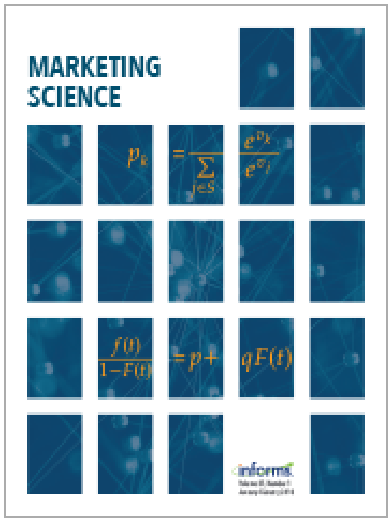


Marketing Science

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

Advertising in Health Insurance Markets

Bradley T. Shapiro

To cite this article:
Bradley T. Shapiro (2020) Advertising in Health Insurance Markets. Marketing Science 39(3):587-611. https://doi.org/10.1287/mksc.2018.1086

informs
http://pubsonline.informs.org/journal/mksc/
MARKETING SCIENCE
Vol. 39, No. 3, May–June 2020, pp. 587–611
ISSN 0732-2399 (print), ISSN 1526-548X (online)

Advertising in Health Insurance Markets

Bradley T. Shapiro[a]

[a] University of Chicago, Chicago, Illinois 60637

Contact: bradley.shapiro@chicagobooth.edu, http://orcid.org/0000-0003-1938-6727 (BTS)

Received: March 5, 2017
Revised: July 1, 2107; November 5, 2017
Accepted: November 7, 2017
Published Online in Articles in Advance: July 18, 2018

https://doi.org/10.1287/mksc.2018.1086

Abstract. The effects of television advertising in the market for health insurance are of distinct interest to both firms and regulators. Regulators are concerned about firms potentially using ads to "cream skim," or attract an advantageous risk pool, as well as the potential for firms to use misinformation to take advantage of the elderly. Firms are interested in using advertising to acquire potentially highly profitable seniors. Meanwhile, health insurance is a useful setting to study the mechanisms through which advertising could work. Using the discontinuity in advertising exposure created by the borders of television markets, this study estimates the effects of advertising on consumer choice in health insurance. Television advertising has a small effect on brand enrollments, making advertising a relatively expensive means of acquiring customers. Heterogeneous effects point to advertising being more effective in less healthy counties, which runs opposite to the concern of cream skimming. Leveraging the unilateral cessation of advertising by UnitedHealthcare, evidence is provided that the small advertising effect is not explained by a prisoner's dilemma equilibrium. An analysis of longer-run effects of advertising shows that advertising effects are short lived, further decreasing the potential of advertising to create long-run value to the firm.

History: Puneet Manchanda served as the senior editor and Carl Mela served as associate editor for this article. This paper has been accepted for the *Marketing Science* Special Issue on Health.
Funding: The author gratefully acknowledges research support from the Beatrice Foods Company Faculty Research Fund at the University of Chicago Booth School of Business.
Supplemental Material: Data are available at https://doi.org/10.1287/mksc.2018.1086.

Keywords: advertising • insurance • health • Medicare Advantage • adverse selection

1. Introduction

Television advertising by health insurance plans is large and growing, rising from about $200 million in 2004 to $400 million in 2012. With the implementation of the Affordable Care Act (ACA) marketplaces and a more broad-based shift toward health plan choice, television advertising by health insurance plans is expected to continue to grow.

Relative to other markets, health insurance advertising faces increased regulatory scrutiny. Historically, the majority of advertising has been for Medicare Advantage (MA) plans, which provide seniors with a private, government-subsidized alternative to traditional Medicare (TM). While the industry does not advertise as much as traditionally dominant advertisers like automobiles or prescription drugs, it spends more than categories such as hair products, air travel, and computers.[1] Regulators in this market have expressed concern that advertising might be used to "cream skim" lower-cost enrollees, inflating government costs, and they have expressed concern about the potential for misleading advertisements, which might induce seniors into purchasing plans they do not need or that are poorly suited to their preferences.[2] As television is difficult to target carefully geographically, the main concern of regulators with regards to targeting is that the advertising copy that makes ads seem especially relevant to the healthy but not the unhealthy. A Kaiser Family Foundation report on MA advertising in 2008 (Cai et al. 2008, p. 2), which reviewed a large sample of MA ads, found that "None of the ads included images of seniors who appeared to be sick or physically frail (e.g., using a walker or cane); only 14 percent of all ad occurrences included images of prescription drugs."

There also exist reasons to think that MA advertisements—and health insurance advertising more generally—might have important benefits. Health insurance advertising could, of course, serve the standard informative function, helping seniors choose plans that better reflect their own preferences. In addition, because MA plans, and many other health plans, are purchased during an open enrollment period, advertising might have the added benefit of alerting consumers about enrollment deadlines. Because of the well-documented inertia in health plan choice (e.g., Handel 2013), advertising might be particularly useful in making consumers aware of other plans, thereby intensifying competition in the market.

This paper takes a first step toward understanding the role of advertising in health insurance markets by estimating the impact of television advertising on MA enrollments, characterizing heterogeneity in these effects, and testing mechanisms through which advertising might work. In particular, this study tests whether advertising is especially useful in expanding the category by moving seniors out of TM and into MA and characterizes whether such expansion happens disproportionately by county characteristics, including health status. Additionally, this paper tests whether competitive advertising cancels out in a kind of prisoner's dilemma equilibrium and whether advertising has long-lived effects.

Television advertising is measured using AC Nielsen's media database, which provides spot-level information on television advertising from 2004 to 2012. MA enrollment is measured at the contract-year level using administrative data from the Centers for Medicare and Medicaid Services (CMS) from 2007 to 2012. These data provide the number of potential MA enrollees for each county, enrollment totals for each plan in each county, and information on premiums, coinsurance, and other characteristics for each plan and county.

Sharp discontinuities in the level of advertising at the borders of geographically based television markets provide exogenous variation in advertising, as in Shapiro (2018), to assess the extent to which advertising increases the demand for an MA plan over either TM or another MA plan, observable heterogeneity in advertising effects by health status, and other factors that are informative to both regulators and managers.

Using this variation, I find a small average lift of brand MA advertising on demand. The point estimate from the preferred specification is statistically significant but small, with an implied cost per conversion (CPC) of about \$1,300. The 95% CPC confidence interval (CI) ranges from \$660 to \$43,300, which rules out very profitable effects of advertising. Meanwhile, a more naïve approach would suggest advertising is a very inexpensive means of converting enrollees, and rival advertising leads to a material reduction in demand.

Additionally, I find that the effect of advertising on moving seniors from TM into MA (category expansion) is small in magnitude. The preferred specification suggests that removing all MA advertising would decrease MA demand by only 0.23 percentage points, from 11.95% to 11.72% of eligible seniors. At the right edge of the 95% confidence interval, eliminating MA advertising results in a drop in MA demand of only 0.53 percentage points. A more naïve approach would produce an effect more than an order of magnitude larger. Given the small main effects, any heterogeneous effects that alter the risk pool in MA must be small.

Indeed, when examining how ad effects vary with observable characteristics, no statistically meaningful relationship is detected between advertising effectiveness and the average health risk of a county. The point estimates suggest that advertising works slightly better on less healthy counties, which works in the opposite direction of cream skimming. Advertising effectiveness is also correlated with lower income, a higher share of elderly population, and a higher share of Asian population.

If advertising works primarily to steal business, firms must keep advertising to maintain the status quo share and avoid competitors stealing their enrollees in a kind of prisoner's dilemma. In other words, advertising and rival advertising could be strategic complements, leading to increased advertising by all with little change or no change in market shares. I address this question by leveraging the unilateral cessation of advertising from 2008 to 2010 by UnitedHealthcare, one of the largest players in this market, providing a direct test of the consequences of removing advertising on brand share. The estimates suggest that no material loss of brand share occurred, providing evidence against the prisoner's dilemma hypothesis. This also provides evidence against the need for TV advertising to enhance other marketing levers, as a cessation in TV advertising would lead to a loss of market share in that case.

Finally, given the sticky nature of health insurance purchase, advertising might be expected to have longer-run effects, as seniors who sign up in one year because of advertising are likely to stick around for subsequent years, which would make a \$1,300 CPC more reasonable. Just how sticky the advertising marginal customers are is important in determining whether the advertising marginal customers provide more or less lifetime value than average customers. Using a stock conception of advertising and an assortment of assumed rates of ad stock persistence, I show that the advertising effect is small, statistically insignificant, and precisely estimated, providing evidence against large long-run effects of advertising on demand, either through state dependence or through advertising carryover. In particular, these results are consistent with advertising marginal customers being lower lifetime value customers than average customers.

Combined, these results suggest that while firms might be trying to cream skim using advertising, they are not particularly successful at drawing a large number of healthy (or any, for that matter) consumers. As such, concerns over the cream skimming and misinformation may be overblown. Of course, these small effects might be due to current regulatory attention. Advertising might affect consumer choices in potentially undesirable ways if regulatory attention were to be reduced. However, given that advertising has a very small effect, additional regulatory scrutiny seems unwarranted. On the firm side, these estimates show that advertising is an expensive means of customer acquisition and provide some guidance on targeting.

While the regulatory implications might be clear, a puzzle remains. If television advertising is ineffective relative to other means of customer acquisition in the short and the long run, why are firms spending hundreds of millions of dollars per year on television advertising? That advertising increases from 2004 to 2012 suggests that firms are not learning over time that advertising is ineffective.[3] It could be that firms have a high cost of measuring their own advertising effectiveness or that there is a difficult-to-overcome principal–agent problem with advertising agencies (Vranica 2016), though it is difficult to say definitively using these data. Both the difficulty in measuring advertising effects for firms (e.g., Lewis and Rao 2015) and that firms as a consequence could make systematic mistakes in advertising strategy (e.g., Blake et al. 2015) have been documented in the advertising effectiveness literature.

The contribution of this paper is an empirical one in the context of an important market. The results of this study also informs our priors on the usefulness of advertising for selective targeting more broadly, as well as the usefulness of advertising in markets for goods with complicated, infrequent choices. While there is some recent research studying advertising targeting in MA markets, this is the first paper that estimates the causal effect of advertising using a natural experiment in the market for health insurance. Aizawa and Kim (2018) shows that firms tend to target advertisements toward healthier consumers, while Mehrotra et al. (2006) find that ad content is targeted toward healthy patients, giving some credence to the regulatory concern. Duggan et al. (2016) find that firms advertise more in markets where the government pays them more per enrollee. Aizawa and Kim (2018) explore how market equilibria are affected if firms can use marketing levers to risk select.[4]

This paper also contributes to literatures on competition in health insurance (Dafny 2010) and in the MA market (e.g., Curto et al. 2015, Cabral et al. 2018, Town and Liu 2003). In particular, Ericson (2014) finds that default plans are persistent, so firms are more likely to offer new plans rather than lower prices. Cooper and Trivedi (2012) find that firms try to gain advantageous selection by offering plan characteristics such as gym memberships. Abaluck and Gruber (2011) find that consumers could save a significant amount of money by switching to the lowest-cost prescription drug plan for them. Despite evidence of cream skimming and social misallocation, this paper shows evidence that further regulation of television advertising is unlikely to solve any of these problems.

This paper also adds to a growing literature on advertising effectiveness (e.g., Johnson et al. 2016, Shapiro 2018, Blake et al. 2015). On the measurement side, all of these studies have shown that failure to consider the endogeneity of advertising can lead to large biases in estimated advertising effects, usually in the upward direction. Documentation of this problem along with attempts at a solution date back at least to the work of Eastlack and Rao (1986), who ran a set of controlled experiments on V-8 vegetable juice advertising. Later, Lodish et al. (1995) used split cable experiments across a number of brands to think about advertising effectiveness more broadly. Additionally, a recent stream of literature has used field experiments online (Blake et al. 2015; Johnson et al. 2016; Sahni 2015a, b; Lewis and Nguyen 2015), instrumental variables (IVs) (e.g., Sinkinson and Starc 2017, Hartmann and Klapper 2018), or natural experiments (e.g., Shapiro 2018, Tuchman 2015, Spenkuch and Toniatti 2018). This paper will follow the third strategy and exploit the random nature of TV market borders.

The use of borders as a source of variation is related to a small but growing literature. Spatial strategies have been used to identify the effects of minimum wages (Dube et al. 2010), the effects of right-to-work laws (Holmes 1998), the effects of schools on home values (e.g., Black 1999, Bayer et al. 2007), and the response of households to changes in electricity prices (Ito 2014). While many of these studies exploit state borders, that is unattractive in this setting, as many health-policy-related factors vary across states. As such, this study will focus on within-state comparisons across the borders of television markets. Additionally, since some designated market areas (DMAs) have few counties, this study will only use those border areas that make up less than 35% of the counties in the DMA.

Since the main result of this paper is a very small effect of advertising, this paper might be seen as presenting a null effect. However, it is an important and informative result with regulatory and managerial implications. In particular, the documented advertising effect is precisely estimated. Indeed, the confidence interval shows that even viewed optimistically, advertising is a relatively expensive means to acquire business. As regulators work with limited resources to find interventions that work, knowing which ones will not work or that will work in only a very limited fashion is necessary. Additionally, advertising spending is in the hundreds of millions of dollars per year in health insurance, making the documentation of a negative result important for firms.

The rest of this paper proceeds as follows. Section 2 describes the markets for advertising and health insurance for the elderly. Section 3 describes the data. Section 4 explains the research design in detail. Section 5 documents the results, and Section 6 provides general discussion and concludes.

2. Background

2.1. Health Insurance for Seniors

Nearly all seniors at or above 65 years old in the United States receive health insurance coverage under the Medicare program. Historically, most seniors have enrolled in what is now called traditional Medicare, which is a public insurance program administered by the CMS. Beneficiaries can go to any provider who is willing to see them. The program is fee-for-services, meaning the providers are paid according to the medical services they provide. In addition to premiums, beneficiaries have to pay deductibles and coinsurance, or purchase supplemental Medigap coverage to cover this cost sharing.

Medicare Advantage was established in the early 1980s to provide a private alternative to TM coverage.[5] MA plans are differentiated from TM in having restricted provider networks, alternative cost-sharing arrangements, and additional benefits, such as vision and dental coverage. MA plans have historically been offered by health maintenance organizations (HMOs). Plans receive a capitation payment from Medicare for each enrolled beneficiary and often charge beneficiaries a supplemental premium. Premiums are determined by the relationship between plan bids to the government and statutory benchmark rates. Since 2000, MA enrollment has risen considerably, from nearly 0% to over 30% of Medicare beneficiaries, fueled in part by legislation that has increased payment to plans and lifted restrictions on entry by non-HMO plans. See McGuire et al. (2011) for an in-depth history of the MA program.

There are six large national firms that make up around 65% of the total MA share: UnitedHealthcare, Aetna, Humana, Cigna, Kaiser Permanente, and Blue Cross and Blue Shield (BCBS). While UnitedHealthcare and BCBS have strength in many markets across the country, other plans have more geographically concentrated historical strength. In addition to these large national firms, there exist more geographically concentrated local plans in many markets.

For most of Medicare's history, very few enrollees had prescription drug coverage. Some had coverage through their MA plan, and some purchased supplemental insurance with drug coverage, but the majority of seniors paid for most prescription drugs out of pocket. The 2004 Medicare Modernization Act changed this with the creation of Medicare Part D. Starting in 2006, seniors with TM could enroll in subsidized, private Part D plans, and seniors who enrolled in MA could use their subsidies for MA plans that provided drug coverage.

2.2. Television Advertising

Firms can purchase advertising space on television in two ways. First, there is an up-front market each summer where advertising agencies and firms make deals for the upcoming year of television. Advertising purchased in the up-front market cannot be "returned" and typically has minimal flexibility in terms of timing, though there is a secondary market that firms sometimes use to offload unneeded advertising space. There is also a spot market, where firms can purchase advertising closer to the date aired.

Ads may be purchased for local or national television. Local advertisements are seen only by households within a particular DMA. A DMA is a collection of counties, typically centered around a major city, and it is defined by the global market research firm AC Nielsen. The DMAs were first defined to allow for the sale of advertising in a way that was straightforward to the advertisers. The DMA location of a county determines which local television stations a consumer of cable or satellite dish gets with her subscription. The original idea was to place counties into the same DMA with the local television station that most people wanted to watch, which oftentimes was just the station that was easiest to pick up over the air; that is, if a county picked up the Cleveland stations over the air more easily than the Columbus stations, it would be placed in the Cleveland DMA. Existing laws and regulations in most circumstances do not allow satellite or cable operators to provide broadcast signals from outside of the DMA in which they reside.[6] Even for over-the-air signals, the Federal Communications Commission moderates the signals to try to keep the signal from each station localized only in its own DMA.[7] There are 210 DMAs in the United States. National advertisements are, in principle, meant to be seen by everyone in the country tuned into a particular network station. However, local affiliate stations do have the leeway to bump national ads in favor of additional programming or local ads, generating some local variation in national ads.

3. Data and Summary Statistics

3.1. Advertising

We use advertising data from AC Nielsen's media database from 2004 to 2012 in this study. The database tracks television advertising at the spot-time-DMA level for every product that advertises on television. The top 130 out of 210 DMAs are indicated as "full discovery markets" by AC Nielsen, meaning all television advertising occurrences are measured using monitoring devices. In many of the smaller DMAs, only advertising occurrences that match ads in the larger markets are included. This study uses each of these full discovery markets that has a monitoring device on every major network affiliate (ABC, NBC, CBS, and FOX), which is 120 DMAs.

In the top 25 DMAs, household impressions are measured from set-top viewing information that is recorded in a random subset of households. In DMAs ranked 26–210, advertising impressions are estimated from quarterly diaries filled out by a random subset of households. While impressions are the main advertising measure of interest, there is some concern that the infrequent and self-reported viewing data may be measured with error. In the appendix, all analyses will be

Table 1. Advertising and Shares by Brand per Year

Brand	Average GRP	% local	Brand share (%)	Number of counties
Aetna	4.14	99.82	11.54	171.7
BCBS	23.23	99.99	30.06	962.9
Cigna	0.041	99.26	8.777	296.8
Humana	10.78	81.39	45.11	1,989
Kaiser	20.24	99.98	30.69	97.61
United	5.00	59.47	28.13	1,024
Other insurers	21.22	100	44.08	1,770
Total	53.67	90.20	100	2,412

repeated using ad occurrences as an alternative measure to see whether the results are consistent.[8] The data also include the total estimated expenditure of the firm on the advertisement, the duration of the advertisement, and very coarse age, race, and gender demographic breakdowns of the impressions data. The data include the parent company of the product advertised, a description of the product being advertised, and a very brief description of the content of the advertising copy.

The largest six firms account for 62% of the total advertising impressions. While UnitedHealthcare and Humana have some national ads, overall 90% of advertising for MA is local. Descriptive statistics about firm-level advertising and shares are presented in Table 1. There is also considerable variation in ad spending within a year, with ads heavily focused during the open enrollment period, which runs from October 15 through December 7 each year. Figure 2 shows this dynamic. In terms of advertising copy, these ads are primarily brand-centric advertisements that are not specific about details of the products. In particular, Cai et al. (2008, p. 1) notes that "[t]he majority of Medicare plan ad occurrences did not convey basic, descriptive information defined in this study to include both plan type and premium amount."

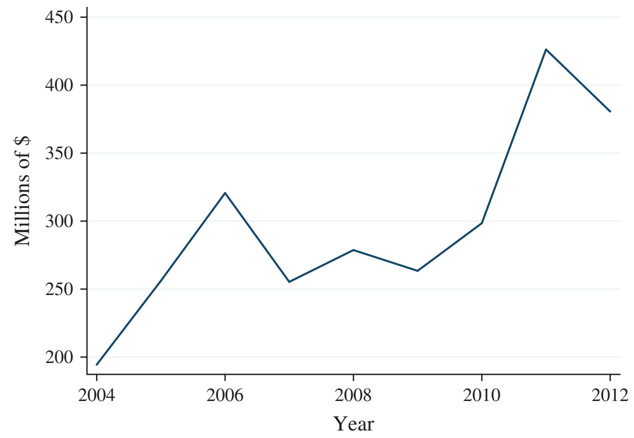
Pairing these data with population data from the U.S. Census, the total number of gross rating points (GRPs) that each advertisement constituted is computed. A GRP is the typical unit of sale between a firm and a television network for advertising space: the total number of advertising impressions divided by the population in the DMA. As such, a yearly increase of one GRP can be interpreted as the average person viewing the ad one additional time over the course of that year. Figure 1 shows the evolution of health insurance advertising spending over the course of the sample. In 2004, health insurance advertising made up about $200 million. By 2012, that number roughly doubled.

3.2. Enrollment

MA enrollments and plan characteristics are measured using data from CMS. Enrollments at the plan-county-month level from 2007 to 2012 are observed. However, with few exceptions, seniors choose their MA plans once per year during open enrollment. The enrollments decided upon by seniors in open enrollment translate into enrollments effective January of the following year. As such, enrollments are measured at the yearly level in February of each year, and those enrollments will be paired with advertising from the prior year when seniors make their choices about the upcoming year's insurance. Plan characteristics such as premiums at the plan-county-year level are also observed.

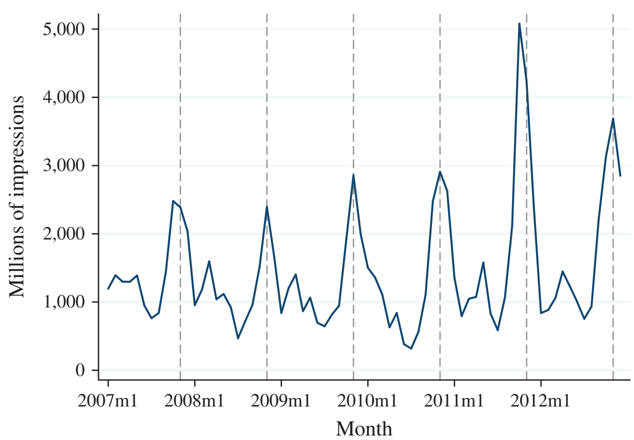
Finally, information from the Census on demographics such as population, Medicare-eligible population, and race are merged at the county level. The Medicare-eligible population is computed using the

Figure 1. (Color online) Health Insurance Advertising by Year



Notes. The figure shows spending on health insurance television advertising in millions of dollars over time. Total television advertising is just under $100 billion per year.
Source. AC Nielsen's media database.

Figure 2. (Color online) Health Insurance by Calendar Month



Notes. The figure shows that a significant amount of advertising is concentrated in the open enrollment period from October 15 to December 7. Dashed lines mark the month of November each year m1, month 1 (January).
Source. AC Nielsen's media database.

share of the population over 65 years old. The United States is generally an aging population: in 2006, the average county had 15.04% of its population over the age of 65, and in 2012, the average county had 17.16% over 65. Data on the Medicare risk scores of each county and statutory benchmark rates, which help to determine the capitation payment rates to MA plans, and plan-level characteristics from the CMS are also included. Risk scores are centered at 1 with a standard deviation of 0.1, with higher scores indicating worse health. Combining all of the data, this study uses enrollment and advertising at the county-plan-year-level from 2007 to 2012, using only data from after the introduction of Medicare's Part D prescription drug benefit.

3.3. Aggregation to Full Data

The final data are aggregated to the brand-county-year level for both advertising and enrollments. For enrollments, this is because the data do not indicate when exactly an enrollee signed up, only when she is currently enrolled. Since enrollments that happen on a particular day during open enrollment in most cases do not take effect until January of the following year, I cannot identify the exact point of enrollment. Since enrollment is aggregated to the brand-county-year in this way, more frequent than yearly observations of advertising are also not particularly helpful, as the time between the ad exposure and the sign-up itself remains unobservable. While a monthly or weekly decay of advertising stock might be a preferable approach for aggregating advertising in an ideal world, being able to identify the month or week of the sign-up is required to initialize the ad stock variable with respect to time. However, separate analysis of advertising only in the open enrollment period will be conducted since that is when both the majority of advertising happens and when the majority of enrollments take place.

4. Research Design

4.1. Endogeneity of Advertising

Identifying the effects of advertising can be difficult, both in terms of statistical power and bias induced by various forms of endogeneity. In terms of power, Lewis and Rao (2015) shows that because of small true advertising effects and often large amounts of noise in purchases, it can be very difficult to pin down advertising effects with precision. In terms of bias, advertising is a firm choice, subject to equilibrium forces and firm maximization. These forces may cause advertising to be correlated with sales for reasons other than a treatment effect of advertising. Firms might also use rules of thumb based on targeting past or expected sales rather than perceived treatment effects, leading to potential concerns about reverse causality. Indeed, most plausible confounds would bias the researcher in favor of finding a larger advertising effect where none (or a smaller one) exists. In the case of correlated firm behaviors, this is nicely illustrated by Lewis et al. (2011).

4.2. Identification Strategy

In this study, sharp discontinuities in the level of advertising at the borders of geographically based television markets provide exogenous variation. This design was first used by Shapiro (2018) to study the effects of television advertising on antidepressant demand, but is also used by Tuchman (2015) to study e-cigarette advertising, as well as Spenkuch and Toniatti (2018) to evaluate the effectiveness of political advertising. Consumers who live on opposite sides of DMA borders face opposite levels of advertising, because of market factors elsewhere in their DMA, but they have similar observable characteristics and choice sets of products. In this way, at the borders, observed advertising is "out of equilibrium" and simulates an experiment.

Capturing this intuition, I estimate the casual effect of advertising on MA enrollment, controlling for unobservable geographic characteristics with border-specific brand-time fixed effects. This allows unobservables to be spatially correlated in ways that are consistent with the take-up of MA across the country. To improve precision and to control for any unobservables that are persistent within counties over time, the panel nature of the data is leveraged using brand-county fixed effects. As regulatory regimes may differ across state lines, I focus on DMA borders that are within a state. The identifying assumption is that there are no unobserved differences in trends across these borders that are simultaneously correlated with changes in advertising and the MA share. At the brand level, I conduct these comparisons only where there exists a matched pair across the border; that is, if UnitedHealthcare is present in a county on one side of a DMA border, it is included in the analysis only if it is present in at least one county on the opposite side of the DMA border.

The top 120 DMAs contain 210 such within-state borders, 164 of which where the border areas make up no more than 35% of the total DMA population. At the brand level, this creates 772 brand-level border experiments where cross-border matches can be made, 573 of which are in border areas that make up no more than 35% of the DMA population. In the main analysis, attention will be restricted to these borders, but sensitivity analysis around the 35% cutoff will be conducted in the appendix. Each of these brand–border pairs will be considered a separate experiment, with the magnitude of the treatment determined by the advertising in each DMA at a given time, measured in GRPs. Only the counties bordering each other while being in the same state will serve as controls for each other, to partial out any local effects that may be increasing or decreasing MA enrollments for both sides of the border, including any national advertising. The level of an observation

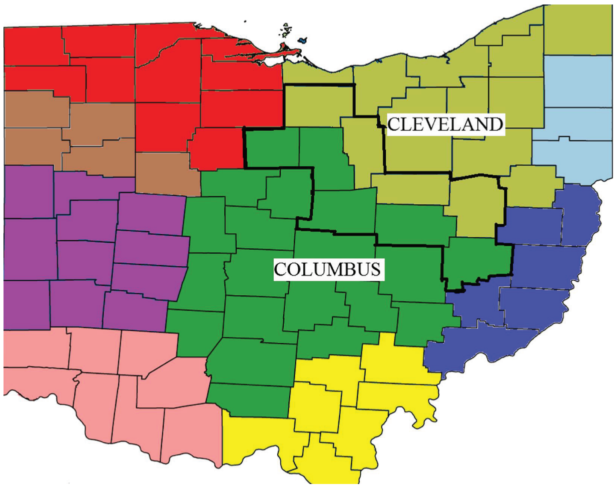


Figure 3. (Color online) Ohio and Its DMAs

Note. Figure shows the Cleveland and Columbus DMAs, highlighting the border region used in identification.

is a county-year in the category-level analysis and a brand-county-year in the brand-level analysis. In each "experiment," one such set of (brand-)counties will be compared with an adjacent set of (brand-)counties directly across the DMA border.

For a leading example, Figure 3 shows the Cleveland and Columbus DMAs in the state of Ohio. The border experiment considered is outlined in bold. I compare how outcomes on the Cleveland side of the border change when the Cleveland DMA receives a change in advertising GRPs relative to the Columbus DMA. Figure 4 shows this dynamic graphically. In panel (A), the time series of advertising for Humana for both the Cleveland and Columbus DMAs are pictured, and in panel (B), the time series of Humana demand is shown for a county on either side of the border. As the Columbus side gets a larger bump in GRPs than the Cleveland side of the border, the county on the Columbus side of the DMA border sees an increase in demand. This is the type of variation that the border approach exploits.

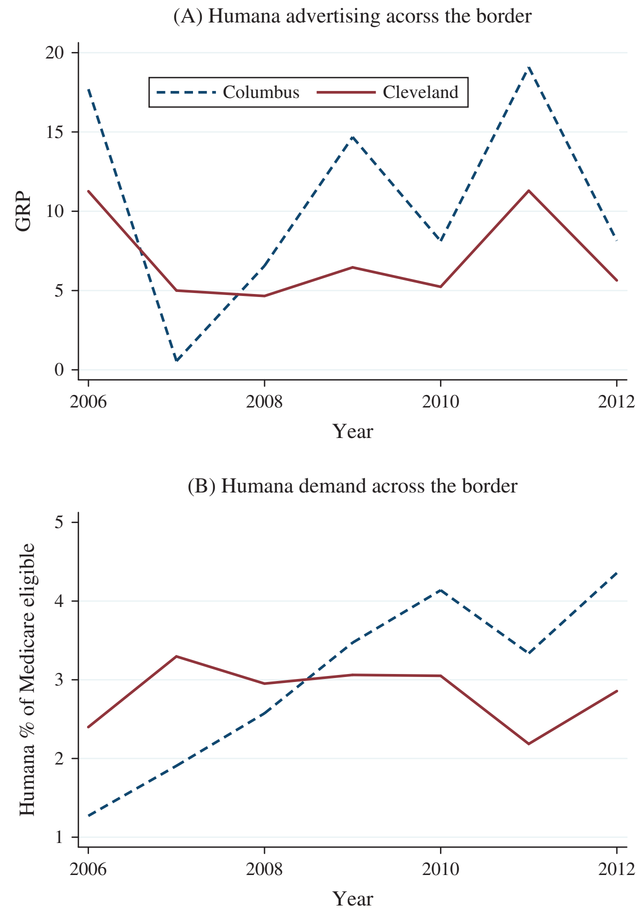
4.3. Econometric Model

To model the main effects of advertising on brand demand, let i index counties, b index borders, j index brand, and t index time. Let s_{bijt} indicate the percentage of Medicare beneficiaries with MA coverage through brand j, and let GRP indicate the level of advertising. The effect of an increase in advertising GRPs on MA brand j demand is estimated with regressions of the form

$$s_{bijt} = \gamma_1 GRP^{\text{own}}_{jit} + \gamma_2 GRP^{\text{rival}}_{jit} + \alpha_{bjt} + \alpha_{ij} + X_{bijt}\alpha_X + \epsilon_{bijt}, \quad (1)$$

where α_{bjt} is border-brand-time fixed effects, α_{ij} is brand-county fixed effects, and X_{bijt} is a vector of brand and county control variables, including demographic, competitive environment, and plan characteristics. In this case, the coefficients of interest, γ_1 and γ_2, capture the causal effects of an increase in own and rival advertisements on brand share, respectively. Since variation is at the brand-DMA level and includes repeated measurements over time for each brand-county, standard errors are clustered by brand-DMA in all brand-level analyses. This base model is easily augmented to study category-level effects, to include interactions with advertising effects, or to use a different measure of either advertising exposure or demand.

Figure 4. (Color online) Border Experiment Illustration

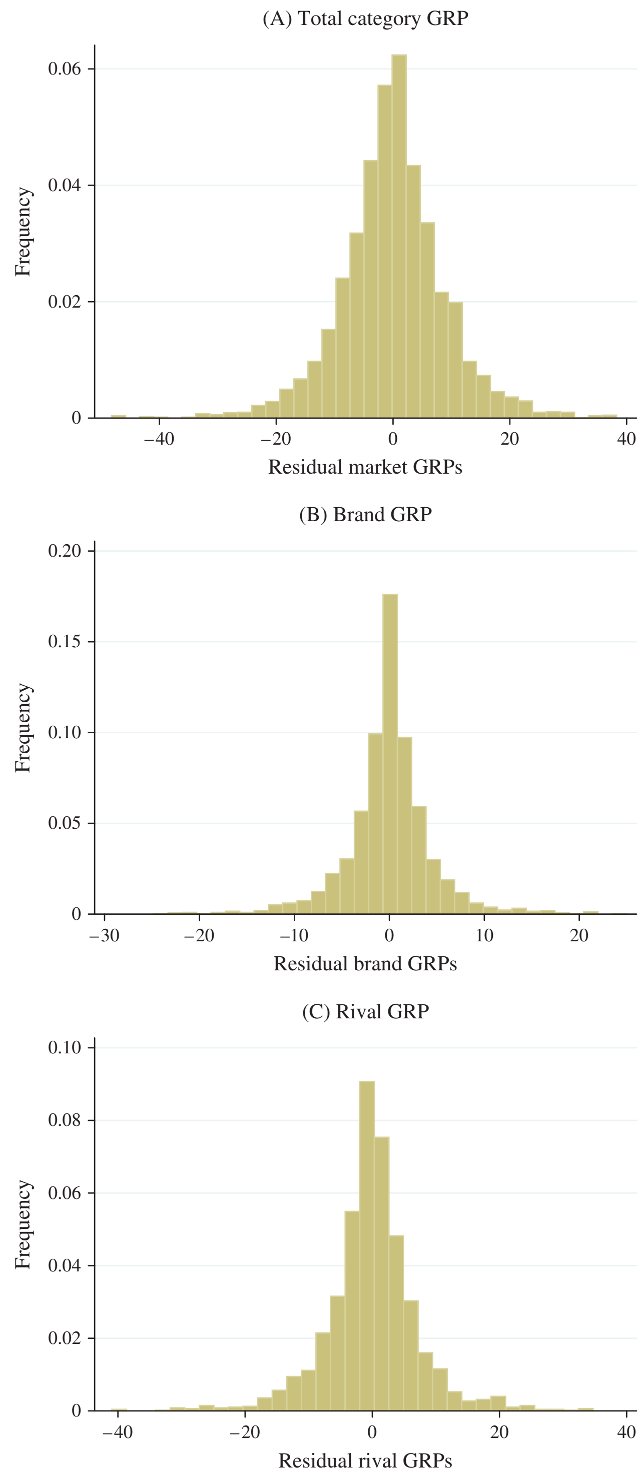


In the brand-level model, there are 573 brand-level border experiments, though not all of them are present in each of the six years in the data. This is because a plan might exit the market during the data or not start in a market and enter later in the sample. The main model contains 2,298 brand-border-year dummies and 2,008 brand-county dummies. While this is quite a lot of fixed effects, there remains sufficient variation to identify an average effect of advertising, as will be shown in Section 5.1. In the category-level model, there are 164 border experiments over the six years, making 984 brand-year dummies and 742 county fixed effects. There remains sufficient variation to identify an average effect of advertising precisely. With this many fixed effects, however, there is insufficient power to estimate very disaggregate heterogeneous treatment effects, such as border-experiment-specific effects.

For this approach to be useful in identifying advertising effects, two conditions must hold. First, there must be sufficient variation in advertising across the borders in the data. If all advertising variation were at the national level over time and local stations rarely used their discretion to displace national ads, the border-specific time fixed effects would sweep away all variation in advertising. Figure 5 shows that there is significant advertising across the borders in both total-market and brand-level GRPs, as well as rival GRPs, by plotting histograms of total, brand, and rival GRPs net of the fixed effects and control variables included in the border approach.[9] In this way, the variation in the histogram reflects the identifying variation in the border approach.

Second, the placement of the borders must be quasi-random with respect to preferences for health insurance, conditional on covariates. Since counties are the relevant markets for MA plans and DMA boundaries coincide with county boundaries, it is also important that plan attributes be either controlled for or quasi-random at the DMA borders in particular; that is, any unobserved plan attribute must follow parallel trends at the border. In the analysis, I will control for plan premiums, drug benefit designs, and entries and exits, as well as demographic variables. As policies related to health insurance and healthcare often vary at the state level, DMA borders that coincide with state borders are excluded, as many policies that may affect preferences change at state borders. The locations of DMA borders were determined historically by AC Nielsen and have rarely changed over time. To provide evidence that observables are not predicted by advertising across the borders, Table 2 shows the estimated coefficient from a regression with observable county and plan characteristics as the dependent variable and DMA-level MA advertising as the independent variable, controlling for brand-border-year fixed effects. As such, this regression tests whether there are discontinuous changes in observables as the DMA border is crossed. None of the county demographics or plan-level variables, including population, share of population over 65, average income, race, average premium, or drug benefit design are predicted by advertising at the border. The only observable that is predicted by advertising differences across the border is in the number of brands competing in the market. An additional GRP is associated with 0.0023 more competitors across the border. As the average number of competitors is 2.52, this is a 0.09% difference. In terms of unobservables, the DMA boundaries were set by AC Nielsen long ago based on which local news station a family was more likely to want to watch. This was largely determined by which television stations were reachable over the air by households at that time. As very few households watch TV over the air today, these boundaries are as good as random.

Figure 5. (Color online) Variation in GRP Changes Across DMA Borders



Note. The figure shows that there is significant variation in total, own, and rival advertising across DMA borders, conditional on (brand-) county and (brand-)border-year fixed effects and control variables.

Table 2. "Balance" Test

	Est.	Standard error	*P*-value	Mean
AvgPremium	−0.0134	0.0218	0.5404	49.25
BasicDrugInc	0.0002	0.0001	0.1334	0.1107
EnhancedDrugInc	−0.0003	0.0002	0.0710	0.7175
DrugDeductible	0.0128	0.0251	0.610	21.75
County benchmark	0.0587	0.0771	0.44779	757
Risk (HCC score)	−0.0026	0.0056	0.6404	94.81
%White	−0.0076	0.0088	0.3868	86.03
%Black	0.0076	0.0088	0.3680	9.738
%Hispanic	−0.0116	0.0071	0.1067	4.914
%Asian	−0.0007	0.0007	0.3584	0.912
AvgIncome	−1.5021	4.3982	0.7334	31,800
%Elderly	0.0026	0.0037	0.4827	16.75
Population	7.5015	56.3943	0.8944	67,200
Brands present	0.0023	0.0007	0.0024	2.52

Notes. Regressions reflect regression with the listed variable as the dependent variable and total GRP as the independent variable and the border-year fixed effects from the border strategy. This tests whether there is a discontinuous change in the observable across the border that is predicted by the advertising level.

The maintained identifying assumption throughout is that trends in any unobservables that correlate with demand must be parallel across the DMA borders.

In terms of confounds of particular concern, I directly control for payments from the government using the county benchmark rates, for entries and exits using the number of brands present in each county-year, and for drug benefit design (enhanced drug benefits included, basic drug benefits included, and drug deductible if applicable). Since physician networks for each plan are not directly observable, brand-county fixed effects control for the persistent element of a brand's network in a given county. The brand-border-year fixed effects control for changes in network design over time that are common to the entire border region. This is a similar approach to that taken by Curto et al. (2015) to control for plan networks. If there are residual changes in networks that are correlated with changes in advertising, those will bias advertising effects up (down) if the favorability of networks is positively (negatively) correlated with changes in advertising. As advertising is typically thought of as positively confounded with other factors that influence demand (Lewis et al. 2011), it might be more likely for any such bias to be in the upward direction.

Aside from premiums and drug benefit design, it is possible that some contamination could still occur if there are other firm strategies that are mechanically correlated with television advertising. In particular, direct mail advertising and online advertising are unobserved. The maintained assumption is that firms do not alter their direct mail or online ad strategies discontinuously at DMA borders as television advertising is altered. Doing so would likely be prohibitively costly from a technical standpoint given the potential return. Additionally, using a different identification strategy, Aizawa and Kim (2018) show that direct mail advertising has no significant effect on MA shares, and any omitted variable bias requires that the omitted variable be driving the outcome as well as being correlated with the endogenous variable. Finally, as pointed out by Lewis et al. (2011), firms tend to correlate their various advertising strategies positively, so if firms were engaged in such highly detailed targeting of other strategies such that they changed discontinuously at borders, these would tend to bias advertising effects upward.

4.4. Features and Limitations

Perhaps the largest feature of this approach is that the observed advertising levels at the border are out of equilibrium; that is, variation is driven by the equilibrium in other markets. At the border of the Cleveland DMA, viewers see Humana ads that are driven by metro Cleveland, despite that at the border, these residents can be quite different. If ads were microtargeted to the county level, these consumers would likely see different ads. Similarly, on the Columbus side of the DMA border, the advertising is largely driven by metro Columbus, which is away from the border, again giving rise to rather different advertising at the Columbus border than if ads could be microtargeted. If metro Columbus and metro Cleveland are sufficiently different from each other, these very similar consumers right on the border will get very different ads, even though their equilibrium microtargeted ads would have been very similar. This gives a reasonable amount of variation away from what would be the equilibrium in the microtargeted world while using the fact that these consumers across the border from one another are very similar to control for unobservable factors driving demand.

By conducting the analysis in this way, it is possible to see advertising levels that are likely to be both well above and well below what would be optimal if firms microtargeted each county individually, which makes the estimated treatment effect approximate an average treatment effect across the advertising response curve for this population. In an experimental approach where the researcher injects some noise into a preexisting equilibrium or targeting rule, the estimated effect will be local only to levels of advertising that are near that equilibrium or targeting rule. If the firm is already optimally allocating advertising spending, the incremental effect from a small amount of noise being added to the equilibrium might be hard to pin down and smaller than the average effect. In an IV approach, the estimated effect will be local only to those people who are affected by the instrument (i.e., the "compliers")—a group that is not always straightforward to characterize for policy or managerial purposes. Additionally, because the border approach does not require the use

Table 3. Selection into the Border Sample

	Difference	Mean	Difference (%)	*P*-value
AvgPremium	0.9007	48.61	1.853	0.4497
BasicDrugInc	0.0092	0.1043	8.82	0.0648
EnhancedDrugInc	−0.0327	0.740	4.4	<0.01
DrugDeductible	3.802	19.17	19.83	
County benchmark	−21.43	772	−2.776	<0.01
Risk (HCC score)	0.1127	95.04	0.119	0.7357
%White	0.4374	85.45	0.512	0.5228
%Black	0.382	9.6947	3.940	0.5616
%Hispanic	−3.1858	7.1323	−44.667	<0.01
%Asian	−0.5278	1.3032	−40.500	<0.01
AvgIncome	−2,760	33,500	−8.239	<0.01
%Elderly	0.7774	16.0474	4.844	<0.01
Population	−75,200	122,000	−61.639	<0.01
Brands present	−0.0857	2.6909	−3.185	0.016

Notes. "Difference" reflects the estimate from a regression with the listed variable as the dependent variable and an indicator for "in the border sample" as the independent variable. This shows how the observable demographics are systematically different for the estimation sample.

of instruments, it is not subject to potential weak instrument bias as well as some of the less desirable finite sample properties of IV estimators.

The border approach falls victim to the familiar local average treatment effect issues that are also common to experiments and IVs. In this case, the estimated effect will be local to those consumers who live in border areas; that is, the compliers will be the set of people that live within the border sample, which is a group that can be characterized and compared with the population at large in a straightforward way to assess whether sample selection is a problem. Table 3 shows how consumers in the border sample are systematically different from consumers outside of the border sample. Note that the average population in a border county is considerably smaller than that in a county outside of the border sample. The border sample also has lower Asian and Hispanic population percentages, slightly less competition, and a larger percentage elderly population. If anything, intuition suggests the larger percentage of elderly consumers in the population (though a small difference) would lead to a higher estimated advertising effectiveness for a product that only elderly consumers purchase. While there are these systematic differences, there is considerable overlap in the support of the distributions of these characteristics between the border sample counties and other counties. As such, the extent to which these characteristics are important to advertising effects may be estimated directly by interacting them with advertising. Additionally, specifications will be run that are identical using the full sample and the border sample to see the likely effect of sample selection in this context.

An additional potential limitation to this approach is that it relies crucially on variation in local advertising, which is often a remnant of the up-front market and might be systematically different from national network or cable advertising. However, in this particular context, that limitation is minimal, as more than 90% of the total advertising is local, as shown in Table 1. MA advertising focuses so much on local ads because of the local nature of health insurance markets. While network TV eyeballs may well be different from spot TV eyeballs, the spot TV eyeballs are the most relevant for this particular market.

A final limitation of this approach is that if there are word-of-mouth effects that are more likely for people in close proximity to one another, the total advertising effects might be underestimated; that is, if one senior sees an ad and then goes to a senior center for a card game and tells her friend about the ad, both seniors would be "exposed" to that ad, even if the friend lived on the opposite side of the DMA border and did not watch the ad on TV. In that case, the estimated effect would be the effect of actually watching the ad on TV over and above the conversation between the friends. Given that the average person communicates on the phone and Internet with people who live far away, it is not clear that word-of-mouth effects are necessarily larger for people in closer proximity. Given the research design, it is difficult to say much about potential word-of-mouth effects.

4.5. Computing Cost of Conversion

Using the above approach produces the average treatment effect of advertising GRPs on demand. A more managerially relevant metric that also helps to frame the economic significance of the estimates is the CPC, which is the acquisition cost of a customer using advertising. To convert the estimates from average lift to CPC, I use the cost of observed advertising from the data, computed at the average cost of an MA GRP in the sample. This is inflated to account for the fact that most brands do not operate in every county in a DMA and, as such, "waste" some percentage of eyeballs. Since the measured treatment effect of advertising is the effect of GRPs on MA share, the CPC is the cost per GRP divided by the product of the estimated advertising effect and the Medicare-eligible population.

Note that the computed CPC from the estimated advertising effects represents a lower bound on the true average CPC for at least two reasons. First, the CPC is a convex function of the effect size. Since CPC is computed at the average ad effect, Jensen's inequality says that the true average CPC is greater than the CPC computed at the average value of effect size. Next, from talking with managers in the industry, nearly all sign-ups for MA plans are consummated by brokers, who extract fees as high as $450 per enrollment with continuing fees if the enrollee renews in subsequent years. Even if the advertising caused the conversion, it is likely that some of the converted enrollees used brokers that were paid commissions, increasing the cost of the average conversion.

CPC estimates are provided for both ends of the 95% confidence interval of average lift. Where the confidence interval in average lift is less than zero, the upper end of the CPC confidence interval is infinity. External information on the cost of acquiring enrollees through price reductions as well as information on average static profitability provides context on whether or not a particular CPC is relatively high or low.[10]

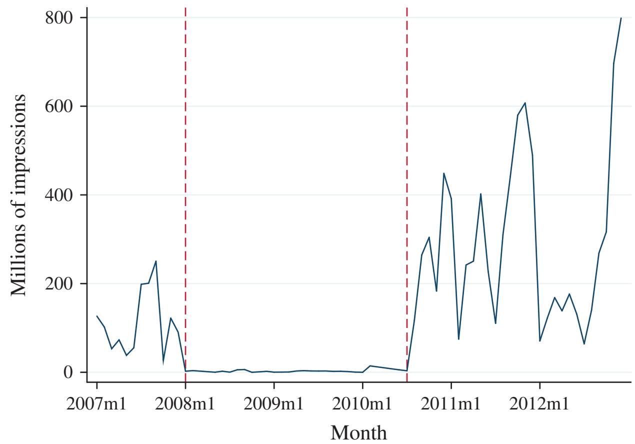
4.6. Mechanisms

Four hypotheses about the mechanisms of advertising effectiveness are tested. First, advertising could primarily drive category enrollments. Next, advertising could work differentially on different populations. Advertising might also be a prisoner's dilemma, whereby firms advertise only to "cancel out" the advertising of their competitors. Finally, advertising might have longer-run effects due to the sticky nature of insurance enrollments as well as potential brand-building effects.

Category demand and heterogeneity are straightforward extensions of the original model. To study category-level demand, the same model is applied, augmented slightly to include only category demand and only total category GRPs. To study heterogeneity, GRPs are interacted with observable characteristics of interest. All of these variables are normalized to have a mean of zero and standard deviation of one for ease of interpretation on the main effect of advertising and are assessed at both the category and brand levels. In particular, regulators are interested in the category-level interaction between GRPs and health status. If advertising disproportionately works on healthy consumers, it might induce adverse selection for TM plans, increasing the costs to the government. For the purposes of targeting and potentially increasing profit, firms should be interested in any heterogeneity in the treatment effect that they may legally be able to target. As scientists, heterogeneous treatment effects may help us to understand something about how exactly advertising works.

While the effects of rival advertising on brand demand are shown in the main effects, a more direct test of the prisoner's dilemma hypothesis is presented, leveraging the fact that UnitedHealthcare unilaterally stopped advertising during 2008 and 2009. In 2006, a *Wall Street Journal* article detailed a scandal involving backdated stock options (Forelle and Bandler 2006). This led to a Securities and Exchange Commission investigation over the course of many years and the eventual resignation of chief executive officer William W. McGuire. In the midst of the commotion and regulatory attention, UnitedHealthcare spent almost nothing on health insurance advertising between 2008 and 2009, as can be seen in Figure 6. If the prisoner's dilemma hypothesis holds, then rival advertising should deteriorate United's brand share over the years in which it does not advertise. Similarly, if advertising is needed in equilibrium to enhance other (unobserved) marketing levers (e.g., the effectiveness of brokers), United's advertising cessation should deteriorate United's share over this period. To control for direct effects of the scandal common to similar counties, the cessation is interacted with the border strategy to provide the causal effect of rival advertising on United brand share.

Figure 6. (Color online) UnitedHealthcare Advertising Expenditure Over Time



To address the long-run effects of advertising, a goodwill stock conception of advertising is used instead of yearly GRP levels. Given the sticky nature of health insurance purchase, advertising might be expected to have longer-run effects, as enrollees who sign up in one year due to advertising are likely to stick around for subsequent years, making advertising more valuable than it would appear in a static sense. In fact, Curto et al. (2015) find that roughly 77% of MA enrollees stick with the same plan each year. Advertising could also have long-run effects on brand equity that would manifest themselves only over a number of years. These two facts combined suggest that the effects of past advertising stock are important to assessing profitability. I measure the long-run effect of advertising using a stock conception assuming different rates of ad stock persistence.[11] I use data from 2006–2010 to initialize the ad stock and estimate the model on the last two years of data (2011–2012) to maximize the number of periods I can use to build the ad stock while still allowing for the use of fixed effects.[12]

If the advertising marginal enrollee were the same as the average enrollee in terms of inertia, then the effect of this year's advertising on next year's demand should be 0.77 as big as the static effect of advertising (or a 0.77 advertising stock persistence parameter), as Curto et al. (2015) find the average enrollee stays with her current plan 77% of the time. If advertising marginal enrollees are less sticky than average, assuming an ad stock persistence parameter of 0.77 will overweight past advertising. Even though past advertising does not work that well, some of the effect of current advertising is attributed to past advertising erroneously. This will bias the estimated effect of advertising stock downward. As such, if the dynamic effect estimated using this method is smaller than the ad effect from the main (static) analysis, I will conclude that the advertising marginal enrollees are less likely to stay in their current plan in a subsequent year than the assumed persistence parameter.

5. Results

5.1. Main Effects

Table 4 presents the estimation of the effect of brand and rival GRPs on the brand demand using Equation (1). To provide economic intuition for the size of the advertising effects, CPC estimates and confidence intervals are presented below the advertising effects.

Table 4. Brand-Level Demand (MA %)

	(1)	(2)	(3)	(4)	(5)
GRP	0.0971***	0.0961***	0.0112	0.0213*	0.0092*
	(0.0228)	(0.0212)	(0.0073)	(0.0099)	(0.0046)
RivalGRP	−0.0163*	−0.0264***	−0.0015	−0.0015	0.0026
	(0.0075)	(0.0072)	(0.0030)	(0.0036)	(0.0028)
AvgPremium		−0.0082	−0.0175***	−0.0193***	−0.0193***
		(0.0047)	(0.0019)	(0.0024)	(0.0023)
BasicDrugInc		1.3331	−0.0293	0.1429	0.4959
		(0.6945)	(0.3493)	(0.4221)	(0.5439)
EnhancedDrugInc		3.7813***	0.5021	0.7904*	0.6827*
		(0.4252)	(0.3001)	(0.3395)	(0.3013)
DrugDeductible		0.0045	0.0005	0.0013	−0.0012
		(0.0025)	(0.0013)	(0.0015)	(0.0019)
BrandsPresent		0.3408*	−0.4129***	−0.4657***	−0.3618***
		(0.1354)	(0.0809)	(0.1117)	(0.0904)
Risk		−0.0021	−0.1126	−0.1288	−0.1871
		(0.2398)	(0.1568)	(0.1911)	(0.1266)
CountyBenchmark		0.5905**	−0.0644	−0.3018*	−0.1384
		(0.2255)	(0.1344)	(0.1213)	(0.1392)
Income		−0.3410**	−0.4107*	−0.6624*	0.0413
		(0.1231)	(0.1692)	(0.2827)	(0.2267)
%Elderly		−0.2369*	−0.4790***	−0.5600*	−0.6212***
		(0.1137)	(0.1414)	(0.2395)	(0.1506)
%White		0.1665	0.0737	0.3806	0.1219
		(0.2431)	(0.2114)	(0.4321)	(0.3864)
%Asian		−0.0606	−1.2631***	−1.3033*	−0.5939
		(0.0895)	(0.3702)	(0.5096)	(0.4248)
%Black		−0.3099	−1.6695*	−2.757	−1.2531
		(0.2381)	(0.8292)	(1.4720)	(1.0614)
%Hispanic		−0.3591**	−1.1465*	−1.5322	−1.4192
		(0.1169)	(0.4895)	(0.9006)	(0.7224)
Brand-year FEs		×	×	×	
Brand-county FEs			×	×	×
Border sample				×	×
Brand-border-year FEs					×
CPC	$123.72	$125.02	$1,076.71	$562.71	$1,298.74
CPC CI	[84.74, 229.10]	[87.27, 220.30]	[472.71, ∞]	[295.15, 6,019.16]	[659.26, 43,300]
Mean MA %	5.4456	5.4364	5.4958	5.0191	5.0191
R-squared	0.07	0.175	0.895	0.854	0.917
Observations	36,764	36,489	35,919	10,651	10,651

Notes. Brand-DMA clustered standard errors are in parentheses. The following variables are averages across plans within a brand-county-year: *AvgPremium*, *BasicDrugInc*, *EnhancedDrugInc*, and *DrugDeductible*. The following variables are normalized to have mean zero and standard deviation one and are common across brands in a county-year: *BrandsPresent*, *Risk*, *CountyBenchmark*, *Income*, *%Elderly*, *%White*, *%Asian*, *%Black*, and *%Hispanic*. FE, Fixed effect.

$^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$.

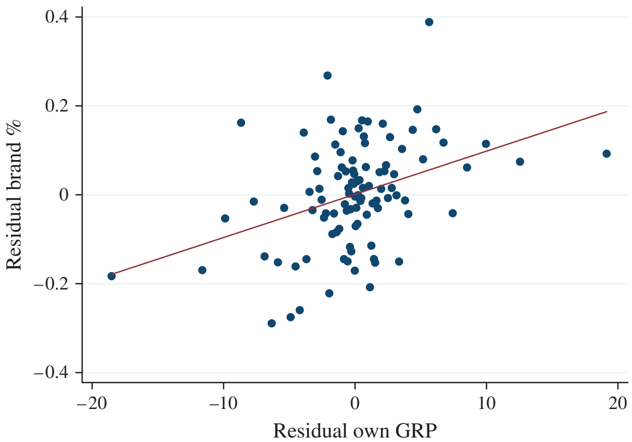
Column (1) presents the naïve regression, where all counties are included, and no fixed effects or controls are used. It suggests a positive and significant effect of advertising on brand demand, with an increase of one GRP associated with an increase in brand demand 0.0971 percentage points, implying an average CPC of $123.72. Effects of this size would make advertising a relatively inexpensive way to acquire customers, if true. Additionally, column (1) shows a small but significant negative impact of rival advertising on demand. Column (2) adds observable control variables for demographics, premiums, whether and to what extent drug benefits are included, and year fixed effects, as well as the number of competitors in the market, average risk scores of the counties, and statutory county benchmark rates. This will control for the possibility that advertising affects preferences but premiums or drug benefits change as a result, masking or inflating the effect of advertising. The effect size is almost unchanged and implies a CPC of $125.02, while the magnitude of the rival effect increases in magnitude to −0.026, suggesting significant damage to the firm from rival advertising. If firms systematically target advertising to markets that are strong for reasons other than advertising, the first two columns would provide spurious positive estimates of the own ad effect as well as spuriously negative effects from rival ads. In column (3), brand-county fixed effects are added to control for time-invariant factors affecting MA strength in a particular county. In this case, identifying variation comes from deviations from the average brand advertising in a market. Controlling for these factors makes the ad effect shrink considerably. The average lift from one GRP is about 0.011 percentage points and implies a CPC of $1,076.71, and the effect of rival ads goes to zero. Meanwhile, the estimated effect is not especially precise, as the 95% confidence interval of the CPC in this specification is [$472.71, ∞]. While $472.71 might be worth spending to acquire a customer, zero average effect (and infinite CPC) of advertising cannot be ruled out. Column (4) runs exactly the same specification as column (3) but limits the sample to that which will be included in the border approach, to highlight the role of sample selection in this context. In this specification, counties in the border sample are included, but a general brand-year fixed effect is included as in column (3) rather than controlling for local demand shocks using the brand-border-year fixed effects. The estimated effect of advertising goes up, suggesting that the sample selection effect to the border would tend to overstate the true average effect in the population.[13] However, if firms target markets of not only historic strength but also recent strength, or target unobserved concurrent local demand shocks, columns (3) and (4) will also overestimate the advertising effect. In column (5), the border approach, using brand-county and brand-border-year fixed effects, is employed to control for all remaining endogeneity issues. The effect of advertising on brand demand is now more precisely estimated and is statistically significant, though small in magnitude. The point estimate implies that an increase of one GRP is associated with an increase in brand demand of 0.0092 percentage points. Evaluated at the point estimate for an average number of GRPs per brand-year (9.3) and average share per brand-year (5.019%), this implies that, on average, advertising accounts for about 1.7% of a firm's customers in a given year. The confidence interval around the ad effect is roughly two-thirds as wide as the fixed effects estimation in column (3). However, despite finding a statistically significant ad effect, the implied CPC is difficult to bound above at a reasonable rate, as the CPC approaches infinity as the ad effect approaches zero. The 95% CPC confidence interval ranges from $659.26 to $43,300, with an average CPC of $1,299. The estimated rival effect is positive, suggesting a positive spillover, but is neither statistically nor economically significant. While a simple correlation would lead the researcher to conclude that advertising is a highly profitable way to shift the elderly into a branded MA plan, a more careful analysis using plausibly exogenous variation shows much more limited advertising effectiveness. Additionally, while the simple correlation would lead the researcher to conclude significant harm from rival advertising, the more careful approach shows no material effect.

In terms of control variables, I restrict attention to the preferred specification in column (5). Unsurprisingly, premiums are negatively associated with brand demand, and generous drug coverage is positively associated with brand demand. The number of brands present is negatively associated with brand demand, suggesting that additional brands do steal from each other and not just from TM.

To provide a graphical illustration of these results, a bin scatter, using 100 bins, of brand MA percentage and brand GRP is presented in Figure 7, and one of brand MA percentage and rival GRP in presented in Figure 8. These amounts are residualized by a brand-county fixed effect and a brand-border-year fixed effect, so they reflect variation across TV market borders, as in column (5).

5.1.1. Discussion. The estimated effect of advertising, while detectable in this setting, is small relative to advertising effects estimated in past literature. The point estimate implies an advertising elasticity of about 0.027. The estimates in Aizawa and Kim (2018) suggest that a 1% increase in advertising leads to a 0.04% increase in enrollments, about 50% larger than what is found here. The estimates in this study are also smaller than the estimates of advertising effectiveness in other contexts. For example, Shapiro (2018) finds advertising elasticities of about 0.04 in the context of antidepressant

Figure 7. (Color online) Brand Bin Scatter—Own GRP

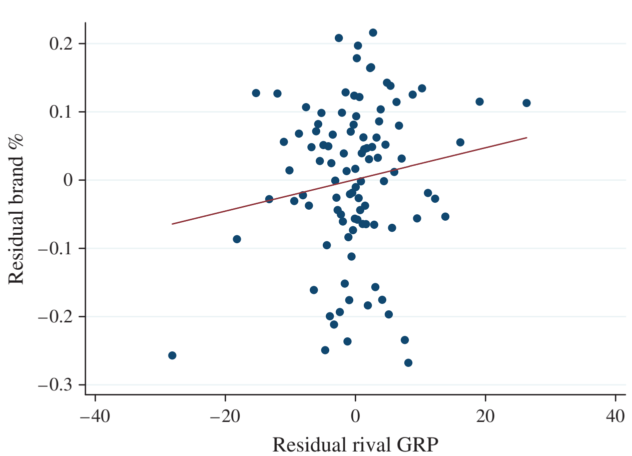


Note. This bin scatter has 100 bins and reflects brand percentage of the Medicare eligible population on the vertical axis and brand GRP on the horizontal axis, net of brand-county and brand-border-year fixed effects.

demand, and Sethuraman et al. (2011) find an average ad elasticity of about 0.12 in a meta-analysis.

To provide further context on the magnitude CPC in terms of profitability, an estimate of how much an incremental customer would be worth to the firm is useful. An incremental customer provides static profits and potentially continuing lifetime value, as customers often stick to plans for many years after the initial sign-up. From the GAO report (Cosgrove et al. 2013), the average static profit for an MA enrollee in 2011 was about \$504. The CPC is bounded below by \$659.26 in the preferred specification, which rules out short-run return on investment using advertising. However, firms might value customers above \$504 if they also provide future profits.[14] In Section 5.2.4,

Figure 8. (Color online) Brand Bin Scatter—Rival GRP



Note. This bin scatter has 100 bins and reflects brand percentage of the Medicare eligible population on the vertical axis and rival GRP on the horizontal axis, net of brand-county and brand-border-year fixed effects.

I directly assess whether the advertising marginal customers provide considerable future profits, either through advertising carryover or through inertia.

Firms may also reveal how much they are willing to pay for customers through other means of conversion. One way to convert additional customers is to lower premiums or add additional benefits. The cost of gaining a marginal customer using prices or benefits is that all inframarginal customers will also receive the premium reduction. The best evidence on price elasticity in MA plans comes from Curto et al. (2015), who document that MA customers have an average premium semielasticity of 0.012. This implies that a reduction in monthly premium of \$1 increases brand enrollments by 1.2%, or an average acquisition cost using price reductions of about \$1,000 on the margin, which is less expensive than advertising by about 20%, on average, though \$1,000 is still within the 95% CPC confidence interval. If firms are carefully setting their prices, they reveal a willingness to pay of about \$1,000 for a marginal enrollee, which should take into account possible inertia among the price marginal.

Overall, any potential effect of advertising on demand implied by these estimates is small relative to previous literature on advertising effects, and even though it is statistically significant, prohibitively large CPC cannot be ruled out. Even at the lower end of the CPC confidence interval, advertising is not profitable in the short run. As explained in Section 4.5, the average CPC computed here is a lower bound. All of this combined with the small estimated static effects implies that firms need advertising marginal customers to stay for multiple years at a minimum for advertising to be profitable.

5.2. Mechanisms

5.2.1. Category Expansion. From a regulatory perspective, the concern about cream skimming comes at the category level. If MA plans can draw the healthy out of TM, it leaves the government with unhealthy consumers. Since the capitation rates received by MA plans depend on the average cost of providing healthcare to a TM consumer, such cream skimming would increase the per-customer medical cost for TM and increase the amount the government pays to MA plans, even holding total expenditures on healthcare provision in the county fixed. Such an effect can only exist insofar as advertising moves seniors out of TM and into MA and the magnitude of the selection effect is limited to the size of the category expansion effect.

Table 5 shows the effects of category total advertising GRPs on the percentage of Medicare-eligible seniors who choose any MA plan over TM. The columns correspond with the specifications in Table 4, and the CPC estimates are interpreted as the cost of moving a senior out of TM and into any MA plan. This is lower than the

Table 5. Category-Level Demand (MA %)

	(1)	(2)	(3)	(4)	(5)
TotalGRP	0.0750***	0.0250	0.0023	0.0070	0.0044
	(0.0195)	(0.0145)	(0.0062)	(0.0073)	(0.0027)
AvgCatPremium		−0.0182	−0.0320***	−0.0348***	−0.0238***
		(0.0126)	(0.0052)	(0.0060)	(0.0034)
BasicDrugInc		−3.9113*	0.0158	−0.6519	−0.2704
		(1.9429)	(0.6522)	(0.7274)	(0.6107)
EnhancedDrugInc		0.6827	0.5217	−0.2037	−0.2903
		(1.0078)	(0.4856)	(0.5011)	(0.4876)
DrugDeductible		0.0027	0.0003	0.0037	−0.0025
		(0.0083)	(0.0018)	(0.0021)	(0.0025)
BrandsPresent		4.9743***	0.9361***	0.8580***	0.8362***
		(0.4451)	(0.1332)	(0.1764)	(0.1200)
Risk		0.0096	0.3947	0.1823	0.0273
		(0.5928)	(0.2728)	(0.3408)	(0.2636)
CountyBenchmark		1.7337**	−0.3927*	−0.9004***	−0.4643**
		(0.6407)	(0.1754)	(0.2034)	(0.1516)
Income		−0.7696*	−1.3187***	−1.4594***	0.2439
		(0.3484)	(0.2366)	(0.3368)	(0.3926)
%Elderly		−0.6389*	−1.1341**	−1.0419*	−1.7231***
		(0.2939)	(0.3369)	(0.5139)	(0.3128)
%White		0.4449	−0.0981	0.3706	−0.141
		(0.6077)	(0.3335)	(0.6327)	(0.6670)
%Asian		−0.0212	−0.9745	−2.0661*	−0.1961
		(0.2508)	(0.7714)	(1.0322)	(0.6304)
%Black		−0.8452	−1.2033	−2.5554	−1.3999
		(0.6269)	(1.2897)	(2.3934)	(1.9052)
%Hispanic		−0.6557	−2.1665*	−3.0306	−1.5592
		(0.3412)	(0.9493)	(1.5910)	(1.2766)
Year FEs		×	×	×	
County FEs			×	×	×
Border sample				×	×
Border-year FEs					×
CPC	\$160.20	\$479.93	\$5,293.70	\$1,721.25	\$2,726.76
CPC CI	[106.10, 326.85]	[224.69, ∞]	[835.02, ∞]	[565.89, ∞]	[1,229.33, ∞]
Mean MA %	12.5388	13.5378	13.5535	11.9524	11.9524
R-squared	0.061	0.309	0.963	0.945	0.978
Observations	17,136	15,535	15,511	4,951	4,951

Notes. Brand-DMA clustered standard errors are in parentheses. The following variables are averages across MA plans within a county-year: *AvgPremium*, *BasicDrugInc*, *EnhancedDrugInc*, and *DrugDeductible*. The following variables are normalized to have mean zero and standard deviation one: *BrandsPresent*, *Risk*, *CountyBenchmark*, *Income*, *%Elderly*, *%White*, *%Asian*, *%Black*, and *%Hispanic*. FE, Fixed effect.
$^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$.

firm-relevant CPC given that a senior converted from TM to MA might choose a rival plan. In column (1), neither fixed effects nor controls are used in a completely naïve approach. This specification suggests that an increase of one GRP corresponds with an increase in category demand of 0.075 percentage points and is statistically significant. The estimate corresponds with a CPC of \$160.20. Column (2) adds in controls for average county-level premiums, demographics, competition, risk scores, and benchmark rates. The estimate shrinks by a factor of three, to 0.025, and is no longer statistically significant. It corresponds with a CPC of \$479.93, and infinite CPC is now part of the confidence interval. Column (3) adds in county fixed effects and the effect size decreases by an order of magnitude to 0.0023, implying a CPC of \$5,293.70. Column (4) uses the same specification as column (3), but only on the sample of counties at the borders of DMAs. The estimate increases slightly to 0.0070, as in the brand-level analysis, but remains statistically insignificant and corresponds with an average CPC of \$1,721.25. Column (5) is the preferred specification and employs the border approach with county fixed-effects and border-year fixed-effects. The estimate shrinks to 0.0044 with a corresponding average CPC of \$2,726.76.

In terms of control variables, I restrict attention to the preferred specification in column (5). The average premium across MA plans in the county is negatively associated with the share of seniors who sign up for MA over TM. If MA is relatively more expensive, more seniors choose TM. As there are more brands present, there are more options for seniors, and more do choose MA. This suggests that additional brands are not only stealing share from each other, but are also expanding the market.

Using advertising to convert TM customers into MA plans has limited potential to create regulatory harm. The point estimate in column (5) implies that the elimination of all MA advertising would move the average percentage of seniors who choose MA over TM from 11.95% to 11.72%. At the right edge of the confidence interval, the elimination of all advertising would move the average MA share from 11.95% to 11.42%. If regulatory bodies relied on the naïve approach in column (1), effects would appear more than an order of magnitude larger and indeed be cause for concern. However, the more careful approach suggests limited potential for category expansion to drive significant shifts in the distribution of seniors picking MA over TM. This implies that even if advertising effects are directionally consistent with cream skimming, other levers, such as capitation payments, may be a more efficient tool to reduce cream skimming than regulating advertising. In addition, with an average CPC of $2,726.76 and no guarantee that the customer will pick the advertised brand, using advertising to bring customers into the category is a managerially costly strategy.

5.2.2. Selection Through Ad Effect Heterogeneity. While the overall effect of advertising on category expansion is limited, it could be that the regulatory concern over using advertising to cream skim is still well founded directionally. Such cream skimming using advertising might be possible if less healthy consumers were unable to focus on advertising or if advertising copy were particularly attractive to healthy rather than unhealthy people. Meanwhile, heterogeneity in the treatment effect is very important to firms in addition to regulators. If advertising is more useful on more profitable patients than average, the estimated CPC could be more easily justifiable. Additionally, the heterogeneity should inform targeting decisions about which markets should receive more advertising attention from firms.

To assess heterogeneity in the treatment effect, advertising GRPs are interacted with variables about the average health risk in the county, the number of brands competing in a county, average premiums, drug benefits, and demographic variables. To assess ad timing, advertising GRPs are also interacted with whether or not the ad took place during the open enrollment period. To assess whether observed advertising is potentially in the flat part of the advertising response curve, a quadratic term of GRPs is also included. These interactions are all included using the border approach with the border sample, (brand-)county fixed effects, and (brand-)border-year fixed effects.

The results are presented in Table 6. Column (1) shows heterogeneity in the treatment effect of total GRP on category demand. First, advertising is not disproportionately moving lower-health-risk patients into MA, which was the main basis for regulatory concern. In fact, the point estimate suggests that advertising works better on less healthy counties, though the effect is small and insignificant. Advertising and premiums have no significant interaction, though advertising appears to work better in counties where a large fraction of plans offer enhanced drug benefits together with MA plans.[15] Advertising in the open enrollment

Table 6. Heterogeneity and Targeting

	(1) MA %	(2) Brand MA %
GRP ×	0.0074	0.0140
	(0.0070)	(0.0111)
GRP	0.00004	0.000005
	(0.00003)	(0.00008)
Risk	0.0007	0.0057
	(0.0033)	(0.0036)
Premium	−0.0018	−0.0029
	(0.0024)	(0.0034)
BasicDrugInc	0.0034	0.0105*
	(0.0028)	(0.0049)
EnhancedDrugInc	0.0060**	0.0041
	(0.0020)	(0.0043)
DrugDeductible	−0.0009	−0.0067
	(0.0018)	(0.0039)
BrandsPresent	−0.0014	0.0014
	(0.0021)	(0.0035)
OpenEnrollment	−0.0265***	−0.0136
	(0.0082)	(0.0114)
%Elderly	0.0042	0.0116**
	(0.0030)	(0.0043)
Income	−0.0057	−0.0123**
	(0.0028)	(0.0043)
%Hispanic	−0.0058	−0.0020
	(0.0076)	(0.0113)
%Asian	0.0072*	0.0217**
	(0.0032)	(0.0066)
%Black	0.0020	0.0043
	(0.0043)	(0.0040)
R-squared	0.968	0.918
Observations	4,591	10,087

Notes. Brand-DMA clustered standard errors are in parentheses. Column (1) uses border-year fixed effects, county fixed effects, and the same control variables as Table 5. Column (2) includes brand-border-year fixed effects, brand-county fixed effects, and the same control variables as Table 4.

$^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$.

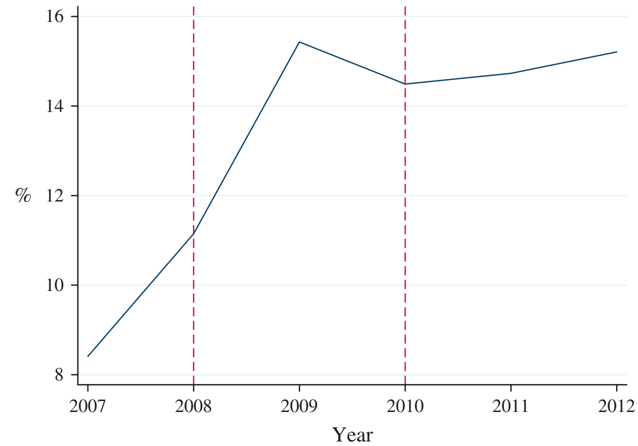
period appears to expand the category less than advertising during the rest of the year in moving customers from TM to MA. It could be that direct brand comparisons rather than general feelings about health insurance brands are more intense during open enrollment, but given that there are many interactions tested, it may also be a false positive. Advertising is also more category expansive in counties with higher shares of Asian population.

Column (2) shows heterogeneity in brand demand response to advertising. The interaction with average risk status in a county is positive, suggesting that advertising works better in less healthy counties, which works against the cream-skimming story, though it is not statistically significant. Advertising appears to work marginally better for brands that offer basic drug plans, but no significant interaction effect shows up either with premiums or with enhanced drug plans. In terms of managerially actionable targeting, advertising works better in counties with a higher share of elderly population, lower average income, and a higher percentage of Asian population. Given that poor health tends to be correlated with low income, the income effect may reflect further health effects that are not captured in the risk scores. Note that this analysis is at the county level, so I cannot definitively rule out that advertising works only on the unusually healthy individuals within the unhealthy counties. However, if advertising works better on the healthy either because of the advertising copy or the ability of the healthy to engage with the ads, that should be reflected in advertising working better in healthier counties. It is difficult to tell a story that advertising on television can be specifically targeted to the healthy individuals in unhealthy counties while not affecting healthy individuals in healthy counties.

The interaction effects also inform the degree to which sample selection to the borders might affect the estimated average treatment effect of advertising. Recall that the border counties had lower percentages of Asian population but also lower income and higher elderly populations. On net, these sample selection effects combined with the interactions roughly offset, suggesting that the average overall effect of advertising is roughly equal in the border sample and the interior of the DMA.

5.2.3. Prisoner's Dilemma. The main results suggest that rival advertising has a small positive effect on brand demand, but that could possibly be true only if the own brand also advertises in equilibrium. It could be that advertising and rival advertising are strategic complements and primarily business stealing in nature, leading to large investments in advertising without much change to brand enrollments. To provide a direct test of this prisoner's dilemma theory of advertising, UnitedHealthcare's two-year cessation

Figure 9. (Color online) UnitedHealthcare Shares Around Cessation



Notes. The figure shows the average UnitedHealthcare brand share over time. Vertical dashed lines represent the years with advertising turned off.

of advertising is leveraged. Note that UnitedHealthcare does reasonably well in terms of brand share over the time in which it is not advertising, as can be seen in Figure 9. The effect of rival ads on United's share during the cessation will also test the hypothesis that advertising is required to enhance the usefulness of some other marketing lever, such as brokers. If advertising makes access to brokers easier or enhances some other marketing lever, failure to advertise will result in lost share. Table 7 presents the results of the regression analysis using United's brand share as the dependent variable. Column (1) presents the correlational results when not using the border strategy, controls, or fixed effects. While the point estimate implies a negative effect of a rival advertising GRP on United's brand share of −0.0778, it is small and implies a CPC, or a cost of stealing an enrollee from United, of $1,465.43, which is rather high. However, the bad press could have directly lowered shares. United also could have lowered premiums to compensate for its lack of advertising.[16] To address these concerns, column (2) includes control variables, including United's premium and drug benefits, the average premiums and drug benefits of other brands, demographics, and year fixed effects. The effect size shrinks and becomes insignificant, with an average effect of −0.02 and a corresponding CPC of $5,697.01. The sign of the estimate flips once county fixed effects are included in column (3), making the average CPC infinite. Column (4), which moves to the border sample but maintains the specification, leaves the effect unchanged. Moving to using the border approach in column (5), the point estimate is still positive, but now closer to zero at 0.015. This leaves the average CPC at infinity and the left edge of the CPC confidence interval at $1,717.54 to steal an enrollee from United while it is not advertising. These

Table 7. United Share (%)

	(1)	(2)	(3)	(4)	(5)
RivalGRP	−0.0778*	−0.0200	0.0354	0.0094	0.0150
	(0.0371)	(0.0281)	(0.0313)	(0.0428)	(0.0415)
AvgPremium		0.0272	−0.0939**	−0.0703	−0.1022
		(0.0662)	(0.0285)	(0.0784)	(0.0642)
RivalAvgPremium		−0.1647***	−0.0675*	−0.0956*	0.0531
		(0.0381)	(0.0271)	(0.0405)	(0.0526)
EnhancedDrug		9.7610***	5.2471	7.2301*	2.6904
		(2.6852)	(2.7744)	(3.3923)	(5.0395)
RivalEnhancedDrug		9.8199	12.0867***	11.0614*	−1.7802
		(5.7220)	(3.2740)	(5.0562)	(8.4487)
BasicDrug		8.2042	3.5255		
		(11.2753)	(4.4739)		
RivalBasicDrug		5.9494	15.4557*	10.4666	1.9554
		(15.2300)	(7.5713)	(11.4358)	(13.8341)
RivalDrugDeductible		−0.0116	0.0441*	0.1002**	−0.0276
		(0.0471)	(0.0199)	(0.0353)	(0.0526)
BrandsPresent		−6.0984***	−3.2153***	−5.6767***	−3.8384
		(1.1553)	(0.7749)	(1.5956)	(2.4311)
Demo controls		×	×	×	×
Brand-year FEs		×	×	×	
County-brand FEs			×	×	×
Border sample				×	×
Brand-border-year FEs					×
CPC	\$1,465.43	\$4,750.38	\$∞	\$∞	\$∞
CPC CI	[757.74, 22,200.00]	[1,484.04, ∞]	[5,946.12, ∞]	[1,803.70, ∞]	[1,895.73, ∞]
Mean united %	25.37	25.30	25.60	22.079	22.079
R-squared	0.015	0.215	0.947	0.900	0.907
Observations	1,867	1,850	1,405	288	288

Notes. Brand-DMA clustered standard errors are in parentheses. Only counties with at least two competitors present are included (a firm cannot steal or lose brand share when it is fixed at 100%). *DrugDeductible* (own drug deductible) is excluded because no UnitedHealthcare plans in the sample include drug deductibles. *BasicDrug* (own basic drug) is excluded in columns (4) and (5), because during this time period, only 0.16% (three) of UnitedHealthcare's contracts have a basic drug benefit, and none of those is in the border sample. FE, Fixed effect.
$^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$.

results provide direct evidence against the prisoner's dilemma explanation theory of advertising and show that other firms in the market did not gain brand share from United's cessation of advertising.

5.2.4. Long-Run Effects. While all of the above analysis provides evidence about the usefulness of advertising in the year it airs, it could still be the case that advertising works over a longer-term horizon. The effects of advertising might be expected to be long lasting, as Curto et al. (2015) find that roughly 77% of MA enrollees stick with the same plan each year. If advertising were responsible for converting an enrollee, it is possible that the enrollee would stick longer than one year, even if advertising were to cease. To test systematically the degree to which advertising effects persist, the border approach is interacted with a stock conception of advertising. In this case, advertising is assumed to persist at a geometric rate, δ, per year for four candidate values of δ: $0.5, 0.6, 0.7$, and 0.8. If enrollees acquired by advertising stuck to plans as frequently as the average enrollee, using $\delta = 0.77$ should leave the short-run advertising effect previously estimated unchanged. If, alternatively, advertising did not have a long-lived effect, making the independent variable a function of past advertising would be similar to introducing right-hand-side measurement error and would attenuate the estimate toward zero. As mentioned previously, if the persistence rate of advertising is 0.61 and the profit per enrollee is as described by the average in the GAO report (Cosgrove et al. 2013), then advertising would be a break-even proposition. Advertising effects could fail to be long lived if either advertising marginal enrollee were less sticky than average or if the advertising had to be maintained to keep the enrollee, as would be the case if advertising were a direct utility shifter.

Table 8 provides the results of the effects of advertising stock on brand demand. Columns (1)–(4) show the analysis with persistence parameters of 0.5, 0.6, 0.7, and 0.8, respectively. In all four columns, the estimated advertising effect is smaller than the estimated short-run effect from before. Additionally, the four columns cannot be statistically distinguished from one another, nor from zero. These estimates provide evidence against long-lived effects of advertising,

Table 8. Brand-Level Demand—Stock (%)

	(1) $\delta = 0.5$	(2) $\delta = 0.6$	(3) $\delta = 0.7$	(4) $\delta = 0.8$
GRPStock	−0.0006	0.0010	0.0030	0.0050
	(0.0042)	(0.0040)	(0.0037)	(0.0035)
RivalGRPStock	0.0017	0.0012	0.0007	0.0002
	(0.0029)	(0.0028)	(0.0026)	(0.0023)
R-squared	0.972	0.972	0.972	0.972
Observations	3,717	3,717	3,717	3,717

Notes. Brand-DMA clustered standard errors are in parentheses. All specifications use brand-border-year fixed effects and county fixed-effects with a full set of control variables. The following included control variables are averages across plans within a brand-county-year: *AvgPremium*, *BasicDrugInc*, *EnhancedDrugInc*, and *DrugDeductible*. The following included control variables are normalized to have mean zero and standard deviation one and are common across brands in a county-year: *BrandsPresent*, *Risk*, *CountyBenchmark*, *Income*, *%Elderly*, *%White*, *%Asian*, *%Black*, and *%Hispanic*.

either from particularly sticky advertising marginal enrollees or from some other long-run effect, such as the building of brand equity. Importantly, they provide evidence that any persistence rate of advertising is below 0.6, indicating that accounting for either sticky advertising marginal enrollees or traditional advertising carryover effects, the average estimated ad effect is not profitable. It could be that advertising marginal enrollees are less sticky than the average customers, or it could be that for the advertising marginal enrollees, advertising goes into the yearly utility function, so the firm must keep advertising up to keep them from leaving in a subsequent year.

An alternative way to consider the dynamic effects of advertising is to specify a model directly incorporating inertia, through the inclusion of a lagged dependent variable. Such a model might not be ideal in this scenario, as there are only six time periods. If such a model included county fixed effects, the inertia parameter would be underestimated. To see this, note that differencing would mechanically cause the error term to be positively correlated with the lagged dependent variable and the outcome. If, alternatively, county fixed effects were excluded in such a model, the inertia parameter would likely be overestimated, as it would be confounded with county-level persistent preferences. Having noted these limitations, I estimate both of these models in Appendix E and find results consistent with the analysis using ad stock performed here. The incorporation of dynamic effects shrinks the estimated static effect of advertising considerably, indicating little long-run effect of advertising overall.

5.2.5. Discussion. The exploration of mechanisms provides evidence against both the long-run effectiveness of advertising in this setting and the prisoner's dilemma hypothesis of advertising effectiveness. It also suggests that the effects on the category overall are small and directionally go against the regulator concerns over cream skimming. From a targeting perspective, there is some heterogeneity in the advertising effect on brand demand that might be actionable to managers. Advertising works better in counties that have a larger share of elderly population as well as counties that have a larger percentage of Asian population. However, some of the heterogeneous effects provide murkier implications. For example, advertising works better in lower-income counties, which tend to be less healthy and might be less profitable.

5.3. Overall Results

All of the results together potentially raise a puzzle. Why exactly does broadly targeted advertising persist in equilibrium in this market? One potential explanation is that even when using numerous control variables, the advertising effect appears to be economically quite large. It is not until quasi-exogenous variation is carefully employed that the ad effect shrinks considerably. Managers could have difficulty in finding good variation to measure the causal effects and mistakenly use spurious correlations as their indicator of how well advertising works. Previous research, such as Blake et al. (2015), has found instances where managers struggled to make correct decisions with regard to advertising because of poor measurement of causal effects.

Additionally, managers in health insurance could give little attention to television advertising by outsourcing decision making to ad agencies. Revenue for all MA plans is over $100 billion per year, and advertising spend is only around $400 million. This implies a relatively small advertising-to-sales ratio compared with other industries. Advertising agencies might well have very different incentives from the firms contracting their services, which could lead to poorly executed advertising strategy.

Even if all incentives were perfectly aligned and firms used quasi-exogenous variation that improved both accuracy and power, as it does here, it is possible that for some values in the confidence interval, advertising could provide positive long-run return on investment (ROI). Since I cannot see incremental firm profits, it is impossible to say for sure using these data. However, it is clear that advertising is a relatively expensive way to acquire customers.

6. Conclusion

In this paper, the effect and mechanisms of television advertising for health insurance are explored. Policy makers are very concerned with advertising in the MA market, as evidenced by their numerous publications governing the proper conduct of advertising by firms. They are mostly concerned with firms trying to cream skim a favorable risk pool through advertising, as well as attempting to mislead seniors. Conversely, there is a concern that regulation may cause inefficiency in the marketplace, as consumers may be exposed to a lower-than-optimal level of information about each plan.

Leveraging the discrete borders of television markets, the effect of advertising is estimated to be small. The 95% confidence intervals suggest that even at the most optimistic, advertising effects imply large customer acquisition costs. Meanwhile, a more naïve approach, even including a large number of control variables, would lead the researcher to conclude that advertising is more than 10 times as effective as the more careful approach. The naïve correlation suggests that advertising is a very inexpensive way to acquire customers relative to static profits and relative to acquiring customers through price reductions.

Furthermore, this study provides evidence on mechanisms. Perhaps easing regulatory concerns, advertising has a small overall effect on the share of seniors who choose MA over TM, and a careful look at how that effect varies by health status reveals no systematic relationship between ad effect and the average health risk of a county. Advertising is found to work better on some types of counties than others, particularly those that have lower income and those with high elderly and Asian shares of the population. UnitedHealthcare's cessation of advertising reveals no significant negative effect from rival advertising when own advertising is stopped for an extended period of time, providing evidence against a prisoner's dilemma theory of advertising. Finally, evidence is provided that, even though this is a particularly sticky market, advertising effects are not particularly persistent. Combined, these results suggest that advertising is a relatively costly way to acquire additional seniors, and regulating advertising would provide limited change in market outcomes.

Puzzles remain for future research. First, the reason for the small advertising effect is unclear. Whether the lack of advertising efficacy is due to the current set of regulations or poor advertising copy is an interesting question with important firm implications. If the regulations make it unlikely for advertising to work better, then managers should scale back advertising on television considerably and focus on other areas of marketing strategy. If poor ad copy can explain the small effects, then managers might be able to improve the effectiveness of television ads. Second, given that advertising is estimated to have limited effectiveness in both the short and long run, it remains a puzzle that firms are spending hundreds of millions of dollars on this form of promotion. Since estimating advertising effects could be costly, it might simply be worth taking the gamble that advertising might work given the industry's large revenues. These results further highlight the results of Lewis and Rao (2015) describing the difficulty in accurately measuring ad effects given that the CPC confidence intervals can be quite large and advertising might be worthwhile at the lower end of these confidence intervals. Conversely, agency issues could cause firms extra difficulty in setting optimal advertising budgets.

Finally, for a policy maker worried about cream skimming through marketing levers, future research might be done looking into the pricing, entry, and plan design decisions of firms in MA. As I find limited evidence of cream skimming through advertising, observed advantageous selection by MA firms must be coming primarily from other levers.

The estimates from this study imply that concerns about cream skimming and deception due to advertising may be overblown. However, with advertising having little effect on enrollments, the concerns about the deadweight loss due to regulating advertising are also mitigated. Finally, estimates in this study suggest that firms are potentially making systematic mistakes in their advertising strategies and highlight the difficulty in making accurate assessments.

Acknowledgments

The author thanks Neale Mahoney, Sanjog Misra, Brett Hollenbeck, Ken Wilbur, Dan Bartels, Garrett Johnson, Carl Mela, Elisabeth Honka, and seminar participants at the 2015 Junior Health Economics Conference, the 2016 Choice Symposium, the 2017 Summer Institute in Competitive Strategy, Princeton University, the University of Pennsylvania, the Toulouse School of Economics, Tilburg University, Duke University, the University of Naples, the Harris School of Public Policy, and the Rady School of Management at the University of California, San Diego, for helpful comments. The author is grateful to Mariel Schwartz for excellent research assistance.

Appendix A. Sensitivity to Border Size

In the main analysis, the sample is restricted to those borders that make up no more than 35% of the DMA as a whole. In this section, I test sensitivity to this cutoff in the analysis of the main effects. Extending this sensitivity check to every regression in the paper is straightforward, but is not provided here for the sake of brevity, available upon request. Results are in Table A.1. While the point estimates get noisier as the threshold moves down to 10% (as fewer observations are used), both the direction and economic significance of the results remain unchanged. The average CPC remains similar and large across all specifications. While the point estimates are not all statistically significant, the confidence intervals around those point estimates are all very near to those in the main analysis.

Additionally, I consider borders that make up only a small population share of the DMA. If the counties in the border area make up less than 35% of the total counties, but still comprises a large fraction of the population of the total DMA, the border areas could be a main contributor to DMA level demand shocks. This would induce bias if firms targeted

Table A.1. Brand Share (%)—Sensitivity to Border Size

	(1) 10%	(2) 20%	(3) 30%	(4) 40%	(5) 50%
GRP	0.0016	0.0048	0.0080	0.0088	0.0072
	(0.0064)	(0.0059)	(0.0045)	(0.0046)	(0.0046)
RivalGRP	0.0056	0.0019	0.0027	0.0026	0.0026
	(0.0051)	(0.0029)	(0.0028)	(0.0027)	(0.0025)
CPC	$7,640	$2,510	$1,500	$1,370	$1,660
CPC CI	[848, ∞]	[738, ∞]	[712, ∞]	[677, ∞]	[737, ∞]
Mean brand %	4.21	4.50	4.73	4.98	5.05
R-squared	0.871	0.895	0.900	0.920	0.925
Observations	1,194	4,544	9,374	11,452	13,353

Notes. DMA clustered standard errors are in parentheses. All specifications use brand-border-year fixed effects and brand-county fixed effects with all controls. Each column corresponds with the maximum percentage of DMA counties the border constitutes.

Table A.2. Brand Share (%)—Sensitivity to Population Share of DMA

	(1) 10%	(2) 20%	(3) 30%	(4) 40%	(5) 50%
GRP	0.0123	0.0116*	0.0109*	0.0097*	0.0103*
	(0.0082)	(0.0055)	(0.0048)	(0.0046)	(0.0046)
RivalGRP	0.0040	0.0034	0.0027	0.0023	0.0029
	(0.0036)	(0.0032)	(0.0030)	(0.0029)	(0.0029)
CPC	$979	$1,040	$1,100	$1,240	$1,170
CPC CI	[425, ∞]	[536, 17,300]	[590, 8,510]	[644, 16,800]	[621, 10,100]
Mean brand %	4.61	4.74	4.98	4.97	4.97
R-squared	0.886	0.894	0.913	0.917	0.916
Observations	4,214	7,812	9,402	10,098	10,211

Notes. DMA clustered standard errors are in parentheses. All specifications use brand-border-year fixed effects and brand-county fixed effects with all controls. Each column corresponds with the maximum percentage of DMA population the border constitutes.
$^{*}p < 0.05$.

those demand shocks. Results are in Table A.2 and are consistent with the main analysis. Moving from restricting the sample to less than 10% population share borders to less than 50% population share borders does not significantly alter the results.

Appendix B. Using Log Enrollments Rather than MA Percentage and Brand Percentage

In the main analysis, the dependent variable in the brand demand analysis is the percentage of eligible seniors who

Table B.1. Brand Share (Log Enrollments)

	(1)	(2)	(3)	(4)	(5)
GRP	0.0119**	0.0139***	0.0033**	0.0046**	0.0017*
	(0.0041)	(0.0034)	(0.0011)	(0.0015)	(0.0008)
RivalGRP	0.0007	−0.0052***	−0.0001	−0.0001	0.0011
	(0.0018)	(0.0012)	(0.0008)	(0.0009)	(0.0006)
Brand-year FEs		×	×	×	
County-brand FEs			×	×	×
Border sample				×	×
Brand-border-year FEs					×
CPC	$194.00	$165.23	$694.47	$500.71	$1,365.69
CPC CI	[115.74, 599.06]	[111.63, 317.85]	[417.27, 2,068.80]	[304.09, 1,416.74]	[691.60, 54,000]
Mean log enrollments	5.6342	5.6401	5.6652	5.3289	5.3289
R-squared	0.014	0.407	0.933	0.899	0.950
Observations	36,004	35,741	35,154	10,406	10,406

Notes. Brand-DMA clustered standard errors are in parentheses. The following variables are included in columns (2)–(5) and are averages across all plans within a brand: *AvgPremium*, *BasicDrugInc*, *EnhancedDrugInc*, and *DrugDeductible*. The following variables are included in columns (2)–(5) and are normalized to have mean zero and standard deviation one: *BrandsPresent*, *Risk*, *CountyBenchmark*, *Income*, *%Elderly*, *%White*, *%Asian*, *%Black*, and *%Hispanic*. FE, Fixed effect.
$^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$.

choose the advertised brand. In this section, the dependent variable is changed to the natural log of brand enrollments, with the CPC computation adjusted to reflect the change in the dependent variable. The analysis of main effects is provided in Table B.1. All results are consistent with the main analysis and produce nearly identical CPC confidence intervals. Extending this to all of the analysis in the paper is straightforward, but for the sake of brevity is not provided here, with the intention of eventually including that in an online appendix.

Appendix C. Using Occurrences Instead of GRP

As noted in Section 3, the main analysis focuses on GRPs as the relevant measure of advertising. This generates some concern about measurement error. Outside of the top 25 DMAs, impressions are measured using self-reported diaries, which may be measured with considerable error. Meanwhile, advertising occurrences (the number of instances of an ad) are mechanically measured in the top 130 DMAs. In this section, occurrences are used as the relevant measure of advertising, and the analysis of main effects is repeated. For the sake of brevity, the exploration of mechanisms using occurrences as the relevant measure of advertising is not in this paper but available from the author on request, with the intention of eventually including that in an online appendix.

More noise might be expected when using occurrences as the relevant measure of advertising; that is, each 30 seconds of ad air time is coded as one occurrence, regardless of how many people see it. As such, an ad on a midnight rerun of I Love Lucy is coded exactly the same as an ad during the nightly news. Response to these two ads would be expected to be very different, potentially generating extra noise in the estimation. This is the reason why the main part of this study uses GRPs: it provides a theoretically reasonable way to weight each ad by how many people actually saw it.

Table C.1 presents the results. ROI measures are adjusted to reflect the average cost of an occurrence instead of the average cost of a GRP, as was the case in Table 4. Results are very similar to those in the main analysis, though with more noise. Without using the border strategy, advertising appears useful in lifting brand demand. When the border strategy is used, the average lift approaches zero. In using this alternative measure of advertising, moving from the simple fixed effects to the border approach provides a more dramatic contrast, highlighting the importance of a careful research design. The preferred specification using the border approach in column (5) shows a very similar CPC and CPC confidence interval as when GRP is used in the main analysis.

Appendix D. Response of UnitedHealthcare and Rivals to UnitedHealthcare's Advertising Cessation

Rival advertising and own and rival pricing are controlled for in the prisoner's dilemma analysis. However, responses to a major decrease in United advertising by these firms using pricing and advertising are of interest on their own. Here, I show these responses graphically and through regression. First, graphically, advertising responses by rivals to United's ad cessation are shown in Figure D.1. It appears that non-United advertising goes down a bit in 2008, increases to about 2007 levels in 2009, and goes back down in 2010. It is difficult to draw too strong conclusions from the time series picture alone. Next, premium responses to United's ad cessation are shown in Figure D.2. Since United is not present in all counties, the difference between counties that never had United and counties that include United can be instructive. It appears that during the advertising cessation, counties with and without United present follow roughly parallel trends in premiums. The counties with United, perhaps unsurprisingly, have lower levels of premiums across the sample. This is perhaps unsurprising for two reasons. First, having a major competitor present might provide downward pricing pressure. Second, as seen in the figure, United plans are characterized by very low premiums relative to the market, which indicates that it is perhaps a very tough price competitor. During the ad cessation, United premiums seem to continue on the preexisting gentle downward trend.

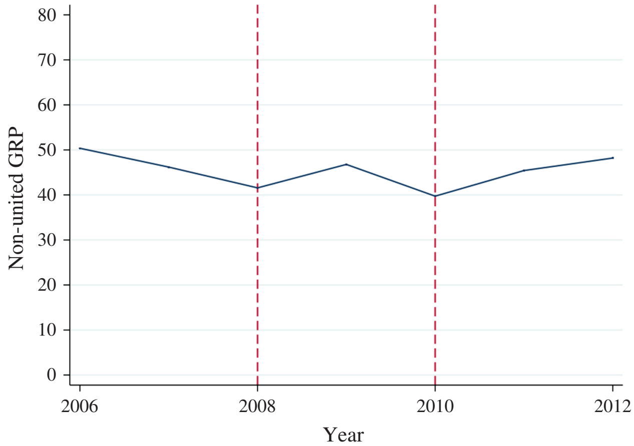
Table C.1. Brand Share (%)—Using Occurrences

	(1)	(2)	(3)	(4)	(5)
Ads	0.4013***	0.3866***	0.1023**	0.1487***	0.0335
	(0.0905)	(0.0989)	(0.0330)	(0.0420)	(0.0215)
RivalAds	−0.0384	−0.1279**	−0.0151	−0.0306	0.0038
	(0.0439)	(0.0461)	(0.0174)	(0.0250)	(0.0157)
Controls		×	×	×	×
County-brand FEs			×	×	×
Border sample				×	×
Border approach					×
CPC	$108.30	$112.42	$424.66	$292.25	$1,297.03
CPC CI	[75.10, 194.14]	[74.87, 225.52]	[260.08, 1,156.42]	[188.18, 653.81]	[574.15, ∞]
Mean brand %	5.4456	5.4365	5.4959	5.0191	5.0191
R-squared	0.031	0.124	0.896	0.854	0.918
Observations	36,764	36,489	35,919	10,651	10,651

Notes. Brand-DMA clustered standard errors are in parentheses. Included in controls are average brand premium, an indicator for zero average brand premium, number of competitors, share white, share black, share Hispanic, share elderly, MA county benchmark rates, and average county Medicare risk scores. FE, Fixed effects.

$^{**}p < 0.01$; $^{***}p < 0.001$.

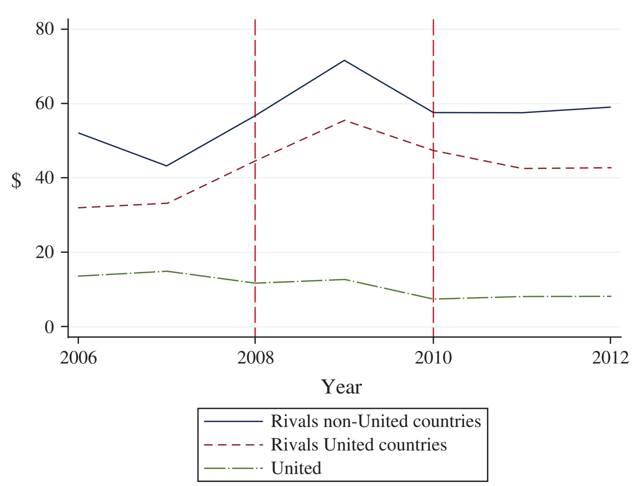
Figure D.1. (Color online) Rival Ad Responses to United



Notes. The figure shows non-United MA advertising over time. Dotted lines indicate years of United ad cessation.

To think more systematically about firm responses, a simple set of regression analyses is undertaken. First, for counties where United is present, I regress premiums of non-United plans on United GRPs, brand-year fixed effects, and brand-county fixed effects. As such, this will show whether firms adjust their premiums considerably when United drastically reduces its advertising in the cessation period. The results are in column (1) of Table D.1. Consistent with Figure D.2, it appears large changes in United advertising have no systematic relationship to changes in rival premiums. The point estimate is small, with each GRP decrease by United associated with a \$0.10 decrease in rival premiums, but the effect is not statistically significant. In column (2), I run the analogous regression but with United premiums as the dependent variable to see whether United itself adjusts premiums in response to a large change in advertising. There is no clear systematic relationship. If anything, the point estimate is negative, indicating an increase of premiums when advertising goes down, but it is again small and statistically insignificant.

Figure D.2. (Color online) Premium Responses to United



Note. The figure shows United and rival premium responses to United ad cessation, which is indicated by dotted lines.

Table D.1. Responses to United Ad Cessation

	(1) Non-United premium	(2) United premium	(3) Non-United GRP
UnitedGRP	0.1032	−0.1279	0.1682*
	(0.2244)	(0.3060)	(0.0840)
Mean DV	44.4122	10.3867	9.1747
R-squared	0.775	0.767	0.770
Observations	13,613	4,969	3,370

Notes. Brand-DMA clustered standard errors are in parentheses. Columns (1) and (2) include only counties in which United is present in the market. Columns (1) and (2) include brand-county fixed effects and brand-year fixed effects. Column (3) includes brand-DMA fixed effects and brand-year fixed effects. DV, Dependent variable.
*$p < 0.05$.

Finally, in column (3), I examine whether rivals adjust their advertising in response to United's decrease of advertising. In this regression, brand GRP is the dependent variable that is regressed on United GRP, brand-DMA fixed effects, and brand-year fixed effects. The point estimate indicates that for every one GRP decrease in advertising by United, rivals decrease advertising by 0.17 GRPs on average. This result is statistically significant, but with a reasonably large confidence interval.

Appendix E. An Inertia Model of Long-Run Advertising Effects

In this appendix, I consider dynamic effects of advertising that enter purely through inertia. The assumption of this model will be that once a customer is converted, by advertising or otherwise, she will remain in the plan in the subsequent year with some probability to be estimated. I incorporate this intuition by specifying a lagged dependent variable model. Note that such a model might not be ideal in this scenario, as there are only six time periods. If such a model included county fixed effects, the inertia parameter would be underestimated. To see this, note that differencing would mechanically cause the error term to be positively correlated with the lagged dependent variable and the outcome. If, alternatively, county fixed effects were excluded in such a model, the inertia parameter would likely be overestimated, as it would be confounded with county-level persistent preferences. Note that these approaches assume that there is a homogeneous rate of inertia and that the advertising marginal customers have the same average rate of inertia as the average customers. This might very well be false, as the advertising marginal customers have been convinced to change plans through advertising and might well be more likely to switch again, particularly in response to future advertising. Having noted these limitations, I estimate both of these models.

Table E.1 presents the estimates from each of these models. In column (1), county fixed-effects are excluded, which leads to an expected overestimate of inertia. Indeed, we find an estimated rate of inertia of 0.97, which is considerably higher than the estimates from either Curto et al. (2015) or Handel (2013). The advertising effect associated with this specification is 0.0009, not statistically significant, and precisely estimated. It is more than an order of magnitude smaller than

Table E.1. Brand Share (MA %)

	(1)	(2)
LagMAPct	0.9703***	0.5990***
	(0.0108)	(0.0424)
GRP	0.0009	0.0024
	(0.0017)	(0.0026)
RivalGRP	0.0001	0.0026
	(0.0007)	(0.0022)
Mean DV	5.0282	5.0951
R-squared	0.938	0.948
Observations	10,641	10,417

Notes. Brand-DMA clustered standard errors are in parentheses. All controls are included as before.
*** $p < 0.001$.

the estimated effect of advertising in the static model. In this case, even though the inertia effect is likely overestimated, the estimated advertising effect completely disappears. In column (2), county fixed effects are included. As noted above, this is likely to lead to an underestimated inertia parameter. The estimated inertia rate in this specification is 0.599. This is smaller than the average inertia of 0.77 found in Curto et al. (2015), which is consistent with an underestimated average rate of inertia. The estimated advertising effect in this specification is 0.0025, which is a factor of 3.5 smaller than the advertising effect estimated in the static model. The estimated effect is not statistically significant and is reasonably precisely estimated. In both specifications, the incorporation of dynamics through inertia makes the estimated advertising effect disappear, providing evidence against substantial long-run gains from advertising.

Endnotes

[1] Advertising for MA makes up a share of total advertising slightly less than proportionate to its revenue share of gross domestic product (GDP). In 2011, MA advertisers spent approximately $426 million on television ads out of a total of $90.7 billion in advertising spend that year, or about 0.47% of total TV advertising. Meanwhile the MA market made about $117 billion in revenue in comparison with a GDP of $15.52 trillion in 2011, or about 0.7% of GDP.

[2] In particular, in the Medicare Marketing Guidelines, the Centers for Medicare and Medicaid Services details that firms may not discriminate based on mental or physical disability, health status, medical history, insurability, or genetic information, indicating a particular concern about cream skimming (https://www.cms.gov/Medicare/Health-Plans/ManagedCareMarketing/Downloads/CY-2018-Medicare-Marketing-Guidelines_Final072017.pdf, p. 8.

[3] In fact, after three years of zero advertising, UnitedHealthcare reentered the advertising market in a significant way in 2011.

[4] In their empirical application, Aizawa and Kim (2018) find advertising to be more than twice as effective overall and more effective on the healthy from 2001 to 2005, which is at odds with the results of this study. This difference could be attributed to differences in research design or differences in sample period, as Aizawa and Kim (2018) use the period from 2001–2005, before sophisticated risk adjustment policies were put into place to try and remove the incentive to cream skim.

[5] MA was previously known as Medicare Part C, or Medicare+ Choice. I use the current naming convention throughout this paper.

[6] See http://www.sbca.com/dish-satellite/dma-tv.htm.

[7] See https://web.archive.org/web/20140724100219/https://www.fcc.gov/encyclopedia/evolution-cable-television.

[8] Additionally, conditional on the fixed effects in the model, ad occurrences predict the measure of impressions precisely and nearly identically in the top 25 DMAs and the DMAs ranked 26–130, suggesting that measurement error is not random, but systematic by DMA. Conditional on the DMA fixed effect, changes in impressions over time appear to be reasonably well measured. Details are available from the author on request.

[9] These control variables include average premiums, drug benefit design, brands present in the county, county benchmark rates, county average risk scores, and demographic variables.

[10] The Government Accountability Office (GAO) computed that average revenues per enrollee were $9,893, with an average profit margin of 4.5% in 2011. Of note is that advertising cost is included in nonmedical expenses that have already been subtracted out of revenues to obtain the 4.5% number. While advertising costs are small relative to other nonmedical expenses such as administrative expenses at about $59 per enrollee, they are added back in to the 4.5% to compute incremental static profit. This amounts to an average one-year profit of $504 per incremental enrollee. More details are available at Cosgrove et al. (2013).

[11] The long-run effect measured is a combination of the inertia from advertising marginal customers and the direct carryover of advertising. Separating the two using aggregate data is infeasible. However, Leone (1995) finds that, correcting for aggregation biases, support in the literature is for advertising to be worn out in six to nine months. If true, this would indicate that the dynamics in this case should come entirely from inertia.

[12] Failure to include data earlier than 2006 to build this ad stock would bias results weakly upward so long as 2005 and previous advertising had a weakly positive effect on demand. To see this, note that all incremental enrollments from ad stock would be attributed to advertising between 2006 and the present, even though 2005 advertising was actually responsible for some of that effect.

[13] If the mechanism through which sample selection might be a problem is through the average population in the county, I note here that running the specification in Table 4, column (3), while only including counties of over 100,000 people, does not significantly change the result—advertising in these urban counties still appears to work less well than advertising at the border. Furthermore, if the advertising effect in column (3) is directly interacted with the county population, the interaction indicates that, if anything, advertising works a bit better in less populous counties. The results of these regressions should alleviate the concern that the small advertising effect found in the border approach is due to the low population in the border counties. Details of these regressions are available from the author on request.

[14] In particular, current advertising could provide future benefit if either advertising marginal consumers this year exhibit inertia and stay enrolled in the future or if this year's advertising directly causes a customer who does not enroll this year to enroll next year. On average, firms would need the advertising effect to persist at a rate of 0.61 (through these mechanisms combined) with no discounting of the future to break even given the GAO estimate of yearly profitability; that is, if 100 customers were converted in 2009 due to 2009 ads, 61 customers would need to be converted in 2010 due to 2009 ads, either from being one of the 100 in 2009 that was sticky or from being directly caused by 2009 ads to enroll in 2010. Note that the relevant inertia is not the average inertia among all customers, but the inertia among those customers who chose the MA plan in question because of advertising (i.e., the advertising marginal). The inertia in that group could be more or less than the average inertia. Average inertia was documented by Curto et al. (2015) in MA to be about 0.77.

Handel (2013) finds a similar rate of inertia in the context employer-provided health insurance but with considerable heterogeneity, indicating that inertia among the advertising marginal customers could be different from the average. If we take the average inertia found in Curto et al. (2015) and assume it is the same as the inertia among the advertising marginal and discount future profits by 10% per year, a new customer would be worth $1,642, implying a long-run ROI of advertising if there are no broker fees for advertising related conversions. However, the advertising marginal might be expected to exhibit considerably less inertia than the average customer, given that advertising caused them to switch in the first place and advertising by a rival might well cause them to switch in the future.

[15] I note that customers most sensitive to generous drug benefit plans are likely to be those needing more drugs and are thus less healthy.

[16] To see how United and rivals adjusted their pricing and advertising strategies in response to the United advertising cessation, please see Appendix D.

References

Abaluck J, Gruber J (2011) Heterogeneity in choice inconsistencies among the elderly: Evidence from prescription drug plan choice. *Amer. Econom. Rev.* 101(3):377–381.

Aizawa N, Kim YS (2018) Advertising competition and risk selection in health insurance markets. *Amer. Econom. Rev.* 108(3):828–867.

Bayer P, Ferreira F, McMillan R (2007) A unified framework for measuring preferences for schools and neighborhoods. *J. Political Econom.* 115(4):588–638.

Black SE (1999) Do better schools matter? Parental valuation of elementary education. *Quart. J. Econom.* 114(2):577–599.

Blake T, Nosko C, Tadelis S (2015) Consumer heterogeneity and paid search effectiveness: A large-scale field experiment. *Econometrica* 83(1):155–174.

Cabral M, Mahoney N, Geruso M (2018) Do larger health insurance subsidies benefit patients or producers? Evidence from Medicare Advantage. *Amer. Econom. Rev.* Forthcoming.

Cai X, Kreps G, McAuley J, Zhao X, Strollo M, Neuman T, Boortz K (2008) Pitching private Medicare plans: An analysis of Medicare Advantage and prescription drug plan advertising. Report, Kaiser Family Foundation. https://kaiserfamilyfoundation.files.wordpress.com/2013/01/7805.pdf.

Cooper AL, Trivedi AN (2012) Fitness memberships and favorable selection in Medicare Advantage plans. *New England J. Medicine* 366(2):150–157.

Cosgrove JC, Brudevold C, George S, Giusto G, O'Donnell B, Morrison ET (2013) Medicare Advantage: 2011 profits similar to projections for most plans, but higher for plans with specific eligibility requirements. U.S. Government Accountability Office, Washington, DC. Accessed July 12, 2018, https://www.gao.gov/assets/660/659836.pdf.

Curto V, Einav L, Levin J, Bhattacharya J (2015) Can managed competition work? The U.S. Medicare Advantage Program. Working paper, National Bureau of Economic Research, Cambridge, MA.

Dafny LS (2010) Are health insurance markets competitive? *Amer. Econom. Rev.* 100(4):1399–1431.

Dube A, Lester TW, Reich M (2010) Minimum wage effects across state borders: Estimates using contiguous counties. *Rev. Econom. Statist.* 92(4):945–964.

Duggan M, Starc A, Vabson B (2016) Who benefits when the government pays more? Pass-through in the Medicare Advantage program. *J. Public Econom.* 141:50–67.

Eastlack JO Jr, Rao AG (1986) Modeling response to advertising and pricing changes for "V-8" cocktail vegetable juice. *Marketing Sci.* 5(3):245–259.

Ericson KM (2014) When consumers do not make an active decision: Dynamic default rules and their equilibrium effects. Working paper, National Bureau of Economic Research, Cambridge, MA.

Forelle C, Bandler J (2006) The perfect payday. *Wall Street Journal.* Accessed July 12, 2018, https://www.wsj.com/articles/SB114265075068802118.

Handel BR (2013) Adverse selection and inertia in health insurance markets: When nudging hurts. *Amer. Econom. Rev.* 103(7):2643–2682.

Hartmann WR, Klapper D (2018) Super bowl ads. *Marketing Sci.* 37(1):78–96.

Holmes TJ (1998) The effect of state policies on the location of manufacturing: Evidence from state borders. *J. Political Econom.* 106(4):667–705.

Ito K (2014) Do consumers respond to marginal or average price? Evidence from nonlinear electricity pricing. *Amer. Econom. Rev.* 104(2):537–563.

Johnson GA, Lewis RA, Reiley DH (2016) When less is more: Data and power in advertising experiments. *Marketing Sci.* 36(1):43–53.

Leone RP (1995) Generalizing what is known about temporal aggregation and advertising carryover. *Marketing Sci.* 14(3, Part 2):G141–G150.

Lewis RA, Nguyen D (2015) Display advertising competitive spillovers to consumer search. *Quant. Marketing Econom.* 13(2):93–115.

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

Lewis RA, Rao JM, Reiley DH (2011) Here, there, and everywhere: Correlated online behaviors can lead to overestimates of the effects of advertising. Ghinita G, Punera K, eds. *Proc. 20th Internat. Conf. World Wide Web* (ACM, New York), 157–166.

Lodish LM, Abraham M, Kalmenson S, Livelsberger J, Lubetkin B, Richardson B, Stevens ME (1995) How TV advertising works: A meta-analysis of 389 real world split cable TV advertising experiments. *J. Marketing Res.* 32(2):125–139.

McGuire TG, Newhouse JP, Sinaiko AD (2011) An economic history of Medicare Part C. *Milbank Quart.* 89(2):289–332.

Mehrotra A, Grier S, Dudley RA (2006) The relationship between health plan advertising and market incentives: Evidence of risk-selective behavior. *Health Affairs* 25(3):759–765.

Sahni NS (2015a) Advertising spillovers: Field-experiment evidence and implications for returns from advertising. *J. Marketing Res.* 53(4):459–478.

Sahni NS (2015b) Effect of temporal spacing between advertising exposures: Evidence from online field experiments. *Quant. Marketing Econom.* 13(3):203–247.

Sethuraman R, Tellis GJ, Briesch RA (2011) How well does advertising work? Generalizations from meta-analysis of brand advertising elasticities. *J. Marketing Res.* 48(3):457–471.

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

Sinkinson M, Starc A (2017) Ask your doctor? Direct-to-consumer advertising of pharmaceuticals. *Rev. Econom. Stud.* Forthcoming.

Spenkuch JL, Toniatti D (2018) Political advertising and election outcomes. *Quart. J. Econom.* Forthcoming.

Town R, Liu S (2003) The welfare impact of Medicare HMOs. *RAND J. Econom.* 34(4):719–736.

Tuchman AE (2015) Advertising and demand for addictive goods: The effects of e-cigarette advertising. Working paper, Northwestern University, Evanston, IL.

Vranica S (2016) Probe said to find rebates lurking behind U.S. ad seals. *Wall Street Journal* (June 2), https://www.wsj.com/articles/ad-agencies-earn-rebates-from-media-companies-for-ad-spending-probe-finds-1464888210?mod=djemCMOToday.