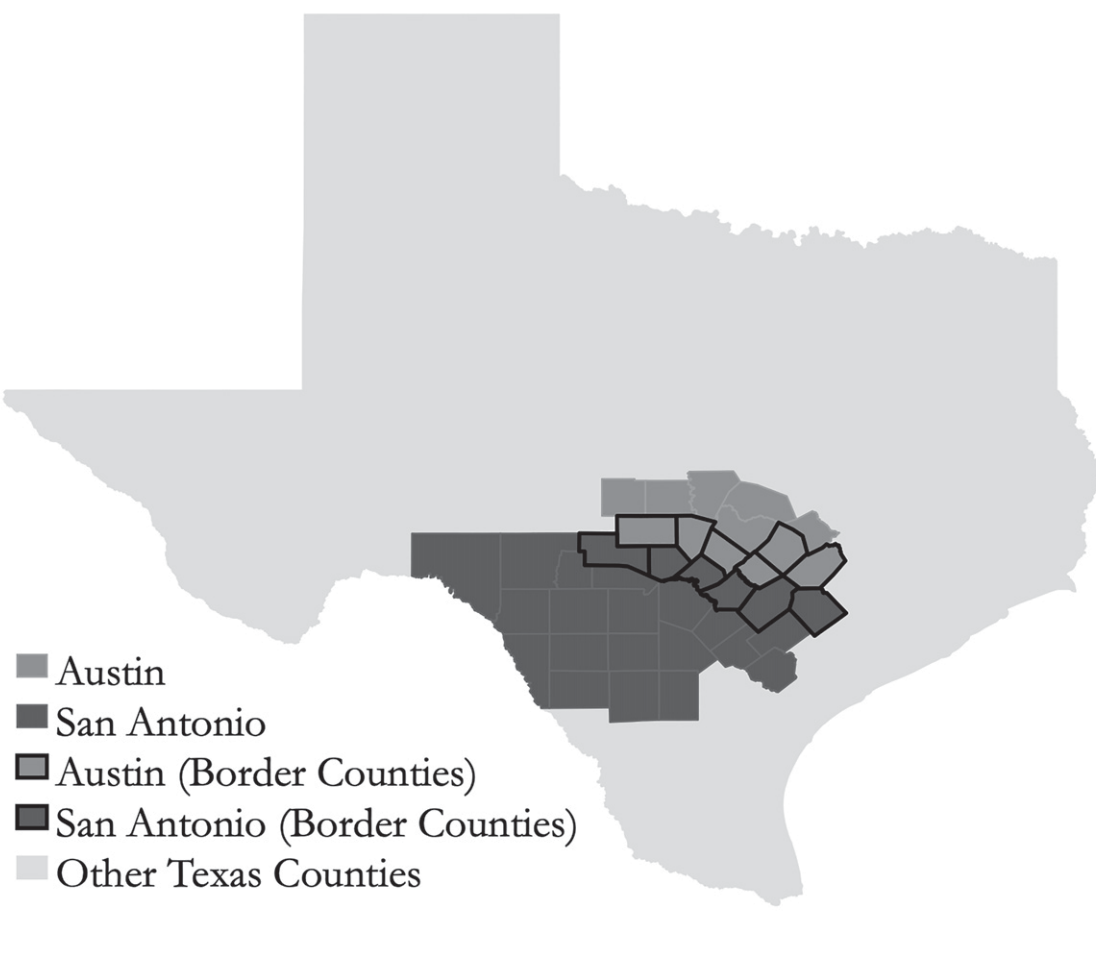
To summarize, the basic specification is immune to targeting based on demographics, global unobservables and endogeneity due to time-invariant, brand-specific local unobservables. Only within-brand-demographic-year, within-online-brand-demographic-year, and within-brand-DMA variation is used to identify the effects of advertising quantity and advertising content. The identifying variation stems from variation in total advertising, that is, both variation in national and variation in local advertising contribute to the identification of the advertising effects. The effects of advertising quantity can be interpreted as ATEs if the researcher is only concerned about endogeneity due to targeting, global unobservables, and endogeneity due to time-invariant, brand-specific local unobservables. For a causal interpretation of the effects of advertising content, Assumption 1 must additionally hold. Because we are not able to provide conclusive evidence for the validity of Assumption 1 in our data, we will interpret the empirical results for advertising content as correlational.

3.2. Border Strategy

In some instances, including brand-DMA fixed effects might not be enough to eliminate endogeneity concerns due to local omitted variables. This is the case when brands make advertising decisions based on unobservables at the brand-market level that vary over time. Such time-varying local "events" include sponsorships of local sports teams or festivals and changes in the focal brand's or competitive brands' local agent network (i.e., openings and closings of agencies). They can also include changes in the communicated content. For example, sponsorships usually only show the brand name and a brand might place more price-related ads because of the opening of a competitor's new local office.

Shapiro (2018) suggests using the regression discontinuity approach to address endogeneity concerns due to time-varying, brand-specific local unobservables.[12] We start by briefly describing the main idea of the regression discontinuity approach using advertising quantity. Figure 1 shows an example of two DMAs in Texas—Austin and San Antonio (in light grey and dark grey, respectively). Note that the border between these two DMAs does not—as they do not for most DMAs—coincide with state borders. Rather, historically, DMAs were centered around a large city or a metropolitan area. The border strategy to deal with

Figure 1. Example of Border Strategy: Austin and San Antonio DMAs



advertising endogeneity considers the six counties directly adjacent to the DMA border belonging to the Austin DMA (in light grey) and the six counties directly adjacent to the DMA border belonging to the San Antonio DMA (in dark grey) as two treatment groups in every month. Although consumers living on different sides of the DMA border are similar, they are being treated with different amounts of advertising. The advertising effect can be identified by comparing how consumers living in the two groups of border counties react differently to differences in advertising quantities. Two types of fixed effects are crucial for the implementation of the regression discontinuity approach: brand-border-DMA and brand-border-month fixed effects. The former controls for persistent differences across different border regions; the latter captures unobserved, time-varying brand-border-region-specific variables.

The intuition for the identification of the advertising content effects is the same as the intuition for the identification of the advertising quantity effect using the regression discontinuity approach: consumers living on different sides of a DMA border are similar but being treated with different amounts of a specific advertising type.[13] Therefore, the effect of that specific advertising type can be identified by comparing how consumers living in the two groups of border counties react differently to differences in spending on that advertising type.

The border strategy requires the following assumption to hold for a causal interpretation of the results:

Assumption 2. *Individuals on both sides of a DMA border are similar.*

Recall that we control for differences in demographics by estimating brand-demographics-year and online-brand-demographics-year fixed effects. Thus, in our empirical analysis, we assume that individuals on both sides of a DMA border are similar *after* controlling for demographics. If the researcher is concerned about targeting, global unobservables, and time-varying, brand-specific local unobservables and Assumption 2 holds (Assumptions 1 and 2 hold), the estimated effects of advertising quantity (advertising content) using the border strategy are causal. Otherwise, the effects should be interpreted as correlational. As pointed out in Section 3.1.3, because we are not able to provide conclusive evidence for the validity of Assumption 1, we will interpret our results for advertising content as correlational.

In the border strategy, the effects of advertising are identified by variation in local advertising. National advertising also contributes to the identification of the effects of advertising but only through functional form, that is, through the logarithmic operationalization of the advertising variables. Ideally, we would like to estimate the overall effect of advertising and not only of local advertising. Such an interpretation of the advertising effects estimated using the border strategy rests on the following (composite) assumption:

Assumption 3. (a) *Advertising affects consumers living in border and nonborder counties similarly;* (b) *national advertising affects consumers in a similar manner as local advertising; and* (c) *the national advertising content composition is similar to the local advertising content composition.*

With regard to Assumption 3(a), it is an open question whether consumer preferences and especially sensitivity to advertising are the same among consumers living in DMA border counties compared with all consumers living in DMAs. Thus, using the border strategy, a "local" average treatment effect is estimated and its generalizability to the whole population requires additional analyses and/or discussion (see Shapiro 2018, Tuchman 2019). We present empirical evidence for the generalizability of the advertising effects in Section 6.1. Regarding Assumption 3(b), it is important to note that insurance brands use the same creatives to advertise nationally and locally. The most common way that ads are "localized" is that a local agent's contact details are shown at the end of a commercial.[14] Further, consumers are unaware which advertising was purchased nationally versus locally. Lastly, with regard to Assumption 3(c), we calculated the correlation based on the distance covariance and find it to equal 0.79 ($p < 0.001$) indicating a strong and close relationship in the national and local content compositions (see, e.g., Josse and Holmes 2016).[15] Thus, we find evidence in support of Assumption 3 (b) and (c) and present evidence in support of Assumption 3(a) in Section 6.1.

Geographic variation, that is, across-DMA-border variation, in our advertising quantity and advertising content variables within a brand and month is crucial for estimation. To put it differently, we need discontinuities in all four advertising measures at DMA borders to be able to identify the effects of advertising. Furthermore, because our data span a time period of seven years, we also need variation in the discontinuities in the four advertising measures over time, that is, across months. We present descriptive evidence that our data on advertising quantity and advertising content contain a sufficient amount of such variation in Web Appendix C.

And lastly, although the border strategy tackles the issue of time-varying, brand-specific local unobservables, it comes at a cost: the effects of advertising can only be estimated using a subsample of the data (consumers living at the borders of DMAs). This raises power concerns that are a well-known issue in the estimation of advertising effects (Lewis and Rao 2015). We discuss whether this is an issue in our empirical analysis in Section 6.1.

To summarize, the specification discussed in this section is immune to targeting based on demographics, global unobservables, and endogeneity due to time-varying, brand-specific local unobservables. Within a brand-demographic-year and an online-brand-demographic-year, the effects of advertising are identified by how consumers living across a DMA border react differently to differences in advertising quantity and advertising content. The identifying variation stems from variation in local advertising, and national advertising contributes to the identification through functional form. If the researcher is concerned about targeting, global unobservables, and time-varying, brand-specific local unobservables and Assumption 2 holds (Assumptions 1 and 2 hold), the estimated effect of advertising quantity (advertising content) using the border strategy are causal. Otherwise, the effects should be interpreted as correlational. We will interpret the effects of advertising quantity as causal and the effects of advertising content as correlational. If Assumption 3 holds, the estimated effects using the border strategy can be interpreted as the overall effects of advertising quantity and advertising content. Otherwise, they should be interpreted as local effects.

4. Data

We combine data from two sources to investigate the relationship between advertising and each stage of the consumers' purchase process. Our data on advertising come from Kantar. Kantar tracks TV advertising expenditures (in dollars and units) at the national and DMA level. We have monthly data from 2010 to 2016. Additionally, Kantar supplied us with the creatives, that is, the files containing the TV commercials.

Our second data come from J.D. Power and Associates who generously shared data from their annual screener surveys and annual "Insurance Shopping Studies" covering consumer behavior from 2010 to 2016. The data sets contain individual-level information on consumers' awareness and consideration sets, the identity of the purchased option, the identity of the previous insurance provider, location and demographic information, survey and shopping months, perceived categorical price information for shoppers, and representativeness weights.[16]

4.1. Data Processing

4.1.1. Advertising Content. We have all ads, that is, creatives, placed by auto insurance companies on TV between 2010 and 2016 (both in Spanish and English), that is, 2,965 unique creatives across 21 auto insurance brands.

To code the content of these creatives, we hired a team of 25 student research assistants during an 18-month time period.[17] These research assistants were trained to code whether a creative (i) talked about prices/rates/discounts, (ii) conveyed (nonprice) product feature information, (iii) focused on the brand name, (iv) had emotional appeal (i.e., humorous/funny/entertaining and/or fear-inducing). We developed these four content types based on previous literature (e.g., Resnik and Stern 1977, Stern et al. 1981) and taking the characteristics of the auto insurance industry into account. A detailed description of each content type with examples is shown in Web Appendix B. Note that creatives can contain more than one piece of content, for example, price-related and emotionally appealing content.

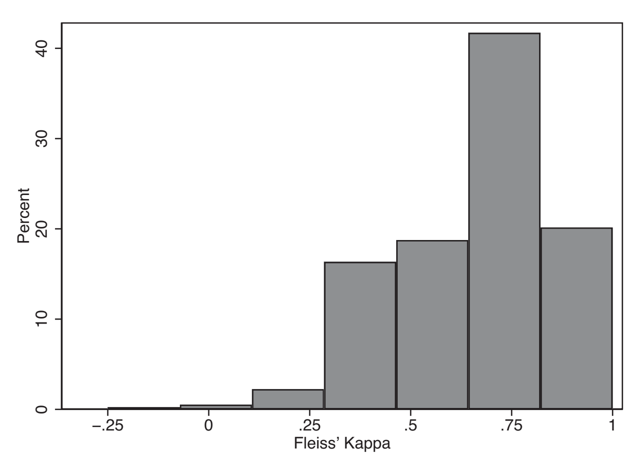
The training of the research assistants was conducted as follows: first, research assistants were screened for language skills (English and Spanish) and basic knowledge of the auto insurance market before employment. Then research assistants received a document containing a written description of each content type and a set of 20 creatives that they coded on their own. Afterward, they met with one of the authors to discuss their coding decisions and to resolve other uncertainties. After this meeting, research assistants started coding creatives.

Each creative was independently coded by at least three research assistants, and we use majority coding across the three research assistants for each creative in the analysis. Fleiss' kappa is a measure of interrater agreement in coding. Figure 2 shows a histogram of Fleiss' kappas for the coded creatives across the three research assistants. The average value is 0.68 with a median of 0.70 indicating substantial agreement.

To ensure that only reliably coded ads are employed, we removed all creatives with Fleiss' kappa smaller than 0.4 and control for spending on ads with missing content in the empirical analysis.

4.1.2. Screener Surveys and Insurance Shopping Studies. The screener surveys conducted by J.D. Power and Associates between 2010 and 2016 provide

Figure 2. Histogram of Fleiss' Kappas (Median = 0.70 and Mean = 0.68)



information on a large number of nonshoppers (on average, 15,000 individuals annually). Nonshoppers are consumers who were not actively shopping for auto insurance during a particular year. From these screener surveys, we have information on nonshoppers' unaided and aided awareness sets and their current insurance provider.[18] Each year, J.D. Power and Associates also conduct Insurance Shopping Studies surveying about 10,000 individuals annually. From these Insurance Shopping Studies, we have information on shoppers' unaided and aided awareness sets, consideration sets, and purchase decisions. Additionally, we also have location and demographic information for all consumers, that is, shoppers and nonshoppers, survey and shopping months, and information on the identity of the previous insurance provider. And lastly, for shoppers, we also have categorical information on insurance premia. It is important to note that both the screener surveys and the Insurance Shopping Studies contain repeated cross-sections of consumers and not a panel of consumers.

The original data contain information on 360,182 individuals with valid Federal Information Processing Standard (FIPS) codes (108,942 shoppers and 251,240 nonshoppers). Unfortunately, detailed location information (beyond the state) was not available for respondents from the 2011 Insurance Shopping Study and the 2014 screener survey so these respondents were dropped. In our empirical analysis, we restrict the sample to respondents living in the top 130 DMAs. Furthermore, we focus on the top 21 brands (measured by revenue) that were consistently part of the surveys from 2010 to 2016 and held a joint market share of about 85%. This focus implies that respondents who were not aware of any of the top 21 brands were removed. These data cleaning steps left us with 357,365 consumers (108,905 shoppers and 248,460 nonshoppers).

Next, we dropped respondents who (i) indicated to be younger than 18 years or older than 75 years, (ii) reported an annual income of over $1,000,000, (iii) stated to own more than four cars, (iv) reported paying a premium of more than $4,000 for a six-month policy, (v) were not a decision-maker regarding the auto insurance purchase, (vi) inconsistently reported their location, and (vii) did not provide demographic information. These data cleaning steps left us with 256,950 consumers (82,823 shoppers and 174,127 nonshoppers). And lastly, for the Insurance Shopping Studies, to avoid any memory issues consumers might develop over time, we restrict our data to consumers who completed the survey two or fewer months after shopping for car insurance. This step left us with our final sample of 197,267 respondents (23,140 shoppers and 174,127 nonshoppers). We reweigh the individuals in our final sample using representativeness weights. The reweighted final sample contains 197,267 consumers (61,477 shoppers and 135,790 nonshoppers). Throughout this paper, we refer to this sample as the "Complete DMA Sample."

In our empirical analysis using the border strategy, we focus on respondents living in counties at the borders of the top 130 DMAs excluding the Bakersfield, California and San Diego, California DMAs. We excluded the Bakersfield, California and San Diego, California DMAs because, in both cases, the whole DMA only contains one county and, therefore, a border could not be defined. Focusing on border counties left us with our final border sample of 77,309 respondents (67,583 nonshoppers and 9,726 shoppers) located in 1,263 different counties. These 1,263 counties belong to 250 different border regions, that is, a cluster of geographically adjacent counties spanning across both sides of a DMA border. We reweigh the individuals in our final sample using representativeness weights. The reweighted final sample contains 77,309 consumers (25,106 shoppers and 52,203 nonshoppers). Throughout this paper, we refer to this sample as the "DMA Border Sample."

4.2. Data Description

4.2.1. Advertising. Insurance companies can place TV advertisements nationally and locally, that is, at the DMA level.[19] On average, a single insurance brand spends about $6 million monthly on national TV advertising placing around 2,000 ads. There is large variation in national TV advertising spending ranging from $0 (Auto Owners, Erie, GMAC) to $36 million (Geico) per month. All 21 brands together spend about $123 million monthly on national TV advertising placing around 43, 000 ads. At the DMA-level, a single insurance brand spends, on average, around $2,400 per DMA and per month placing about 12 TV ads. Across all DMAs, all 21 brands spend, on average, around $11 million per month on DMA-level advertising placing about 53,120 TV ads. Eighty-nine percent of brands' TV advertising spending is utilized on national advertising, and the remaining 11% are spent on DMA-level advertising.[20] Whilst national TV advertising has increased during the observation period, DMA-level TV advertising has decreased.

Focusing on the top 130 DMAs, average total monthly advertising spending per household was $0.05 with brands such as Erie, GMAC, and Safeco spending $0 and brands such as Progressive and Geico spending $0.20 and $0.33, respectively, per household (see column (i) in Table 2). Following Shapiro (2018), we use the logarithm of total advertising spending per household, as our measure of advertising intensity in the empirical analyses. As robustness checks of our advertising measure, we also use total advertising units and DMA-level ad expenditure per household in $ as measures of advertising. Descriptive

Table 2. Monthly TV Advertising Quantities

Brand	(i) Total ad expenditure per household in $	(ii) Total ad units	(iii) Local ad expenditure per household in $
21st Century	0.0141	2,029	0.0004
AAA	0.0095	43	0.0095
Allstate	0.1230	4,703	0.0021
American Family	0.0041	13	0.0041
Amica	0.0112	244	0.0094
Auto Owners	0.0002	1	0.0002
Erie	0.0001	0	0.0001
Esurance	0.0956	3,528	0.0005
Farmers	0.0096	487	0.0014
Geico	0.3303	10,275	0.0142
GMAC	0.0000	0	0.0000
Hartford	0.0090	253	0.0015
Liberty Mutual	0.1117	5,041	0.0016
Mercury	0.0008	50	0.0003
MetLife	0.0008	26	0.0000
Nationwide	0.0553	1,374	0.0023
Progressive	0.2018	9,337	0.0182
Safeco	0.0000	1	0.0000
State Farm	0.1241	5,013	0.0042
Travelers	0.0068	258	0.0010
USAA	0.0388	742	0.0017
Average	0.0546	2,068	0.0035

statistics for these two variables are shown in columns (ii) and (iii) in Table 2.

And lastly, Table 3 depicts the percentages of TV ads (weighted by spending) for each brand that contain a specific mix of content over the whole study period. On average, 11% of ads contain only informational content; 34% of ads contain only noninformational content; and the remaining 55% contain both informational and noninformational content.

4.2.2. Consumer Shopping Behavior.

We first discuss consumer characteristics and then consumer shopping behavior. In Table 4, we display descriptive statistics for all consumers from the Complete DMA sample and the DMA border sample (columns (i) and (iv)) as well as nonshoppers (columns (ii) and (v)) and shoppers (columns (iii) and (vi)) separately. Overall, consumers from the Complete DMA sample and the DMA border sample are very similar in terms of their characteristics. The two largest differences are that the percentage of married consumers in the DMA border sample is higher by 1.66% and, not surprisingly, the percentage of consumers living in an urban area is lower by 3.37%. Because the two samples exhibit very similar characteristics, we focus on the Complete DMA sample in the following. Among all consumers, about 80% of respondents are between 25 and 65 years old, 42% are male, and 56% are married. Sixty percent of respondents have a college degree, and 25% of respondents have an annual income of more than $100K. Comparing the two subgroups of shoppers and nonshoppers (columns (ii) and (iii) in Table 4), we find shoppers to be more likely male, married, Black, and Hispanic than nonshoppers. For shoppers (only), we have additional information on insurance-related variables: 45% of shoppers were also shopping for homeowner's or renter's insurance and 7% of shoppers indicated having a poor credit history. Further, 3% and 4% of shoppers reported having had two or

Table 3. Percentages of Spending on Each Advertising Type During Study Period

Brand	(i) Informational only	(ii) Noninformational only	(iii) Both informational and noninformational
21st Century	0.00	0.00	100.00
AAA	0.04	20.10	79.86
Allstate	42.69	29.69	27.62
American Family	0.00	88.21	11.79
Amica Mutual	1.38	10.90	87.73
Auto Owners	0.00	99.33	0.67
Erie	100.00	0.00	0.00
Esurance	17.17	10.02	72.81
Farmers	1.32	69.92	28.76
Geico	0.30	0.94	98.76
Hartford	1.31	15.70	82.98
Liberty Mutual	10.52	1.21	88.27
Mercury	0.62	12.96	86.42
MetLife	0.00	0.45	99.55
Nationwide	22.85	21.46	55.68
Progressive	1.26	8.23	90.50
Safeco	0.00	100.00	0.00
State Farm	21.86	42.23	35.90
Travelers	0.00	54.32	45.68
USAA	0.00	99.73	0.27
Average	11.07	34.27	54.66

Table 4. Descriptive Statistics

	(i)	(ii)	(iii)	(iv)	(v)	(vi)
	Complete DMA sample			DMA border sample		
Demographics	All consumers	Nonshoppers	Shoppers	All consumers	Nonshoppers	Shoppers
Age ≤ 25 years	0.0472	0.0415	0.0598	0.0455	0.0395	0.0580
26 Years < age ≤ 45 years	0.3703	0.3642	0.3839	0.3598	0.3548	0.3702
46 Years < Age ≤ 65 years	0.4342	0.4409	0.4194	0.4440	0.4515	0.4284
Age > 65 years	0.1482	0.1533	0.1368	0.1507	0.1542	0.1434
Male	0.4173	0.4056	0.4433	0.4170	0.4038	0.4445
Black	0.0472	0.0408	0.0612	0.0432	0.0375	0.0550
Hispanic	0.0333	0.0236	0.0548	0.0309	0.0212	0.0511
Asian	0.0661	0.0723	0.0525	0.0633	0.0714	0.0466
Married	0.5593	0.5527	0.5739	0.5759	0.5708	0.5863
College degree	0.5994	0.6334	0.5243	0.5870	0.6239	0.5103
Income greater than $100K	0.2492	0.2536	0.2394	0.2468	0.2524	0.2351
Lived in urban area			0.1847			0.1510
Someone under 25 years insured under the policy			0.1360			0.1357
Shopped for homeowner's insurance			0.3452			0.3410
Shopped for renter's insurance			0.1046			0.0978
Shopped for Life Insurance			0.0360			0.0386
Shopped for personal umbrella insurance			0.0589			0.0568
Two or more accident(s) in the last three years			0.0314			0.0288
Two or more tickets in the last three years			0.0369			0.0385
Poor credit history			0.0704			0.0758
Same insurer as in previous year			0.5139			0.5201

more accidents and tickets, respectively, during the last three years.

Next, we discuss consumer shopping behavior for consumers in the Complete DMA sample. The statistics on consumer shopping behavior for consumers in the DMA border sample are very similar to those in the Complete DMA sample, and we show them in Web Appendix D. In the Complete DMA sample, 31% of consumers are shoppers and the remaining 69% of consumers are nonshoppers.[21] This proportion of shoppers is consistent with proportions reported by other sources: 46% of consumers reported having shopped for auto insurance during the past 12 months according to a 2015 comScore survey;[22] 25% of consumers reported having shopped for auto insurance during the past 12 months according to a 2017 Princeton Research Survey Associates International survey; 33% of consumers reported having shopped for auto insurance during the past 12 months according to the 2012 McKinsey Auto Insurance Customer Insights Research report.

Among shoppers, 49% of consumers switch their auto insurance provider after the shopping occasion under study and the remaining 51% of consumers remain with their previous insurance provider.[23] Projecting to the whole population, we find that 13% of all consumers switch their auto insurance provider in a year. The 2012 McKinsey Auto Insurance Customer Insights Research similarly report found about one-third of shoppers or 13% of the total population to switch insurance providers.

The average number of auto insurance brands consumers are aware of is 4.21 for unaided awareness and 12.43 for aided awareness. As expected, nonshoppers are aware of fewer brands than shoppers: 3.84 versus 5.04 (difference statistically significant at $p < 0.001$) for unaided awareness and 12.32 versus 12.65 (difference statistically insignificant) for aided awareness. We next turn to the brands that consumers are aware of (see Table 5). The probability that a consumer is aware of any brand is 20% (unaided) and 59% (aided) (columns (i) to (ii) in Table 5). Across all consumers, the brand-specific awareness probabilities range from 1% (GMAC) to 72% (State Farm) for unaided awareness and 9% (Auto Owners, Erie) to 96% (State Farm) for aided awareness. Further, we compare the brand-specific awareness probabilities for shoppers and nonshoppers (columns (iii) to (vi) in Table 5). Not surprisingly, shoppers have, on average, a higher probability of being aware of any brand than nonshoppers. We provide more details on consumers awareness and consideration in Web Appendix E.

Table 6 contains consideration and purchase shares for all brands as well as conversion rates for consideration, that is, conditional on being aware of a brand the proportion of consumers that consider the brand, and for purchase, that is, conditional on considering a brand the proportion of consumers that choose the brand.[24] Conditional on unaided awareness, the conversion rates to consideration vary from 49% (Farmers) to 95% (GMAC) with an average of 64%. Conditional on aided awareness, the conversion rates to consideration

Table 5. Awareness Probabilities

	(i) All consumers	(ii)	(iii) Nonshoppers	(iv)	(v) Shoppers	(vi)
Brand	Unaided	Aided	Unaided	Aided	Unaided	Aided
21st Century	0.0790	0.5046	0.0558	0.4975	0.1305	0.5184
AAA	0.2621	0.8528	0.2646	0.8728	0.2565	0.8136
Allstate	0.6427	0.9493	0.6209	0.9477	0.6909	0.9524
American Family	0.0818	0.3555	0.0743	0.3562	0.0986	0.3542
Amica Mutual	0.0240	0.1356	0.0120	0.0862	0.0506	0.2323
Auto Owners	0.0187	0.0880	0.0123	0.0734	0.0327	0.1167
Erie	0.0240	0.0878	0.0183	0.0730	0.0366	0.1166
Esurance	0.1009	0.6239	0.0591	0.5957	0.1935	0.6790
Farmers	0.4194	0.8806	0.4288	0.8950	0.3985	0.8524
Geico	0.5997	0.9482	0.5548	0.9445	0.6992	0.9555
GMAC	0.0096	0.2268	0.0033	0.2267	0.0237	0.2269
Hartford	0.0853	0.6573	0.0633	0.6557	0.1339	0.6604
Liberty Mutual	0.1338	0.7917	0.0917	0.7774	0.2270	0.8196
Mercury	0.0710	0.3060	0.0763	0.3572	0.0591	0.2057
MetLife	0.0495	0.7600	0.0354	0.7582	0.0805	0.7635
Nationwide	0.1898	0.8259	0.1579	0.8152	0.2605	0.8467
Progressive	0.4546	0.9151	0.3887	0.9051	0.6008	0.9346
Safeco	0.0444	0.3541	0.0299	0.3436	0.0765	0.3745
State Farm	0.7175	0.9635	0.7158	0.9653	0.7212	0.9600
Travelers	0.0981	0.7274	0.0739	0.7177	0.1518	0.7463
USAA	0.1037	0.4807	0.0991	0.4581	0.1138	0.5248
Average	0.2004	0.5921	0.1827	0.5868	0.2398	0.6026

vary from 10% (MetLife) to 42% (Geico) with an average of 26%. And lastly, conditional on consideration, the conversion rates to purchase range from 15% (Geico) to 62% (Auto Owners) with an average of 28%. To summarize, there is substantial variation within a purchase stage and across purchase stages both in shares and in conversion rates across brands.

5. Models and Estimation

In our main empirical analysis, we estimate a total of 24 linear probability models that vary across three dimensions (advertising, endogeneity, purchase stage).[25] Here, we aim at concisely and clearly presenting them. Because our main focus is on advertising content, we show the advertising content regressions in detail—the advertising quantity regressions only differ by the definition of the advertising variable.[26] Further, we structure this section by types of endogeneity that are addressed in a specification—the same way as Section 3.

5.1. Basic Specification

In these specifications, we address the types of endogeneity discussed in Section 3.1, that is, targeting based on observables, global unobservables, and time-invariant, brand-specific local unobservables. The models presented in this section are estimated using the Complete DMA Sample.

Let Y_{ibt} be the binary dependent variable of interest, that is, (unaided or aided) awareness, consideration or purchase indicator for consumer i, brand b, and month t. Then the (unaided or aided) awareness model is given by

$$\begin{aligned} Y_{ibt} = {} & \beta_1 \log\left(1 + A^{f}_{bd,t-1}\right) + \beta_2 \log\left(1 + A^{f,nf}_{bd,t-1}\right) \\ & + \beta_3 \log\left(1 + A^{nf}_{bd,t-1}\right) + \varrho^{D_i}_{bg} + \varphi^{D_i}_{bg}\mathbb{I}^{o}_{i} + \mu_{bd} \\ & + \zeta_{bst} + \psi_{bd,t-1} + \tau^{survey}_{g} + \epsilon_{ibt} \quad \forall\, b \neq b_{CI}, \end{aligned} \tag{1}$$

where $\log(1 + A^{f}_{bd,t-1})$, $\log(1 + A^{f,nf}_{bd,t-1})$, and $\log(1 + A^{nf}_{bd,t-1})$ capture the logarithms of total TV advertising spending per household on ads with only informational content, ads with both informational and noninformational content, and ads with only noninformational content, respectively, by brand b in DMA d. Note that consumers must be aware of the company they purchase an insurance policy from. Similar to Honka et al. (2017), we therefore exclude a consumer's current insurance provider b_{CI} from all awareness model estimations.

Moreover, we include the following fixed effects to tackle targeting and global unobservables: first, we include brand-demographics-year fixed effects, $\varrho^{D_i}_{bg}$, where g represents year. Although we do not observe other potentially targeted offline marketing activities such as direct mail, as long as the targeting is based on demographics, our brand-demographics-year fixed effects control for it. Second, we include online-brand-demographics-year fixed effects, $\varphi^{D_i}_{bg}$. The dummy variable $\mathbb{I}^{o}_{i}$ is individual specific and indicates whether a

Table 6. Consideration, Purchase, and Conversion Probabilities (Shoppers Only)

Brand	Considered	Chosen	Aware (unaided) → considered	Aware (aided) → considered	Considered → chosen
21st Century	0.1328	0.0411	0.8973	0.2530	0.3090
AAA	0.2336	0.0955	0.8295	0.2860	0.4090
Allstate	0.3408	0.0553	0.5009	0.3577	0.1623
American Family	0.2114	0.0904	0.7046	0.2538	0.4277
Amica	0.0465	0.0181	0.7894	0.1923	0.3885
Auto Owners	0.0610	0.0381	0.8549	0.3006	0.6239
Erie	0.1210	0.0736	0.8046	0.3497	0.6083
Esurance	0.1732	0.0536	0.8527	0.2534	0.3095
Farmers	0.2229	0.0540	0.4909	0.2531	0.2423
Geico	0.3957	0.0612	0.5972	0.4162	0.1547
GMAC	0.0298	0.0159	0.9474	0.1243	0.5330
Hartford	0.1156	0.0427	0.7610	0.1734	0.3691
Liberty Mutual	0.1574	0.0458	0.7018	0.1918	0.2910
Mercury	0.0864	0.0429	0.8180	0.2857	0.4972
MetLife	0.0776	0.0370	0.7662	0.1006	0.4772
Nationwide	0.1571	0.0477	0.5590	0.1830	0.3036
Progressive	0.3733	0.0669	0.6553	0.3998	0.1792
Safeco	0.0894	0.0519	0.8610	0.2240	0.5805
State Farm	0.3665	0.0543	0.5162	0.3825	0.1481
Travelers	0.1233	0.0491	0.7023	0.1615	0.3978
USAA	0.0940	0.0479	0.7653	0.1758	0.5099
Average	0.1765	0.0498	0.6391	0.2627	0.2822

consumer spends more than the median amount of hours per week online (14 hours). Although we do not observe online search advertising in our data, these fixed effects control for the amount of exposure to targeted online advertising as long as the targeting is based on demographics. Recall that brand, year, and brand-year fixed effects are subsumed in the brand-demographics-year and online-brand-demographics-year fixed effects. Third, μ_{bd} are a brand-DMA fixed effects that control for time-invariant, brand-specific local unobservables.

Additionally, we include several control variables: brand-state-month fixed effects, ζ_{bst}, capture changes at the state level such as changes in insurance rates, that is, premium levels, or regulations over time. For a small portion of ads, we either did not receive the creative files, the quality of the creative files was too bad for coding or the ads were not reliably coded and thus excluded from the empirical analysis. We control for spending (per household) on ads with missing content using $\psi_{bd,t-1} = \beta_4 \log(1 + A^m_{bd,t-1})$. Additionally, we also include survey fixed effects, τ_g^{survey}, to control for any systematic differences across surveys. And finally, ϵ_{ibt} is a standard normally distributed error term.

The conditional consideration model is given by

$$\begin{aligned} Y_{ibt} &= \beta_1 \log\left(1 + A^f_{bd,t-1}\right) + \beta_2 \log\left(1 + A^{f,nf}_{bd,t-1}\right) \\ &\quad + \beta_3 \log\left(1 + A^{nf}_{bd,t-1}\right) + \varrho^{D_i}_{bg} + \varphi^{D_i}_{bg}\mathbb{I}^o_i + \mu_{bd} \\ &\quad + \zeta_{bst} + \psi_{bd,t-1} + \tau^{survey}_g + \epsilon_{ibt} \qquad \forall\, b \in S^{aware}_i. \end{aligned} \tag{2}$$

Note that we condition on consumers' awareness sets S_i^{aware} when estimating the effects of advertising on consumers' consideration decisions, that is, we only include the set of brands for each consumer that the consumer is aware of. We do so once using consumers' unaided and once consumers' aided awareness sets. Similarly, in the following model describing conditional purchase, we condition on each individual consumer's consideration set S_i^{cons}

$$\begin{aligned} Y_{ibt} &= \beta_1 \log\left(1 + A^f_{bd,t-1}\right) + \beta_2 \log\left(1 + A^{f,nf}_{bd,t-1}\right) \\ &\quad + \beta_3 \log\left(1 + A^{nf}_{bd,t-1}\right) + \delta_1 \mathbb{I}_{Y_{ibt}} + \delta_2 \mathbb{I}^p_{ibt} + \varrho^{D_i}_{bg} \\ &\quad + \varphi^{D_i}_{bg}\mathbb{I}^o_i + \mu_{bd} + \zeta_{bst} + \psi_{bd,t-1} + \tau^{survey}_g + \epsilon_{ibt} \quad \forall\, b \in S^{cons}_i, \end{aligned} \tag{3}$$

where $\mathbb{I}_{Y_{ibt}}$ captures state dependence and is operationalized as a dummy variable indicating whether brand b chosen in time period t is the same brand that consumer i chose in the previous policy period. The variable $\mathbb{I}^p_{ibt}$ is also a dummy variable that indicates whether brand b offered the lowest premium for consumer i in time period t among the brands consumer i considered and is a self-reported variable. Thus, whereas the brand-state-month fixed effects ζ_{bst} capture average premium changes across all consumers in a state (for a company and year), the dummy variable $\mathbb{I}^p_{ibt}$ is specific to each consumer and his or her consideration set.

Lastly, we also estimate an unconditional purchase model, that is, full information model. The model is defined as in Equation (3) but without conditioning on consumers' consideration sets. That is, consumers choose a brand for purchase among all brands in the market and not only among the brands in their consideration sets.

5.2. Border Strategy

In the specifications shown in this section, we control for all endogeneity types presented in the previous section and, additionally, also for endogeneity due to time-varying, brand-specific local unobservables. The models discussed in this section are based on the discussion in Section 3.2 and are—in contrast to the previous section—estimated using the DMA border sample.

The (unaided and aided) awareness model is given by

$$\begin{aligned} Y_{ibt} = {} & \beta_1 \log\left(1 + A^{f}_{bd,t-1}\right) + \beta_2 \log\left(1 + A^{f,nf}_{bd,t-1}\right) \\ & + \beta_3 \log\left(1 + A^{nf}_{bd,t-1}\right) + \varrho^{D_i}_{bg} + \varphi^{D_i}_{bg} \mathbb{I}^{o}_{i} + \nu_{bdr} \\ & + \eta_{brt} + \zeta_{bst} + \psi_{bd,t-1} + \tau^{survey}_{g} + \epsilon_{ibt} \qquad \forall\, b \neq b_{CI}, \end{aligned} \tag{4}$$

where ν_{bdr} are brand-border-DMA fixed effects (r represents border) and capture persistent differences across different border regions.[27] Further, η_{brt} are brand-border-month fixed effects and represent unobserved border-region-specific trends. Next, the conditional consideration model is given by

$$\begin{aligned} Y_{ibt} = {} & \beta_1 \log\left(1 + A^{f}_{bd,t-1}\right) + \beta_2 \log\left(1 + A^{f,nf}_{bd,t-1}\right) \\ & + \beta_3 \log\left(1 + A^{nf}_{bd,t-1}\right) + \varrho^{D_i}_{bg} + \varphi^{D_i}_{bg} \mathbb{I}^{o}_{i} + \nu_{bdr} \\ & + \eta_{brt} + \zeta_{bst} + \psi_{bd,t-1} + \tau^{survey}_{g} + \epsilon_{ibt} \qquad \forall\, b \in S^{aware}_{i}, \end{aligned} \tag{5}$$

and the conditional purchase model is given by

$$\begin{aligned} Y_{ibt} = {} & \beta_1 \log\left(1 + A^{f}_{bd,t-1}\right) + \beta_2 \log\left(1 + A^{f,nf}_{bd,t-1}\right) \\ & + \beta_3 \log\left(1 + A^{nf}_{bd,t-1}\right) + \delta_1 \mathbb{I}_{Y_{ibt}} + \delta_2 \mathbb{I}^{p}_{ibt} + \varrho^{D_i}_{bg} \\ & + \varphi^{D_i}_{bg} \mathbb{I}^{o}_{i} + \nu_{bdr} + \eta_{brt} + \zeta_{bst} + \psi_{bd,t-1} \\ & + \tau^{survey}_{g} + \epsilon_{ibt} \qquad \forall\, b \in S^{cons}_{i}. \end{aligned} \tag{6}$$

The unconditional purchase model is defined as in Equation (6) but without conditioning on consumers' consideration sets.

6. Results and Discussion

Before interpreting individual coefficient estimates, we start this section by taking an overall look at the empirical results from our basic specification and the border strategy. We discuss econometric concerns and provide additional empirical evidence in Section 6.1 to determine which of these two specifications we view as our main one and whether to view the estimated advertising effects as local or overall effects. In Sections 6.2–6.4, we address more econometric concerns related to within-brand variation in spending on an advertising type, alternative advertising operationalizations, and selection based on individual- and brand-specific unobservables. We then interpret our main set of results in Sections 6.5 and 6.6. Finally, in Section 6.7, we show that advertising with both informational and noninformational content influences a subset of consumers.

6.1. Model Specifications

We present our results for the basic specification in Table 7 and for the border strategy in Table 8. Note that we report both the number of observations and the "effective" number of observations, that is, the number of observations remaining after observations collinear with included fixed effects have been dropped, and that all standard errors are clustered at the DMA-level.[28]

The results from the basic specification and the border strategy are not identical. The main differences are the magnitudes of the estimated advertising effects and only informational advertising having a significant effect on unaided awareness in the basic specification, whereas this coefficient is insignificant when using the border strategy. This observation brings up the following question: which set of results is the "correct" one? There are three potential reasons for why the results from the basic specification and the border strategy differ: (a) the DMA border sample has less statistical power to estimate advertising effects; (b) consumers living anywhere in DMAs and consumers living in border counties of DMAs are different, that is, have different preferences and especially a different sensitivity to advertising; and (c) the results are different because of time-varying, brand-specific local unobservables, that is, it is important to account for this type of endogeneity. In the following, we discuss each of these three reasons and present (no) evidence for each of them.

If statistical power is one of the reasons, we would expect the standard errors in the border strategy to be larger than those in the basic specification because of the smaller sample size and, as a consequence of the larger standard errors, a subset of the coefficients that were significant in the basic specification to be significant in the border strategy. If statistical power is the *only* concern, we would expect the estimated coefficients in both analyses to be of similar magnitude. Comparing the results from the basic specification and the border strategy in Tables 7 and 8, we find that the standard errors are overall larger in the border strategy than in the basic specification pointing to a loss in statistical power. Further, we find a subset of the coefficients that were significant in the basic specification (some at $p < 0.10$) to be significant in the

Table 7. Results for Basic Specification (Complete DMA Sample)

Variables	(i) Unaided awareness	(ii) Aided awareness	(iii) Consideration conditional on unaided awareness	(iv) Consideration conditional on aided awareness	(v) Choice conditional on consideration	(vi) Full information
Panel A: Advertising quantity						
Advertising spending per household in $[a]	0.1847*	0.1110***	−0.0664	0.1441	−0.0066	0.0379***
	(0.0506)	(0.0532)	(0.2457)	(0.1568)	(0.1532)	(0.0180)
Same insurer as in previous year					0.2017*	0.8157*
(Yes = 1)					(0.0067)	(0.0333)
Insurer provided the best price					0.7305*	0.4660*
(Yes = 1)					(0.0090)	(0.0224)
Brand-demographics-year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Online-brand-demographics-year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Brand-DMA fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Number of observations	4,142,607	2,703,582	117,914	286,878	71,167	3,623,553
Effective number of observations[b]	4,139,751	2,700,516	109,200	280,133	62,001	3,621,098
Panel B: Advertising content						
Spending per household in $ on ads with...						
... Informational content only[a]	0.7083**	0.3933***	−0.0526	−0.4703	−0.6891	−0.1391
	(0.2607)	(0.1795)	(0.7914)	(0.6321)	(0.7570)	(0.2020)
... Noninformational content only[a]	0.0782****	0.0157	−0.3681	0.4625	0.4078	0.0490****
	(0.0426)	(0.0184)	(0.5439)	(0.3369)	(0.4582)	(0.0284)
... Both informational and noninformational content[a]	0.0774	−0.0868	−0.2106	0.0767	−0.1634	−0.0060
	(0.1058)	(0.0762)	(0.3062)	(0.2240)	(0.1605)	(0.0436)
Same insurer as in previous year					0.2016*	0.8157*
(Yes = 1)					(0.0067)	(0.0333)
Insurer provided the best price					0.7305*	0.4660*
(Yes = 1)					(0.0091)	(0.0224)
Brand-demographics-year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Online-brand-demographics-year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Brand-DMA fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Number of observations	4,142,607	2,703,582	117,914	286,878	71,167	3,623,553
Effective number of observations[b]	4,139,751	2,700,516	109,200	280,133	62,001	3,621,099

Notes. Controls included in advertising quantity regressions: brand-state-month fixed effects and survey fixed effects. Controls included in advertising content regressions: brand-state-month fixed effects, spending with missing content, and survey fixed effects. Standard errors in parentheses (clustered at the DMA level).

[a]Measured on a logarithmic scale.

[b]After dropping observations because of collinearity with fixed effects.

*< 0.001; **< 0.01; ***< 0.05; ****< 0.10.

border strategy – again, consistent with the explanation of a loss in statistical power. Lastly, some of the estimated coefficients change in magnitude (e.g., spending on ads with only informational content in the unaided awareness regression) pointing to a loss in statistical power not being the only explanation for the difference in results.

Next, we evaluate the potential reason that consumers living anywhere in DMAs and consumers living in border counties of DMAs are different, that is, have different preferences and especially a different sensitivity to advertising. We do so by estimating an intermediate model: a model that includes the same set of fixed effects as in the basic specification, that is, tackles time-invariant, brand-specific local unobservables but not time-varying, brand-specific local unobservables, and is estimated using the DMA border sample. The results are shown as Intermediate Analysis 1 in Table 9. We find that the results for both advertising quantity and advertising content shown in this intermediate model in Table 9 are extremely similar in terms of coefficient estimates to those in the basic specification in Table 7. Thus, different sensitivity to advertising between consumers in the complete DMA sample and consumers in the DMA border sample does not explain the difference in results. This finding also has implications for the generalizability of advertising results found using the border strategy, that is, to the interpretability of the estimated advertising effects as applying to the whole population, in our empirical context. We view this result as strongly supporting Assumption 3(a) from Section 3.2 and the notion that the advertising effects found using the DMA border sample (and regression discontinuity approach) can be interpreted as the effect of advertising in the whole population and not only for consumers living at DMA borders.

Table 8. Results for Border Strategy (DMA Border Sample)

Variables	(i) Unaided awareness	(ii) Aided awareness	(iii) Consideration conditional on unaided awareness	(iv) Consideration conditional on aided awareness	(v) Choice conditional on consideration	(vi) Full information
Panel A: Advertising quantity						
Advertising spending per household in $[a]	0.3009*	0.1745***	−0.3085	−0.2981	0.0590	0.0628***
	(0.0617)	(0.0841)	(0.3528)	(0.3137)	(0.2191)	(0.0290)
Same insurer as in previous year					0.1936*	0.8464*
(Yes = 1)					(0.0201)	(0.0339)
Insurer provided the best price					0.7263*	0.4488*
(Yes = 1)					(0.0241)	(0.0272)
Brand-demographics-year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Online-brand-demographics-year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Brand-border-DMA fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Brand-border-month fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Number of observations	2,105,376	1,376,025	63,743	155,464	38,720	1,808,446
Effective number of observations[b]	2,078,790	1,348,641	40,242	117,424	19,994	1,785,270
Panel B: Advertising content						
Spending per household in $ on ads with...						
... Informational content only[a]	−0.2980	0.7102**	−0.1569	−1.0660	−0.2063	0.1348
	(0.2115)	(0.2444)	(0.8491)	(0.5749)	(1.4829)	(0.4053)
... Noninformational content only[a]	0.2306*	0.0034	−0.1468	−0.1672	−0.0734	0.8633***
	(0.0520)	(0.0327)	(0.3037)	(0.3187)	(0.2816)	(0.3452)
... Both informational and noninformational content[a]	0.0555	0.1030	0.3515	0.0187	−0.1733	0.0309
	(0.1893)	(0.1245)	(0.5393)	(0.4994)	(0.4495)	(0.0721)
Same insurer as in previous year					0.1934*	0.8464*
(Yes = 1)					(0.0201)	(0.0339)
Insurer provided the best price					0.7264*	0.4487*
(Yes = 1)					(0.0241)	(0.0272)
Brand-demographics-year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Online-brand-demographics-year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Brand-border-DMA fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Brand-border-month fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Number of observations	2,105,376	1,376,025	63,743	155,464	38,720	1,808,446
Effective number of observations[b]	2,078,790	1,348,641	40,242	117,424	19,994	1,785,270

Notes. Controls included in advertising quantity regressions: brand-state-month fixed effects and survey fixed effects. Controls included in advertising content regressions: brand-state-month fixed effects, spending with missing content, and survey fixed effects. Standard errors in parentheses (clustered at the DMA level).
[a]Measured on a logarithmic scale.
[b]After dropping observations because of collinearity with fixed effects.
*< 0.001; **< 0.01; ***< 0.05.

The remaining explanation is that endogeneity due to time-varying, brand-specific local unobservables—on top of endogeneity due to time-invariant, brand-specific local unobservables—matters. Although this is the residual explanation, we also want to show some direct evidence for it. To achieve this, we estimate a second intermediate model shown as Intermediate Analysis 2 in Table 9. In it, compared with the Intermediate Analysis 1, we add brand-border-month fixed effects, that is, this specification uses the regression discontinuity approach (and the DMA border sample is used for estimation). The difference to the border strategy specification in Table 8 is that, in Intermediate Analysis 2, we estimate brand-DMA fixed effects instead of brand-border-DMA fixed effects. The inclusion of brand-border-month fixed effects captures time-varying, brand-specific local unobservables—the source of the endogeneity concerns leading up to the use of the border strategy. Although the results in the Intermediate Analysis 1 were very similar to those from the basic specification (Table 7), the results from the Intermediate Analysis 2 are very similar to those from the border strategy (Table 8). This finding points to the difference in results in the basic specification and the border strategy being primarily due to the border strategy accounting for endogeneity due to time-varying, brand-specific local unobservables and to the importance of doing so when measuring the effects of advertising in our empirical context.

We conclude that it is important to tackle endogeneity concerns due to global unobservables, targeting based on demographics, *and* time-varying, brand-specific local

Table 9. Intermediate Results

Variables	(i) Unaided awareness	(ii) Aided awareness	(iii) Consideration conditional on unaided awareness	(iv) Consideration conditional on aided awareness	(v) Choice conditional on consideration	(vi) Full information
Panel A: Intermediate analysis 1: Basic specification–DMA border sample						
Advertising quantity						
Advertising spending per household in $[a]	0.1725**	0.1278****	−0.3186	0.0716	0.0705	0.0488***
	(0.0627)	(0.0646)	(0.3253)	(0.2100)	(0.2035)	(0.0229)
Same insurer as in previous year					0.2030*	0.8229*
(Yes = 1)					(0.0117)	(0.0370)
Insurer provided the best price					0.7253*	0.4566*
(Yes = 1)					(0.0155)	(0.0269)
Brand-demographics-year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Online-brand-demographics-year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Brand-DMA fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Number of observations	2,105,376	1,376,025	63,743	155,464	38,720	1,808,446
Effective number of observations[b]	2,102,583	1,373,127	57,875	149,100	33,123	1,806,040
Advertising content						
Spending per household in $ on ads with...						
... Informational content only[a]	0.9261**	0.6016***	−1.0492	−0.9652	0.1282	−0.1830
	(0.3318)	(0.2502)	(0.6101)	(0.5438)	(0.7043)	(0.2999)
... Noninformational content only[a]	0.0957****	0.0146	−0.4254	0.0142	0.2234	0.0596****
	(0.0535)	(0.0236)	(0.4678)	(0.4196)	(0.3444)	(0.0336)
... Both informational and noninformational content[a]	0.0203	−0.0454	−0.3783	0.0420	0.0086	0.0267
	(0.1403)	(0.1069)	(0.3604)	(0.2573)	(0.3396)	(0.0647)
Same insurer as in previous year					0.2030*	0.8283*
(Yes = 1)					(0.0118)	(0.0368)
Insurer provided the best price					0.7253*	0.4635*
(Yes = 1)					(0.0155)	(0.0265)
Brand-demographics-year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Online-brand-demographics-year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Brand-DMA fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Number of observations	2,105,376	1,376,025	63,743	155,464	38,720	1,808,446
Effective number of observations[b]	2,102,583	1,373,127	57,875	149,100	33,123	1,806,040
Panel B: Intermediate analysis 2: Basic specification & brand-border-month fixed effects–DMA border sample						
Advertising quantity						
Advertising spending per household in $[a]	0.3055*	0.1895***	−0.3265	−0.3082	0.0696	0.0938**
	(0.0629)	(0.0860)	(0.3489)	(0.3083)	(0.2163)	(0.0312)
Same insurer as in previous year					0.1938*	0.8463*
(Yes = 1)					(0.0200)	(0.0339)
Insurer provided the best price					0.7259*	0.4488*
(Yes = 1)					(0.0239)	(0.0272)
Brand-demographics-year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Online-brand-demographics-year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Brand-DMA fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Brand-border-month fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Number of observations	2,105,376	1,376,025	63,743	155,464	38,720	1,808,446
Effective number of observations[b]	2,078,916	1,349,019	40,587	118,269	20,217	1,785,388
Advertising content						
Spending per household in $ on ads with...						
... Informational content only[a]	−0.3197	0.7094**	−0.1986	−1.0465	−0.1549	0.1650
	(0.2114)	(0.2439)	(0.7962)	(0.5509)	(1.4516)	(0.4063)
... Noninformational content only[a]	0.2319*	0.0048	−0.1506	−0.1987	−0.0758	0.8624***
	(0.0517)	(0.0328)	(0.3006)	(0.3071)	(0.2704)	(0.3427)
... Both informational and noninformational content[a]	0.0610	0.1166	0.3252	0.0090	−0.1098	0.0203
	(0.1884)	(0.1232)	(0.5386)	(0.4967)	(0.4359)	(0.0714)
Same insurer as in previous year					0.1936*	0.8463*
(Yes = 1)					(0.0200)	(0.0339)
Insurer provided the best price					0.7260*	0.4487*
(Yes = 1)					(0.0239)	(0.0272)

Table 9. (Continued)

Variables	(i) Unaided awareness	(ii) Aided awareness	(iii) Consideration conditional on unaided awareness	(iv) Consideration conditional on aided awareness	(v) Choice conditional on consideration	(vi) Full information
Brand-demographics-year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Online-brand-demographics-year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Brand-DMA fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Brand-border-month fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Number of observations	2,105,376	1,376,025	63,743	155,464	38,720	1,808,446
Effective number of observations[b]	2,078,916	1,349,019	40,587	118,269	20,217	1,785,388

Notes. Controls included in advertising quantity regressions: brand-state-month fixed effects and survey fixed effects.
Controls included in advertising content regressions: brand-state-month fixed effects, spending with missing content, and survey fixed effects. Standard errors in parentheses (clustered at the DMA level).
[a]Measured on a logarithmic scale.
[b]After dropping observations because of collinearity with fixed effects.
*< 0.001; **< 0.01; ***< 0.05; ****< 0.10.

unobservables when estimating the effects of advertising for the auto insurance industry. Further, we find consumers' sensitivity to advertising to be similar among consumers living in DMA border counties and among consumers living anywhere in DMAs providing evidence for the generalizability of advertising effects found using the DMA border sample, that is, the interpretability of the estimated effects as overall advertising effects.

6.2. Within-Brand Variation in Spending on an Advertising Type

In Section 3.1.3, we posed Assumption 1 stating that, in order to be able to make causal claims for *all* brands regarding the effects of different types of advertising, we have to assume that the within-brand variation in a type of advertising is the same for a brand which employs this type of advertising and a brand which does not (if it were to utilize it). Although this assumption is not directly verifiable, we provide suggestive evidence here that it may hold in our specific empirical context. We do so by estimating our main specification (border strategy) for advertising content with separate advertising coefficients for brands which employ one or two types of advertising and brands that employ all three types of advertising.[30] If Assumption 1 holds, we would expect the advertising effects for both groups of brands to be similar.

The results are shown in Table 10. The interaction effects between each advertising type and brands utilizing all three advertising types are insignificant, that is, we do not find significant differences in the effects of advertising content for brands that employ a subset of advertising types compared with brands that employ all three advertising types. Further, our main results hold: ads with only noninformational content affect unaided awareness and ads with only informational content affect aided awareness. These results suggest that Assumption 1 holds; however, they are not conclusive. To be careful, we therefore interpret the results for advertising content as correlational.

6.3. Alternative Advertising Operationalizations

Previous literature has sometimes operationalized the advertising variable as advertising goodwill (e.g., Shapiro 2018) or also included competitive advertising in the empirical analysis (e.g., Anderson et al. 2016). Here, we show that our main results are robust to the inclusion of competitive advertising and to an advertising goodwill specification.

We start by discussing how we include competitive advertising in the estimations. Recall that our data contain information on a large number of brands (21). Therefore estimating a separate effect for each competitor's advertising is not feasible. For that reason, we re-estimate all our models using the border strategy twice: once, for each advertising variable, also including the corresponding sum of all competitors' advertising spending as an additional independent variable and once, again for each advertising variable, also including the sum of the four largest auto insurance brands (Allstate, Geico, Progressive, State Farm) and the sum of all other auto insurance brands as two additional independent variables. The results for the advertising quantity and advertising content regressions using these alternative model specifications are shown in Tables F-1 and F-2 in Web Appendix F. Although competitive advertising has significant effects, more importantly, our results for the focal brand are very similar to those from our main specification (border strategy) presented in Table 8.

Next, we estimated our main specification (border strategy) with a goodwill operationalization of the advertising variables. We included lagged advertising up to 12 months and, to determine the appropriate

Table 10. Advertising Content Results for Awareness by Number of Content Types (Border Strategy)

Variables	(i) Unaided awareness	(ii) Aided awareness
Spending per household in $ on ads with...		
... Informational content only[a]	0.0864	0.9848***
	(0.3969)	(0.4862)
... Noninformational content only[a]	0.3482****	−0.0277
	(0.1919)	(0.2671)
... Both informational and noninformational content[a]	0.1542	0.1520
	(0.2299)	(0.1250)
Brands utilizing three ad types × spending per household in $ on ads with...		
... Informational content only[a]	−0.5368	−0.0516
	(0.5237)	(0.7483)
... Noninformational content only[a]	−0.1356	0.0337
	(0.1920)	(0.2664)
... Both informational and noninformational content[a]	−0.4617	−1.4527
	(0.3704)	(1.0893)
Brand-demographics-year fixed effects	Yes	Yes
Online-brand-demographics-year fixed effects	Yes	Yes
Brand-border-DMA fixed effects	Yes	Yes
Brand-border-month fixed effects	Yes	Yes
Number of observations	2,105,376	1,376,025
Effective number of observations[b]	2,078,790	1,348,641

Notes. Controls included: brand-state-month fixed effects, spending with missing content, and survey fixed effects. Standard errors in parentheses (clustered at the DMA level).

[a]Measured on a logarithmic scale.

[b]After dropping observations because of collinearity with fixed effects.

*< 0.001; **< 0.01; ***< 0.05; ****< 0.10.

carryover factor, we conducted a grid search with 11 different carryover factors for each regression (see, e.g., Shapiro et al. 2021). The examined carryover factors ranged from 0 to 1 (with increments of 0.1) and, for each regression, we picked the one with the smallest predicted mean square error. The results are shown in Table G-1 in Web Appendix G and are qualitatively similar to those from our main specification (border strategy) displayed in Table 8.

6.4. Selection Based on Individual- and Brand-Specific Unobservables

A potential concern are individual- *and* brand-specific unobservables that are correlated across the three purchase stages (awareness, consideration, purchase). Not accounting for such variables can lead to selection issues and biased estimates in the conditional consideration and conditional choice stages of the advertising quantity and advertising content regressions (but is not a concern for the awareness stage—see Maddala 1983 or Wachtel and Otter 2013). Because we model three stages of consumers' purchase process, selection could potentially occur twice: as consumers move from awareness to consideration and as consumers move from consideration to purchase. However, in our specific empirical content of the auto insurance industry, selection can only happen *once* when consumers move from awareness to consideration/shopping, that is, decide whether to actively shop or to passively remain insured with their previous insurance company. The reason is that having auto insurance is mandatory and thus all consumers who shop also have to buy an insurance policy, that is, no purchase is not an option. Although we do not account for the potential existence of such unobserved, individual and brand-specific variables in our modeling and estimation decisions, we discuss here why selection does not appear to be a concern in our specific empirical setting.

Positively (negatively) correlated, unobserved, individual- and brand-specific variables can lead to an overestimation (underestimation) of the effect of advertising in the conditional consideration and conditional choice stages. Given that we find the effects of advertising quantity and different types of advertising content in the conditional consideration and conditional purchase stages to be insignificant (see columns (iii)–(v) in Table 8), overestimation is not a concern. However, underestimation of the effects of advertising quantity and advertising content in the conditional consideration and conditional purchase regressions might be a potential concern.

We address selection concerns with two robustness checks. The results for conditional consideration and conditional purchase are shown in Tables H-1 and H-2 in Web Appendix H. Recommendations by family and friends are an example of an individual- and brand-specific variable. In our data, we observe whether a consumer received a recommendation but not the valence of the recommendation. Therefore, in the first robustness check whose results are shown in Table H-1, we drop all consumers who report having received a recommendation. Furthermore, we also observe in our data whether there is a brand a consumer would never consider. In a second robustness check whose results are shown in Table H-2, we control for the presence of such a brand. In both robustness checks, we find our advertising coefficient estimates for conditional consideration and conditional purchase in the advertising quantity and advertising content regressions to be similar to those from our main model specification. We therefore conclude that selection is not a concern in our specific empirical application.

6.5. Discussion of Advertising Intensity Results

We now turn to the estimates from the regressions investigating the effects of advertising quantity. In discussing our results, we focus on our main specification shown in the top half of Table 8. We find total advertising spending per household to significantly affect consumers' unaided and aided awareness (see columns (i) and (ii) in the top half of Table 8). To understand the magnitudes of these advertising effects, we calculate average advertising elasticities and find them to be 0.07 and 0.01 for unaided and aided awareness, respectively.

Next, we describe the effects of advertising intensity on the other stages of the purchase process. Recall that, when estimating the effects of advertising quantity on consideration, we only use data on shoppers because nonshoppers do not shop and thus do not form consideration sets. In columns (iii) and (iv) in the top half of Table 8, we show the results for consideration conditional on unaided and aided awareness, respectively. In both cases, advertising does not have significant effects on consideration. Turning to the purchase stage, we find total advertising spending per household to have an insignificant effect on conditional purchase (column (v) in Table 8). Compared with the awareness and conditional consideration regressions, we include two additional variables in the conditional purchase regression (column (v) in Table 8): a dummy variable indicating whether a brand is a consumer's previous insurance provider and a dummy variable indicating whether a brand offered a consumer the lowest premium. The parameter estimates for both variables are significant and have the expected signs: consumers are more likely to purchase an insurance policy from their previous insurance provider and are also more likely to pick the insurance brand that offers them the lowest premium.

Lastly, in column (vi) in Table 8, we compare our results to those from an unconditional purchase model, sometimes also called a full information model, that is, a model in which consumers are aware of and consider all brands in the market for purchase. We find advertising to have a small, but significant positive effect on purchase (elasticity: 0.03). The advertising elasticity is in line with those found in Shapiro et al. (2021) for a large number of products. However, this result stands in contrast to the result from the conditional purchase model in column (v) in Table 8 in which advertising did not have a significant effect. Further, under full information, the effect of the previous insurer dummy is overestimated and the effect of the best price dummy is underestimated. Thus, similar to previous literature (e.g., Goeree 2008, Pires 2016), we find that not accounting for consumers' limited information leads to quantitatively and qualitatively different results.

To recap, we find that advertising intensity positively affects consumer purchase behavior. However, it does so not by directly affecting consumer purchase decisions but rather indirectly by affecting consumers' (unaided and aided) awareness. These findings are consistent with those in Honka et al. (2017) in the context of retail banks. Further, these results are also consistent with advertising professionals' recent demands for companies to focus on consumer awareness rather than consumer engagement.[31]

6.6. Discussion of Advertising Content Results

We now discuss our estimates from the advertising content regressions. Here again, we focus on the results from our main specification (border strategy) presented in the lower half of Table 8. Columns (i) and (ii) in Table 8 show the parameter estimates for unaided and aided awareness. Our results reveal that advertising with only noninformational content has a significant positive effect on unaided awareness (elasticity: 0.02), whereas advertising with only informational content has a positive significant effect on aided awareness (elasticity: 0.02).

It is well known that unaided and aided awareness do not refer to the same concept (see, e.g., Bagozzi and Silk 1983). Unaided awareness or (brand) recall captures situations in which a consumer must independently produce previously acquired information, whereas aided awareness or (brand) recognition describes cases in which a consumer is given possible choices and must indicate which one was previously seen (Lynch and Srull 1982, Alba et al. 1991). Multiple theories exist regarding the relationship between recall and recognition. For example, Gillund and Shiffrin (1984) suggest that recall is driven by a search process that operates slowly and uncertainly. In contrast, recognition is executed through a complex direct access process in which many memory images are contacted together, that is, recall and recognition differ in the memory retrieval process they use. On the other hand, Bagozzi and Silk (1983) and Finn (1992) argue that memory is a multidimensional construct and that recall and recognition access different aspects of memory. Previous (experimental) research has also shown that some factors such as delay and position influence recall and recognition similarly (see, e.g., Ratcliff and Murdock 1976, Raaijmakers and Shiffrin 1981), whereas other factors, such as types of rehearsal and context shifts influence recall and recognition differently (see, e.g., Woodward et al. 1973, Smith et al. 1978). If different types of ads contain different factors that affect recall and recognition differently, one type of ad might influence recall but not recognition and vice versa.

We first discuss our results for brand recall (unaided awareness): a significant effect of ads with only noninformational content and an insignificant effect of ads with only informational content. Recall that only noninformational ads are ads with brand name focus and/or emotionally appealing content. Brand name focused ads are creatives that either dominantly and/or frequently show the brand name or contain no information on car insurance but that mention the brand name (e.g., TV program sponsorships, public service messages). As the term says, emotionally appealing creatives appeal to consumers' emotions, often contain a story or an unexpected event, and allow for imagination and inspiration (Heath and Heath 2008). Advertisements with only noninformational content are commonly more creative than advertisements with only informational content. Our results are in line with previous causal research that has shown that creativity increases brand recall, but not recognition (see, e.g., Till and Baack 2013). Evidence on whether emotional ads increase recall and/or recognition is mixed. Previous experimental studies often included one and not both memory measures and suffer from small sample sizes (see, e.g., Thorsen 1991, Leigh et al. 2006, Mehta and Purvis 2006). Our results suggest that emotionally appealing content increases brand recall, but not recognition.

Next, for brand recognition (aided awareness), we find an insignificant effect of ads with only noninformational content and a significant effect of ads with only informational content. Our results are in line with previous causal research that has shown that familiarity, that is, the number of product-related experiences and exposures, increases recognition (Alba et al. 1991). Further, the accessibility of product attributes/positioning affects which brands are recognized as members of a product category (McCloskey and Glucksberg 1979, Alba et al. 1991). Recall that informational advertising contains descriptions of price and nonprice product attributes, whereas noninformational advertising does not. Thus, informational advertising increases the accessibility of these product attributes and accessibility, in turn, increases brand recognition.

In columns (iii) and (iv) in the lower half of Table 8, we show the results for consideration conditional on unaided and aided awareness, respectively, and the results for conditional purchase in column (v). No advertising type has a significant effect on conditional consideration or conditional purchase. Further, in the conditional purchase regression (column (v) in lower half of Table 8), the coefficient estimates for the previous insurer and lowest price dummies are very similar to those in top half of Table 8 (column (v)) where we showed results from modeling the effects of advertising intensity.

Similar to the analyses for advertising intensity, we also estimate a full information model, that is, unconditional purchase model, for advertising content. The results are shown in column (vi) in lower half of Table 8. We find a small and significant effect of only noninformational advertising (elasticity: 0.06). This result is consistent with Cronqvist (2006) and Bertrand et al. (2010) who find noninformational content to influence consumers' decision-making for financial services. Considering a wider range of industries and contexts, Sahni et al. (2018) and Lee et al. (2018) also found that noninformational content can affect email openings, sales, and customer engagement.

Taking a step back, our results also speak to the concepts of informative and persuasive advertising (Bagwell 2007). Thorsen Informative advertising informs consumers about product existence and (price and nonprice) product features, whereas persuasive advertising alters consumers' tastes and creates spurious product differentiation (Bagwell 2007). *Any* type of advertising that conveys the existence of a product to consumers, that is, makes consumers aware of a product, has an informative effect. To put it differently, the effect of *both* informational and noninformational advertising content on consumer awareness is informative. We find this informative effect in our empirical application.

To understand whether advertising has an informative and/or persuasive effect in the consideration and purchase stages of the purchase process, the content of ads has to be observed. If noninformational ad content were to affect conditional consideration and/or conditional purchase, the effect of advertising could be interpreted as persuasive. If informational ad content were to affect conditional consideration and/or purchase, the effect of advertising could be interpreted as informative. In our empirical application, we do not find a significant effect of any type of advertising in the conditional consideration and conditional purchase regressions. Thus, overall, our results are consistent with an informative effect of advertising in the auto insurance industry.

To summarize, we find advertising only containing noninformational content to increase unaided awareness, whereas advertising only containing informational content increases aided awareness. We do not find significant effects of advertising containing both informational and noninformational content. This last finding prompts the question why companies spend a large portion of their advertising budgets on advertising that contains both informational and noninformational content if that type of advertising is not effective. We provide an explanation in the following section.

6.7. For Whom is Advertising with Both Informational and Noninformational Content Effective?

Companies must allocate their spending between customer acquisition and customer retention. The customer lifetime value (CLV) literature provides a framework on how to optimally do so (e.g., Blattberg and Deighton 1996, Rust et al. 2004, Reinartz et al. 2005). Whereas for many industries spending on retention has a bigger impact on CLV than spending on acquisition, this is not the case for industries that are characterized by a very high retention rate. If customers are very loyal, spending on acquisition has a bigger impact on CLV than spending on retention: companies will compete fiercely to acquire customers and benefit from the high retention rate in later time periods (e.g., Blattberg et al. 2001, Villanueva and Hanssens 2007).

The retention rate in the auto insurance industry is very high. According to industry sources[32] and Israel (2005), around 87%–90% of customers stay with their insurance provider after a one-year contract period among the most established auto insurance companies and this percentage increases with customer tenure. In our data, we observe an average retention rate of 88%, that is, only 12% of consumers are switchers. For the auto insurance industry, Honka (2014) finds that this high retention rate is mostly driven by consumer search costs, followed by customer satisfaction, and switching costs. After removing these three retention drivers, she finds that the baseline retention rate, which also includes the effects of advertising, is around 9%. To put it differently, advertising is not an important driver of customer retention.

We therefore investigate here whether advertising plays an important role in customer acquisition. We do so by looking at the heterogeneous effects of advertising for shoppers vs. nonshoppers.[33] 32% of consumers in the DMA Border Sample are shoppers. The results are shown in Table 11.[34] Overall, our results for advertising quantity are consistent with those from the main specification: advertising intensity has significant effects on awareness for both shoppers and nonshoppers. However, advertising intensity has significantly larger effects on unaided awareness for shoppers than for nonshoppers with elasticities of 0.08 and 0.06, respectively. The results for advertising content are also consistent with those from the main specification: advertising only containing noninformational content increases unaided awareness, whereas advertising only containing informational content increases aided awareness. Additionally, for shoppers, we also find advertising with both informational and noninformational content to have a significant effect on unaided awareness (elasticity: 0.02).

Shopping is a well-known situational factor that increases consumers' involvement with a product or product category (Zaichkowsky 1985). Increased involvement motivates consumers to pay more attention to ads and increases their willingness to process larger amounts of content (Celsi and Olson 1988). Ads with both informational and noninformational content contain more content than ads with only informational content or ads with only noninformational content. Consumers who (plan to) shop are willing to process this larger amount of content.

The influence of advertising with both informational and noninformational content is not limited to shoppers only. Other groups of consumers, who are relatively involved with auto insurance, are willing to process larger amounts of information and are influenced by this type of advertising as well. Here, we show examples of two such groups: high-risk consumers and consumers with a change in circumstances.[35] We define high-risk consumers as consumers who had accidents or tickets during the last three years, who have a poor credit history, or who are younger than 25 years. Consumers have to satisfy one of these criteria to be classified as high-risk consumers (8.5% of consumers). High-risk consumers are more involved with auto insurance than the remainder of the population because of a higher probability of being dropped as a customer and because of paying higher premia and higher potential savings (from switching). Consumers with a change in their circumstances are individuals who either added/dropped a vehicle to their policy, added/dropped a driver to their policy, or had a change in their family (e.g., marriage, divorce) during the last 12 months (6.0% of consumers). Consumers with a change in circumstances must contact their insurance company to update their insurance policy, which foregrounds auto insurance.[36]

The results for advertising quantity, which are consistent with those from our main specification, are shown in Table J-1 in Web Appendix J. Similar to the results for shoppers, we also find that these two groups of consumers are more affected by advertising quantity in their unaided awareness (and also in aided awareness for consumers with change in circumstances) than the remainder of the population. In the following, we focus on the results for advertising content displayed in Table 12. For both groups of consumers, the results from our main specification hold: ads with only noninformational content increase unaided awareness and ads with only informational content increases aided awareness. Additionally, for high-risk consumers, we

Table 11. Effect Heterogeneity—Shoppers and Nonshoppers

	(i) Unaided awareness	(ii) Aided awareness
Panel A: Advertising quantity		
Advertising spending per household in $[a]	0.2296*	0.1903***
	(0.0668)	(0.0912)
Shopper dummy × advertising spending per household in $[a]	0.3049*	−0.0553
	(0.0493)	(0.0487)
Brand-demographics-year fixed effects	Yes	Yes
Online-brand-demographics-year fixed effects	Yes	Yes
Brand-border-DMA fixed effects	Yes	Yes
Brand-state-month fixed effects	Yes	Yes
Number of observations	2,105,376	1,376,025
Effective number of observations[b]	2,078,790	1,348,641
Panel B: Advertising content		
Spending per household in $ on...		
... Informational content[a]	−0.4160	0.6904***
	(0.2567)	(0.3164)
... Noninformational content[a]	0.2229*	−0.0057
	(0.0482)	(0.0380)
... Both informational and noninformational content[a]	−0.0537	0.1204
	(0.1852)	(0.1303)
Shopper dummy × spending per household in $ on...		
... Informational content[a]	0.2824	0.0102
	(0.2765)	(0.2929)
... Noninformational content[a]	0.0665	0.0974
	(0.1767)	(0.1313)
... Both informational and noninformational content[a]	0.3242*	−0.0359
	(0.0515)	(0.0397)
Brand-demographics-year fixed effects	Yes	Yes
Online-brand-demographics-year fixed effects	Yes	Yes
Brand-border-DMA fixed effects	Yes	Yes
Brand-state-month fixed effects	Yes	Yes
Number of observations	2,105,376	1,376,025
Effective number of observations[b]	2,078,790	1,348,641

Notes. Controls included in advertising quantity regressions: brand-state-month fixed effects and survey fixed effects. Controls included in advertising content regressions: brand-state-month fixed effects, spending with missing content, and survey fixed effects. Standard errors in parentheses (clustered at the DMA level).

[a]Measured on a logarithmic scale.

[b]After dropping observations because of collinearity with fixed effects.

*< 0.001; **< 0.01; ***< 0.05.

also find advertising with both informational and noninformational content to have a significant effect on unaided awareness (elasticity: 0.04). For consumers with a change in circumstances, advertising with both informational and noninformational content significantly affects both unaided and aided awareness with elasticities of 0.04 and 0.01, respectively. Furthermore, consumers with a change in circumstances are also significantly more susceptible to advertising with only informational content and advertising with only noninformational content than the remainder of the population.

We conclude that advertising with both informational and noninformational content significantly increases shoppers' awareness. This is the case because shoppers are more involved with auto insurance than nonshoppers. Involvement has been shown to increase consumers' attention to ads and their willingness to process larger amounts of content. Other groups of relatively involved consumers, such as high-risk consumers or consumers with a change in circumstances, are affected by advertising with both informational and noninformational content as well.

7. Additional Robustness Checks

We evaluate the robustness of our results using several checks. First, we evaluate the robustness of our results with respect to an alternative operationalization of the advertising quantity variable. Here, we re-estimate our models using the logarithm of total advertising *units* as our measure of advertising intensity. The results are shown in the top third of Table K-1 in Web Appendix K and are similar to our main results. Second, we investigate the robustness of our results with respect to our use of total advertising spending.

Table 12. High-Risk Consumers and Consumers with Change in Circumstances

	(i) Unaided awareness	(ii) Aided awareness	(iii) Consideration conditional on unaided awareness	(iv) Consideration conditional on aided awareness	(v) Choice conditional on consideration
Part A: High-risk consumers					
Spending per household in $ on ads with...					
... Informational content only[a]	−0.3547	0.7064**	−0.1340	−1.1232	−0.2230
	(0.2072)	(0.2407)	(0.8950)	(0.5880)	(1.5062)
... Noninformational content only[a]	0.2272*	0.0121	−0.1325	−0.1504	−0.0619
	(0.0494)	(0.0348)	(0.3077)	(0.3085)	(0.2912)
... Both informational and Noninformational content[a]	0.0456	0.0999	0.3328	−0.0088	−0.1490
	(0.1899)	(0.1243)	(0.5423)	(0.4994)	(0.4565)
High-risk consumer × Spending per household in $ on...					
... Informational content only[a]	0.3511	−0.0012	−0.3511	−0.2105	−0.0240
	(0.2383)	(0.2634)	(0.6383)	(0.4642)	(0.6298)
... Noninformational content only[a]	0.0310	−0.1258	−0.1996	−0.1495	−0.2653
	(0.1408)	(0.1443)	(0.4762)	(0.2628)	(0.3534)
... Both informational and noninformational content[a]	0.1640**	0.0372	0.1221	0.2584***	0.1066****
	(0.0555)	(0.0364)	(0.1306)	(0.1017)	(0.0570)
Same insurer as in previous year (Yes = 1)					0.1935*
					(0.0203)
Insurer provided the best price (Yes = 1)					0.7264*
					(0.0241)
Brand-demographics-year fixed effects	Yes	Yes	Yes	Yes	Yes
Online-brand-demographics-year fixed effects	Yes	Yes	Yes	Yes	Yes
Brand-border-DMA fixed effects	Yes	Yes	Yes	Yes	Yes
Brand-border-month fixed effects	Yes	Yes	Yes	Yes	Yes
Number of observations	2,105,376	1,376,025	63,743	155,464	38,720
Effective number of observations[b]	2,078,790	1,348,641	40,242	117,424	19,994
Part B: Change in circumstances					
Spending per household in $ on ads with...					
... Informational content only[a]	−0.3451	0.5107****	0.0365	−0.9146	−0.4995
	(0.2190)	(0.2710)	(0.8523)	(0.5664)	(1.3873)
... Noninformational content only[a]	0.2208*	−0.0027	−0.2510	−0.3111	−0.0512
	(0.0503)	(0.0332)	(0.3014)	(0.3322)	(0.2864)
... Both informational and noninformational content[a]	0.0422	0.0918	0.3314	0.0008	−0.1579
	(0.1886)	(0.1243)	(0.5424)	(0.4983)	(0.4358)
Change in circumstances × Spending per household in $ on ads with...					
... Informational content only[a]	0.0855	0.7455**	−0.7644	−0.6859	0.8601
	(0.3207)	(0.2351)	(0.7497)	(0.3656)	(0.6370)
... Noninformational content only[a]	0.6150*	0.3822**	0.8504***	0.7182**	−0.2738
	(0.1542)	(0.1403)	(0.3331)	(0.2569)	(0.2849)
... Both informational and noninformational content[a]	0.1983*	0.0981**	0.1089	0.0695	−0.0007
	(0.0385)	(0.0329)	(0.0773)	(0.0547)	(0.0767)
Same insurer as in previous year (Yes = 1)					0.1936*
					(0.0200)
Insurer provided the best price (Yes = 1)					0.7263*
					(0.0242)
Brand-demographics-year fixed effects	Yes	Yes	Yes	Yes	Yes
Online-brand-demographics-year fixed effects	Yes	Yes	Yes	Yes	Yes
Brand-border-DMA fixed effects	Yes	Yes	Yes	Yes	Yes
Brand-border-month fixed effects	Yes	Yes	Yes	Yes	Yes
Number of observations	2,105,376	1,376,025	63,743	155,464	38,720
Effective number of observations[b]	2,078,790	1,348,641	40,242	117,424	19,994

Notes. Controls included: brand-state-month fixed effects, spending with missing content, and survey fixed effects. Standard errors in parentheses (clustered at the DMA level).

[a]Measured on a logarithmic scale.

[b]After dropping observations because of collinearity with fixed effects.

*< 0.001; **< 0.01; ***< 0.05; ****< 0.10.

Here, we re-estimate our models using the logarithm of DMA-level advertising spending per household as our measure of advertising intensity. The results are shown in the middle third of Table K-1 in Web Appendix K. Overall, our results are qualitatively robust to this alternative operationalization. And lastly, we investigate the robustness of our results with respect to our use of total advertising in the ad content regressions. Here, we re-estimate our models using the logarithm of DMA-level advertising spending per household on only informational ads, only noninformational ads, and both informational and noninformational ads. The results are shown in the lower third of Table K-1 in Web Appendix K and confirm that our findings are robust to this alternative operationalization.

8. Limitations and Future Research

There are several limitations to our paper and opportunities for future research. First, although we observe the shopping month, we do not observe the shopping date. This is a limitation of our data. Therefore, we estimate the effects of advertising in month $t-1$ on outcome variables in month t. However, consumers are also likely influenced by advertising during month t in the days prior to the shopping date. Thus, our estimates on the effects of advertising should be interpreted as lower bounds. Second, we estimate the effects of only informational, only noninformational, and both informational and noninformational ads. However, ads containing both informational and noninformational content could potentially be divided in at least three subgroups: ads containing more informational than noninformational content, ads containing equal amounts of informational and noninformational content, and ads containing less informational than noninformational content. We leave such a more detailed examination of ads with both informational and noninformational content for future research.

Third, our results for conditional consideration (and conditional purchase) are based on the assumption that the set of people who are aware of a brand is comparable across DMAs and months. However, people who are aware of a brand in DMAs with a lot of advertising might not be comparable to people who are aware of a brand in DMAs with little advertising. It is possible that people who become aware of a brand with a lot of advertising have a lower preference for that brand than people who are aware of the brand with no advertising. This introduces a negative correlation between conditional consideration (and conditional purchase) and advertising and the advertising coefficients might be biased downward in the conditional consideration and conditional purchase regressions. Fourth, our results are based on data from one industry. Although our results are broadly consistent with those found by previous literature for other financial services (when comparisons can be made), it is likely that the generalizability of our results decreases the further one moves away from the financial services sector. This represents a limitation of our data. And lastly, we have information on four content pieces: price and nonprice product features, brand name focus, and emotional appeal, which we aggregate to informational and noninformational content. Exploring how the effects vary across the four content pieces is left for future research.

9. Conclusion

Understanding how advertising influences consumers' decision making is crucial for companies. Marketing managers must not only decide how much to spend on advertising but also what to communicate to consumers in their advertisements. In this paper, we study how TV advertising quantities and TV advertising content affect each stage in the consumer purchase funnel in the context of the U.S. auto insurance industry. We find advertising content to matter—a lot. This finding should give pause to marketing managers and increase their focus on employing advertising content that is effective in achieving their marketing goals.

To summarize, our results show that advertising quantity primarily affects consumer awareness and has no discernible effects on conditional consideration or conditional purchase. However, when studying the influence of different types of advertising content, that is, ads with only informational content, ads with only noninformational content, and ads with both informational and noninformational content, we find a more nuanced set of results: advertising only containing noninformational content increases unaided awareness, whereas advertising only containing informational content increases aided awareness. We do not find significant effects of advertising containing both informational and noninformational content. Because many companies spend a significant portion of their budgets on advertising with both informational and noninformational content, we investigate whether this type of advertising influences certain groups of consumers. We find it to affect unaided (and, in some cases, aided) awareness of shoppers and other groups of relatively involved consumers, such as high-risk consumers and consumers with a change in circumstances.

Acknowledgments

The authors thank J.D. Power and Associates for supplying them with data. The authors especially thank Pradeep Chintagunta, Brett Hollenback, Peter Rossi, Stephen Spiller, and Stephan Seiler for extensive and helpful discussions. The authors also thank Xiaojing Dong, Günter Hitsch, Sylvia Hristakeva, Robert Kent, Jura Liaukonyte, Carl Mela, David Muir, Thomas Otter, Brad Shapiro, Matt Shum, Jeremy Tobacman, Ken Wilbur, and Nathan Wilson; seminar participants at Duke University, INSEAD, Rotterdam School of Economics, University of Michigan, and Tulane University; the paper's discussants Pedro Gardete, Tatsuo Tanaka, and Zsolt Sandor; and the participants of the 2017 International Industrial Organization Conference, the 8th Workshop on Search and Switching Costs, the 2017 Boulder Summer Conference on Consumer Financial Decision Making, the 2017 Marketing Science/INFORMS conference, the 2018 Frank M. Bass University of Texas, Dallas Frontiers of Research in Marketing Science (FORMS) conference, the 2018 Marketing Dynamics conference, and the 11th Workshop on the Economics of Advertising for their comments. The feedback from two anonymous reviewers, an associate editor, and the senior editor is also gratefully acknowledged. All errors are the authors' own.

Endnotes

[1] For example, see Ackerberg (2001), Ackerberg (2003), Narayanan et al. (2005), Ching and Ishihara (2012), Chan et al. (2013), Lovett and Staelin (2016), Hastings et al. (2017), Shapiro (2018), and Shapiro et al. (2021).

[2] By definition, the brand name is mentioned/shown in all ads. The content type "brand name focus" does not capture the simple mention of the brand name in an ad. It captures the *focus* on the brand name in an ad, for example, frequent repetitions/showings of the brand name in a TV ad, multiple/large prints of the brand name in a magazine ad, ads with no product-related information.

[3] In this paper, we use the terms "consider," "search," and "shop" interchangeably.

[4] In this paper, we view shopping, that is, requesting at least one price quote from an insurance company that is *not* the consumer's current insurance provider, as a prerequisite for an (active) purchase decision that involves deciding whether to switch auto insurers.

[5] Some information is (logically) unavailable for nonshoppers. For example, because nonshoppers do not shop, they do not form consideration sets and do not make an (active) purchase decision but remain passively insured with the same insurance company. Further, because nonshoppers do not shop, we have no information on their shopping month and (quoted) insurance premia.

[6] Throughout this paper, we refer to unobservables that are specific to the brand, to the year, and to the brand-year as "global" unobservables.

[7] There is little empirical work on the complementary (Stigler and Becker 1977, Becker and Murphy 1993) and signaling views (Nelson 1970, Nelson 1974) of advertising. Recent exceptions are Tuchman et al. (2018) for complementarity and Sahni and Nair (2020) for signaling.

[8] The data are also different. Whereas Honka et al. (2017) only have data on consumer (aided) awareness, consideration, and purchase for one year and only data on advertising quantities, our data used in this paper span a time period of seven years and also includes unaided awareness. Further, we not only have information on advertising quantities but also on advertising content.

[9] Technically, we do not estimate the fixed effects but difference them out in the estimation. The demographic groups for which we estimate fixed effects are as follows: age <25 years, age between 25 and 45 years, age between 45 and 65 years, male, shopped for homeowner insurance, has college degree, income of more than $100K.

[10] We mark a brand as using an advertising type in a year if it spent more than 1.5% of its annual advertising spending on that advertising type.

[11] If Assumption 1 does *not* hold, that is, if brands employ the type(s) of advertising content that is/are most effective for them, the estimated effects of advertising content are likely biased upwards.

[12] Note that this approach also addresses endogeneity concerns due to time-invariant, brand-specific local unobservables. Brand-DMA fixed effects are subsumed in the brand-border-DMA fixed effects estimated in the regression discontinuity approach.

[13] The regression discontinuity approach would not work with a percentage specification because the percentage of one ad type increases when the percentage of another ad type decreases.

[14] Local promotions are uncommon in the auto insurance industry. The reason is that this industry is heavily regulated at the state level. For example, rating schedules that determine all insurance premia have to be submitted to state insurance commissioners and are publicly available. Insurance companies cannot change insurance rates without submitting the change to state insurance commissioners. This makes the process cumbersome and slow, and most insurance companies do not make changes to their rating schedules more than twice a year. Furthermore, many DMAs span multiple states; communicating a rate decrease that applies to a part but not the whole DMA would be difficult. Because of these issues, insurance companies usually advertise how much people save by switching when trying to promote that they have low insurance rates. Such promoted savings are typically *not* localized.

[15] We also calculated the RV and Procustes coefficients (Josse and Holmes 2016). Their values are 0.62 and 0.79 (both $p < 0.001$), respectively. Using data from the four largest brands only, the correlation based on the distance covariance, the RV, and the Procustes coefficients are 0.84, 0.68, and 0.79, respectively, that is, the distributions of national and local advertising content are even slightly more similar among the top four brands.

[16] J.D. Power and Associates measures unaided awareness, aided awareness, consideration, purchase, previous insurer, and price as follows: unaided awareness – "When you are thinking of auto and home insurance, which companies come to mind?"; aided awareness – "Please review the list below and select ALL the insurance companies that you recognize."; consideration – "From which of the following insurance companies did you receive a quote?"; purchase – "Which company is your current auto insurer?"; previous insurer – "Which company was your auto insurer prior to [pipe in current insurer]?"; price – "Which auto insurer offered you the lower price/premium?".

[17] Alternatively, one could rely on machine learning algorithms to code advertising content. We decided not to do that for several reasons: For example, creatives are deposited in different formats (e.g., gif, swf, and flm for an internet display advertisement), different resolutions, and in different languages (English and Spanish). Based on our observation, well-trained research assistants are more effective in identifying content types and picking up noninformational cues, such as humor. To predict whether a human will find a photograph or video funny/entertaining remains a challenging machine learning task. Computational humor is sometimes considered to be an "AI-complete" problem (Binsted et al. 2006). Although there has been some progress in identifying visual humor (e.g., Chandrasekaran et al. 2016), we decided to use trained research assistants because auto insurance advertisements often are nuanced and specific to the social context. For example, a cartoon in which a car is destroyed by a superhero fight (such as in a TV commercial by Mercury) could be considered funny, but a car damaged completely in real life is typically horrifying.

[18] Because nonshoppers do not shop, they do not form consideration sets and do not make an (active) purchase decision but remain passively insured with the same insurance company.

[19] Across the different media channels, auto insurance brands spend 80% of their total advertising budget on TV, 7% on internet (display advertising), 6% on the radio, 7% on print, and less than 1% on outdoor advertising.

[20] Note that the percentage of spending on local TV advertising in the auto insurance industry is *above average* compared with many other brands/categories: Shapiro et al. (2021) study advertising effects across 288 brands using the regression discontinuity approach. They find that, on average, brands spend 8.7% of their TV advertising expenditures on local TV advertising with a median of 3.7%. In Shapiro (2021), the percentage of TV advertising spent on local advertising is 7% in the antidepressants category.

[21] The percentage of shoppers increased from 29% in 2010–2012 to 36% in 2014–2016.

[22] See https://www.comscore.comInsightsPress-Releases201511comScore-Releases-2015-US-Online-Auto-Insurance-Shopping-Report; https://www.huffingtonpost.comentrypaying-too-much-for-auto-insurance-many-consumers_us_58c2dbede4b070e55af9ee2b; https://www.mckinsey.commediamckinseydotcomclient_serviceFinancial%20ServicesLatest%20thinkingInsuranceWinning_share_and_customer_loyalty_in_auto_insurance.ashx.

[23] The percentage of switchers among shoppers increased from 45% in 2010–2012 to 49% in 2014–2016.

[24] Note that these shares were calculated based on the available information in our data, that is, the number of individuals who bought an insurance policy from a brand. In contrast, many outlets and websites publish market shares based on each brand's revenue.

[25] Twenty-four linear probability models = 2 ad kinds (advertising quantity, advertising content) x 2 endogeneity types (described in Sections 3.1 and 3.2) × 6 stages (unaided awareness, aided awareness, consideration conditional on unaided awareness, consideration conditional on aided awareness, conditional purchase, unconditional purchase).

[26] The advertising quantity variable is defined as the sum of spending on advertising with only informational content, advertising with only noninformational content, advertising with both informational and noninformational content, and advertising with missing content.

[27] The brand-border-DMA fixed effects subsume the brand-DMA fixed effects.

[28] Our results—in terms of which coefficients are significant and which ones are not—are robust to alternative clusterings at the individual and at the DMA-border level.

[29] For this analysis, we classify a brand as employing a type of advertising content if its spending on it is at least 1.5% over the whole study period (based on Table 3). Given this condition, four brands in our data utilize all three types of advertising content: Allstate, Esurance, Nationwide, and State Farm.

[30] See, for example, http://www.adweek.com/brand-marketing/advertisers-need-to-stop-chasing-engagement-and-get-back-to-focusing-on-awareness/?utm_content=buffer42f1f&utm_medium=social&utm_source=facebook.com&utm_campaign=buffer.

[31] For example, see 2012 McKinsey Auto Insurance Customer Insights Research Report (https://www.mckinsey.commediamckinseydotcomclient_serviceFinancial%20ServicesLatest%20thinkingInsuranceWinning_share_and_customer_loyalty_in_auto_insurance.ashx).

[32] Whether to shop is likely an endogenous decision made by consumers. Thus, we interpret the results of the shoppers versus nonshoppers analysis as correlational.

[33] We do not show the results for conditional consideration and conditional purchase in Table 11 because they are identical to those in Table 8 since nonshoppers, by definition, do not shop and thus the conditional consideration and conditional purchase regressions are estimated only using shoppers.

[34] Group membership, that is, being a high-risk consumer or a consumer with a change in circumstances, is determined by events that are exogenous to auto insurance brand choice.

[35] The overlap between high-risk consumers and consumers with a change in circumstances is very limited: only 1.4% of consumers belong to both groups. Obviously, there is also some overlap between these two groups of consumers and shoppers: 31% of shoppers are also high-risk and/or had a change in circumstances.

References

Ackerberg D (2001) Empirically distinguishing informative and prestige effects of advertising. *RAND J. Econom.* 32(2):316–333.

Ackerberg D (2003) Advertising, learning, and consumer choice in experience good markets: An empirical examination. *Internat. Econom. Rev.* 44(3):1007–1040.

Agarwal S, Ambrose B (2018) The effect of advertising on home equity credit choices. Teitelbaum J, Zeiler K, eds. *Research Handbook on Behavioral Law and Economics* (Edward Elgar, Northampton, MA), 122–151.

Alba J, Hutchinson J, Lynch J. (1991) Memory and decision making. Kassarjian H, Robertson T, eds. *Handbook of Consumer Behavior* (Prentice Hall, Hoboken, NJ), 1–49.

Allenby G, Rossi P (2019) Inference for marketing decisions. Dube J-P, Rossi P, eds. *Handbook of the Economics of Marketing* (North-Holland, Amsterdam), 69–150.

Anderson S, Ciliberto F, Liaukonyte J, Renault R (2016) Push-me pull-you: Comparative advertising in the OTC analgesics industry. *RAND J. Econom.* 47(4):1029–1056.

Angrist J, Pischke J-S (2009) *Mostly Harmless Econometrics: An Empiricist's Companion*, 1st ed. (Princeton University Press, Princeton, NJ).

Bagozzi R, Silk A (1983) Recall, recognition, and the measurement of memory for print advertisements. *Marketing Sci.* 2(2):95–134.

Bagwell K (2007) The economic analysis of advertising. Armstrong M, Porter RH, eds. *The Handbook of Industrial Organization*, vol. 3 (North-Holland, Amsterdam), 1701–1844.

Becker G, Murphy K (1993) A simple theory of advertising as a good or bad. *Quart. J. Econom.* 108:942–964.

Bertrand M, Karlan D, Mullainathan S, Shafir E, Zinman J (2010) What's advertising content worth? Evidence from a consumer credit marketing field experiment. *Quart. J. Econom.* 125(1):263–306.

Binsted K, Nijholt A, Stock O, Strapparava C, Ritchie G, Manurung R, Pain H, Waller A, O'Mara D (2006) Computational humor. *IEEE Intelligent Systems* 21(2):59–69.

Blattberg R, Deighton J (1996) Manage marketing by the customer equity test. *Harvard Bus. Rev.* 74(4):136–144.

Blattberg R, Getz G, Thomas J (2001) *Customer Equity: Building and Managing Relationships As Valuable Assets*, 1st ed. (Harvard Business Review Press, Boston).

Cameron C, Trivedi P (2005) *Microeconometrics: Methods and Applications*, 1st ed. (Cambridge University Press, New York).

Celsi R, Olson J (1988) The role of involvement in attention and comprehension processes. *J. Consumer Res.* 15(2):210–224.

Chamberlin E (1933) *The Theory of Monopolistic Competition: A Re-Orientation of the Theory of Value* (Harvard University Press, Cambridge, MA).

Chan T, Narasimhan Ch, Xie Y (2013) Treatment effectiveness and side effects: A model of physician learning. *Management Sci.* 59(6):1309–1325.

Chandrasekaran A, Vijayakumar A, Antol S, Bansal M, Batra D, Zitnick C, Parikh D (2016) We are humor beings: Understanding and predicting visual humor. *Proc. IEEE Conf. Comput. Vision Pattern Recognition* (IEEE, Piscataway, NJ), 4603–4612.

Ching A, Ishihara M (2012) Measuring the informative and persuasive roles of detailing on prescribing decisions. *Management Sci.* 58(7):1374–1387.

Cronqvist H (2006), Advertising and portfolio choice. Working paper, Ohio State University, Columbus, OH.

Cunningham S (2020) *Causal Inference*, 1st ed. (Yale University Press, New Haven, CT).

Finn A (1992) Recall, recognition, and the measurement of memory for print advertisements: A reassessment. *Marketing Sci.* 11(1):95–100.

Gillund G, Shiffrin R (1984) A retrieval model for both recognition and recall. *Psych. Rev.* 91(1):1–67.

Goeree M (2008) Limited information and advertising in the US personal computer industry. *Econometrica* 76(5):1017–1074.

Gurun U, Matvos G, Seru A (2016) Advertising expensive mortgages. *J. Finance* 71(5):2371–2416.

Hastings J, Hortaçsu A, Syverson CH (2017) Advertising and competition in privatized social security: The case of Mexico. *Econometrica* 85(6):1723–1761.

Heath Ch, Heath D (2008) *Made to Stick*, 2nd ed. (Random House, New York).

Honka E (2014) Quantifying search and switching costs in the U.S. auto insurance industry. *RAND J. Econom.* 45(4):847–884.

Honka E, Tsai Y (2019) A descriptive analysis of advertising content: Insights from the U.S. auto insurance industry. Working paper, University of California, Los Angeles.

Honka E, Hortaçsu A, Vitorino M (2017) Advertising, consumer awareness, and choice: Evidence from the U.S. banking industry. *RAND J. Econom.* 48(3):611–646.

Israel M (2005) Tenure dependence in consumer-firm relationships: An empirical analysis of consumer departures from automobile insurance firms. *RAND J. Econom.* 36(1):165–192.

Josse J, Holmes S (2016) Measuring multivariate association and beyond. *Statist. Surveys* 10:132–167.

Lee D, Hosanagar K, Nair H (2018) Advertising content and consumer engagement on social media: Evidence from Facebook. *Management Sci.* 64(11):5105–5131.

Leigh J, Zinkhan G, Swaminathan V (2006) Dimensional relationships of recall and recognition measures with selected cognitive and affective aspects of print ads. *J. Advertising* 35(1):105–122.

Lewis R, Rao J (2015) The unfavorable economics of measuring the returns to advertising. *Quart. J. Econom.* 130(4):1941–1973.

Liaukonyte J, Teixeira T, Wilbur K (2015) Television advertising and online shopping. *Marketing Sci.* 34(3):311–330.

Lovett M, Staelin R (2016) The role of paid and earned media in building an entertainment brand: Reminding, informing, and enhancing enjoyment. *Marketing Sci.* 35(1):142–157.

Lynch J, Srull T (1982) Memory and attentional factors in consumer choice: Concepts and research methods. *J. Consumer Res.* 9(1): 18–37.

Maddala G (1983) *Limited-Dependent and Qualitative Variables in Econometrics*, 1st ed. (Cambridge University Press, New York).

McCloskey M, Glucksberg S (1979) Decision processes in verifying category membership statements: Implications for models of semantic memory. *Cognitive Psych.* 11(1):1–37.

Mehta A, Purvis S (2006) Reconsidering recall and emotion in advertising. *J. Advertising Res.* 46(1):49–56.

Mullainathan S, Schwartzstein J, Shleifer A (2008) Coarse thinking and persuasion. *Quart. J. Econom.* 123(2):577–619.

Narayanan S, Manchanda P, Chintagunta P (2005) Temporal differences in the role of marketing communication in new product categories. *J. Marketing Res.* 42(3):278–290.

Nelson P (1970) Information and consumer behavior. *J. Political Econom.* 78(2):311–329.

Nelson P (1974) Advertising as information. *J. Political Econom.* 82(4): 729–754.

Pires T (2016) Costly search and consideration sets in storable goods markets. *Quant. Marketing Econom.* 14(3):157–193.

Raaijmakers J, Shiffrin R. (1981) Order effects in recall. Long A, Baddeley A, eds. *Attention and Performance*, vol. 9 (Erlbaum, Hillsdale, NJ), 403–415.

Ratcliff R, Murdock B (1976) Retrieval processes in recognition memory. *Psych. Rev.* 83(1):190–214.

Reinartz W, Thomas J, Kumar V (2005) Balancing acquisition and retention resources to maximize customer profitability. *J. Marketing* 69(1):63–79.

Resnik A, Stern B (1977) An analysis of information content in television advertising. *J. Marketing* 41(1):50–53.

Rust R, Lemon K, Zeithaml V (2004) Return on marketing: Using customer equity to focus marketing strategy. *J. Marketing* 68(1): 109–127.

Sahni N, Nair H (2020) Does advertising serve as a signal? Evidence from a field experiment in mobile search. *Rev. Econom. Stud.* 87(3):1529–1564.

Sahni N, Wheeler CH, Chintagunta P (2018) Personalization in email marketing: The role of non-informative advertising content. *Marketing Sci.* 37(2):236–258.

Shapiro B (2018) Positive spillovers and free riding in advertising of prescription pharmaceuticals: The case of antidepressants. *J. Political Econom.* 126(1):381–437.

Shapiro B (2021) Promoting wellness or waste? Evidence from antidepressant advertising. *Amer. Econom. J. Microeconom.* Forthcoming.

Shapiro B, Hitsch G, Tuchman A (2021), TV advertising effectiveness and profitability: Generalizable results from 288 brands. *Econometrica*. Forthcoming.

Smith S, Glenberg A, Bjork R (1978) Environmental context and human memory. *Memory Cognition* 6(4):342–353.

Stern B, Krugman D, Resnik A (1981) Magazine advertising: An analysis of its information content. *J. Advertising Res.* 21(2):39–44.

Stigler G, Becker G (1977) De gustibus non est disputandum. *Amer. Econom. Rev.* 67(2):76–90.

Thorsen E (1991) Likeability: Ten years of academic research. *Eighth Annual ARF Copy Research Workshop*.

Till B, Baack D (2013) Recall and persuasion: Does creative advertising matter? *J. Advertising* 34(3):47–57.

Tuchman A (2019) Advertising and demand for addictive goods: The effects of E-cigarette advertising. *Marketing Sci.* 38(6):913–1084.

Tuchman A, Nair H, Gardete P (2018) Television ad-skipping, consumption complementarities and the consumer demand for advertising. *Quant. Marketing Econom.* 16(2):111–174.

Tucker C (2015) The reach and persuasiveness of viral video ads. *Marketing Sci.* 34(2):281–296.

Villanueva J, Hanssens D (2007) Customer equity: Measurement, management and research opportunities. *Foundations Trends Marketing* 1(1):1–95.

Wachtel S, Otter Th (2013) Successive sample selection and its relevance for management decisions. *Marketing Sci.* 32(1):170–185.

Woodward A, Bjork R, Jongeward R (1973) Recall and recognition as a function of primary rehearsal. *J. Verbal Learn. Verbal Behav.* 12(6):608–612.

Wooldridge J (2001) *Econometric Analysis of Cross Section and Panel Data*, 1st ed. (MIT Press, Cambridge, MA).

Zaichkowsky J (1985) Measuring the involvement construct. *J. Consumer Res.* 12(3):341–352.