Advertising in Health Insurance Markets*

Bradley T. Shapiro†

December 14, 2016

Abstract

The effects of television advertising in market for health insurance are of distinct interest to both firms and regulators, due to both the sheer size of the market and some of its more particular characteristics when it comes to consumer choice. 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. On the other hand, ads could provide useful information or remind people to reconsider their options, making regulation potentially welfare-reducing. Using the discontinuity in advertising exposure created by the borders of television markets, this study estimates television advertising to have zero average lift on both the share of seniors who choose private Medicare Advantage (MA) plans over government-provided Traditional Medicare (TM), as well as brand share conditional on MA purchase with enough precision to exclude positive ROI from the 95% confidence interval. Leveraging the unilateral cessation of advertising by United Healthcare, further evidence is provided that this result is not explained by a prisoner's dilemma equilibrium. Additionally, advertising is not more effective in counties with a healthier population, potentially easing the concern over cream skimming. The lack of average effect cannot be attributed to long-run effects of advertising. While the average effect of advertising is zero and firms tend to target broadly, evidence is provided that advertising is significantly more useful for increasing brand share in the most competitive markets, suggesting that firms could improve targeting to gain ROI.

*I thank 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, Duke University, the University of Naples, and the Harris School of Public Policy for helpful comments. I am grateful to Mariel Schwartz for excellent research assistance.

†Chicago Booth. Email: bradley.shaprio@chicagobooth.edu

1 Introduction

Television advertising by health insurance plans is large and growing, rising from about $250 million in 2004 to $500 million in 2012. With the implementation of the Affordable Care Act (ACA) Marketplaces and a more broad-based shift towards 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 that provide seniors with a private, government-subsidized alternative to Traditional Medicare (TM). 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 consumers into purchasing plans they did not need or were poorly suited to their preferences.[1]

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 consumers choose plans that better reflect their own preferences. In addition, because MA, and many other health plans, are purchased during an open enrollment periods, advertising might have the added benefit of alerting consumers about enrollment deadlines. And 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 towards studying advertising in health insurance markets by estimating the impact of television advertising on MA enrollments. Television advertising is measured using the AC Nielsen 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 Center for Medicare & Medicaid Services (CMS) over 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, co-insurance, 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 (2016), to assess the extent to which advertising increases the share of seniors that choose a private MA plan over a government-run TM

[1] https://www.cms.gov/Medicare/Health-Plans/ManagedCareMarketing/FinalPartCMarketingGuidelines.html

plan, as well as the extent to which advertising influences the choice of one brand of plan over another. To assess the potential that competitive ads simply cancel each other out in a kind of prisoner's dilemma, I leverage the fact that United Healthcare, one of the top firms in this market, unilaterally ceased advertising from 2008-2010. These sources of variation combined allow for the estimation of both short-run and cumulative effects of advertising as well as heterogeneity in advertising effects.

Using this variation, I find zero average lift of aggregate MA advertising on aggregate MA enrollment. The point estimate from the preferred estimation is near zero and precisely estimated. The 95% return on investment (ROI) confidence confidence interval has values that range from -126% to -45.5% from moving seniors out of TM and into MA. Meanwhile, a more naive approach would suggest a large positive ROI of advertising that comes from shifting seniors from TM to MA. One potential explanation for the small and precise aggregate effect is that advertising mainly works to steal business from other MA plans. Firms must keep advertising to maintain the status quo share and avoid competitors taking their enrollees. Previous research on plan choice in the employer-sponsored insurance market finds that consumers are substantially more likely to switch among managed care plans than between fee-for-service and managed care alternatives (e.g., Bundorf, Levin and Mahoney, 2011; Handel, 2013). Since MA plans are predominately managed care plans and TM is fee for service, similar substitution patterns might be expected in this setting. Using the border approach at the brand level, I find that advertising provides on average very little lift to brand shares conditional upon purchase of an MA plan. The 95% confidence interval in the preferred specification finds the advertising ROI from business stealing to be between -135% and -68.5%. Effects of competitor advertising are correspondingly small.

Three possibilities are assessed to determine the cause of the finding that advertising has no economically meaningful effects on average: 1) that competitive advertising cancels out in what amounts to a prisoner's dilemma, 2) that advertising provides benefits that can only be observed over a long period of time and 3) that advertising in this setting works on some populations but was poorly targeted by firms over the course of the sample. First, the fact that United Healthcare, one of the largest players in this market, unilaterally ceases advertising from 2008-2010 is leveraged. If advertising were a prisoner's dilemma, the advertising cessation would be predicted to result in decreased brand share for United. No such decrease in brand share occurs. Second, the long-run impact of advertising is assessed using a stock conception of advertising assuming different rates of ad stock decay as well as directly testing the effect of lagged advertising on current enrollments. No matter the conception of long-run effects posited, no economically meaningful effects are detected, and confidence intervals remain precise around zero. Finally, by interacting advertising with a number of firm, market and demographic variables, I find that advertising is potentially useful for brand stealing in the most competitive markets. As firms tend to target less competitive markets, our results provide evidence that firms are not successfully targeting their advertising. The results show that advertising does not work especially well on any particular demographic, nor is it more useful when prices are lower or when patients are healthier, perhaps easing the regulatory concern over cream skimming. There is also little evidence that the average zero effects are attributable to being on the flat part of the advertising response curve.

These results suggest that while firms might be trying to cream skim using advertising, they are not particularly successful at influencing choices of any consumers, healthy or otherwise. As such, concerns over the cream skimming and misinformation may be overblown. Since advertising does not affect MA plan choice, it is unlikely to affect the risk pool of MA enrollees. Potentially misleading advertisements, to the extent they exist, do not seem to affect choices. Of course, these zero 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 advertising has no effect, additional regulatory scrutiny seems unwarranted. On the firm side, there is room for improvement simply by targeting ads to the most competitive markets.

While the regulatory implications might be clear, a large puzzle remains. If television advertising is effective in neither the short- nor the long-run, why are firms spending hundreds of millions of dollars per year on television advertising? That advertising increases from 2004-2012 suggests that firms are not learning over time that advertising is ineffective.[2] It could be that firms have a high cost measuring their own advertising effectiveness or that there is a difficult-to-overcome principal-agent problem with advertising agencies,[3] 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, Nosko and Tadelis (2015)) have been documented in the advertising effectiveness literature.

The contribution of this paper is an empirical one, finding negative ROI from advertising in the market for health insurance for the elderly. Health insurance is an important market in particular, but

[2] In fact, after three years of zero advertising, United Health care re-enters the advertising market in a significant way in 2011.

[3] http://www.wsj.com/articles/ad-agencies-earn-rebates-from-media-companies-for-ad-spending-probe-finds-1464888210?mod=djemCMOToday

the results of this study should also should 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 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 (2015) shows that firms tend to target advertisements towards healthier markets while Mehrotra, Grier and Dudley (2006) find that ad content is targeted towards healthy patients, giving some credence to the regulatory concern. Duggan, Starc and Vabson (2016) find that firms advertise more in markets where the government pays them more per enrollee. Through a structural model, Aizawa and Kim (2015) find advertising to be effective overall and more effective on the healthy, which is at odds with this study. The magnitude of the Aizawa and Kim (2015) effect is replicated when the border strategy is not used. However, using the discontinuity in advertising at the DMA borders, the estimated advertising effect disappears.[4]

This paper also contributes to literatures on competition in health insurance (Dafny, 2010) and in the Medicare Advantage market (e.g., Curto et al., 2014; Cabral, Mahoney and Geruso, 2014; 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. Given these results, there might be concern that equilibrium outcomes are not allocatively efficient, and in fact, 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 mis-allocation, 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, some of which documents similarly small advertising effects (e.g., Blake, Nosko and Tadelis, 2015), while others document significant effects, even in the presence of the aforementioned principle-agent issues in the industry (e.g., Johnson, Lewis and Reiley, 2015; Shapiro, 2016). 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. In particular, as firms often target areas of historic and recent strength, reverse causality is a large concern in the estimation of advertising effects.

[4]This difference could be attributed to differences to research design or differences in sample period, as Aizawa and Kim (2015) use the period from 2001-2005, before sophisticated risk adjustment policies were put into place to try and remove the incentive to cream skim.

Documentation of this problem along with attempts at at a solution date back at least to Lodish et al. (1995), which uses split cable experiments to think about advertising effectiveness. Additionally, a recent stream of literature has used either field experiments online (e.g., Blake, Nosko and Tadelis, 2015; Johnson, Lewis and Reiley, 2015; Sahni, 2015*a*,*b*; Lewis and Nguyen, 2015), instrumental variables (e.g., Sinkinson and Starc, 2015; Hartmann and Klapper, 2014) or natural experiments (e.g., Shapiro, 2016; Tuchman, 2015; Spenkuch and Toniatti, 2015). This paper will follow the latter 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, Lester and Reich, 2010), the effects of right-to-work laws (Holmes, 1998), the effects of schools on home values (e.g., Black, 1999; Bayer, Ferreira and McMillan, 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 DMAs have few counties, we will only use those border areas that make up less than 35% of the counties in the DMA.

While this paper might be seen as presenting a null effect, it is an important and informative null result with regulatory and managerial implications. In particular, the documented effect is precisely estimated and very near zero using a plausible research design as opposed to a noisy estimate that is indistinguishable from zero. Indeed, the null hypothesis of positive ROI from advertising can be rejected with 95% confidence. As regulators work with limited resources to find interventions that work, knowing which ones will not work 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 (and clearly not obvious) for firms. Finally, failure to report zero results on important questions when the research design is sound could potentially bias our understanding of the world, with false positives that occur by chance being published and true negatives being left out, making any meta-analysis biased in favor of the false positives.

The rest of the 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 sixty-five years old in the U.S. receive health insurance coverage under the Medicare program. Historically, most seniors have enrolled in what is now called Traditional Medicare (TM), which is a public insurance program administered by the Center for Medicare and Medicare Services (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 (MA) 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. MA enrollment has skyrocketed in the last 15 years, fueled in part by legislation that has that increased payment to plans and lifted restrictions on entry by non-HMO plans. Since 2000, MA enrollment has risen considerably from nearly 0% to over 30% of Medicare Beneficiaries. See McGuire, Newhouse and Sinaiko (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: United Healthcare, Aetna, Humana, Cigna, Kaiser Permanente, and the Blue Cross & Blue Shield (BCBS) plans. While United 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 are 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 Medicare Advantage 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 Medicare Advantage could use their subsidies for MA plans that provided drug coverage.

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

2.2 Television Advertising

Firms can purchase advertising space on television in two ways. First, there is an upfront market each summer where advertising agencies and firms make deals for the upcoming year of television. Advertising purchased in the upfront 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. National advertisements are seen by everyone in the country tuned into a particular station, while local advertisements are only seen by households within a particular designated market area (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 that a consumer of cable or satellite dish gets with his or 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 often times was just the station that was easiest to pick up over the air. That is, if a county picks 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 FCC 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.

3 Data and Summary Statistics

3.1 Advertising

Advertising data from AC Nielsen's Media database from 2004-2012 is used in this study. The database tracks television advertising at the spot-time-DMA level for every product which advertises on television. For data quality reasons, the top 130 DMAs, which are indicated as "full discovery markets" by AC Nielsen, are used in this study. In the top 130 markets, all television advertising occurrences are measured using monitoring devices, while in many of the smaller DMAs, only advertising occur-

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

[7] http://www.fcc.gov/encyclopedia/evolution-cable-television

rences that match ads in the larger markets are included. 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 analysis will be repeated using ad occurrences as an alternative measure to see if 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 roughly 75% of the total advertising expenditure. Some of them, such as United Healthcare and Humana, spend a fraction of their advertising budgets on national ads, while the more local companies and the smaller national brands, which have more local strengths, primarily focus on local advertising. Even the large national brands often advertise locally, as they often have strengths in particular markets. As such, the number of Humana advertisements that a household sees is the sum of the national Humana advertisements (which every household in the country tuned into the station see) and the local Humana advertisements (which only the households in that particular DMA tuned into the station sees). Descriptive statistics about firm-level advertising and shares are presented in Table I.

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 households who watched an ad 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 I shows the evolution of health insurance advertising over the course of our sample as a fraction of total television advertising. In 2004, health insurance advertising makes up about 0.25 percent of a $100 billion total of television advertising. By 2012, that number has roughly doubled.

As mentioned previously, advertising is split between local, DMA-level and national advertis-

[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. Please contact the author if you would like further details.

ing. There is also considerable variation in ad spending within a year. Figure II shows both of these dynamics. Roughly 65% of advertising spending is local rather than national, which is unsurprising given that the business is inherently a local one. Advertising spending is also highly concentrated during the open enrollment period, which runs from October 15 through December 7 each year.

3.2 Enrollment

MA enrollments and plan characteristics are measured using data from the Centers for Medicare and Medicaid Services (CMS). Enrollments at the plan-county-month level from 2007-2012 are observed. However, with few exceptions, consumers choose their MA plans once per year during open enrollment. The enrollments decided upon in open enrollment will 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 consumers make their choices about the upcoming year's insurance. Plan characteristics such as premiums and deductibles 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. Data on the Medicare risk scores of each county, which help to determine the capitation payment rates to MA plans, and plan-level characteristics from CMS are also included. Combining all of the data, this study uses enrollment and advertising at the county-plan-year-level from 2007-2012, using only data from after the introduction of Medicare's Part D prescription drug benefit.

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 due to 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 large advertising effect where none (or a smaller one) exists.

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 in Shapiro (2016) to study the effects of television advertising on antidepressant demand, but is also used in Tuchman (2015) to study e-cigarette advertising, as well as in Spenkuch and Toniatti (2015) to study political advertising. Consumers who live on different sides of DMA borders face different levels of advertising, due to 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 while observing differences in advertising. To both improve precision and to control for any unobservables that are persistent within counties over time, the panel nature of the data is leveraged by the inclusion of 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 which are simultaneously correlated with changes in advertising and the MA share.

The top 130 DMAs contain 236 such within-state borders, 173 of which where the border areas make up no more than 35% of the total DMA population. In the main analysis, attention will be restricted to these borders, but sensitivity analysis around this cutoff will be conducted in the appendix. Each of these borders 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 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.

4.3 Econometric Model

Let i index counties, b index borders, j index brand and t index time. Let s_{bit} indicate the percentage of Medicare beneficiaries with MA coverage and let GRP indicate level of advertising. The effect of an increase in aggregate advertising GRPs on the choice between TM and MA is estimated with regressions of the form

$$s_{bit}^{MA} = \gamma_1 \sum_{j \in J} GRP_{jit}^{own} + \alpha_{bt} + \alpha_i + X_{bit}\alpha_X + \epsilon_{bit}, \tag{1}$$

where α_{bt} are border-time fixed effects, α_i are county fixed effects, and X_{bit} is a vector of control variables, including demographic, competitive enivronment and plan characteristics. The coefficient of interest γ_1 captures the casual effect of an increase in MA advertisements on the share of seniors choosing MA over TM. Since the variation is at the DMA level and includes repeated measurements over time for each county, standard errors are clustered by DMA in all category level analysis.

For a brand-level analysis, let s_{bjit} indicate the brand share of brand j. The effect of an increase in brand-level GRPs on a brand's share is estimated with regressions of the form

$$s_{bijt} = \gamma_1 GRP_{jit}^{own} + \gamma_2 GRP_{jit}^{rival} + \alpha_{bjt} + \alpha_{ij} + X_{bijt}\alpha_X + \epsilon_{bijt}, \tag{2}$$

where α_{bjt} are border-brand-time fixed effects, α_{ij} are brand-county fixed effects, and X_{bijt} is a vector of brand and county control variables, including demographic, competitive enivronment and plan characteristics. In this case, the coefficients of interest, γ_1 and γ_2, capture the casual 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 analysis.

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, the border-specific time fixed effects would sweep away all variation in advertising. Figures III, IV and V show that there is significant advertising across the borders in both total-market and brand-level GRPs, as well as rival GRPs.

Second, the placement of the borders must be quasi-random with respect to preferences for health insurance. As policies related to health insurance and health care 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 location of DMA borders were determined historically by AC Nielsen and have been changed rarely over time.

Table II presents evidence on the validity of the identifying assumption by examining differences in observable factors across these borders. In particular, Table II shows the estimated coefficient from a regression with observable characteristics as the dependent variable and DMA-level MA advertising as the independent variable. Consumer, county and plan characteristics that might be plausibly correlated with demand are used as dependent variables. None of the county demographics or plan level variables, including population, population over 65, average income, race or average premium are predicted by advertising at the border. In terms of unobservables, there are extensive arguments for the quasi-randomness of DMA borders in, Shapiro (2016), Spenkuch and Toniatti (2015) and Tuchman (2015). The maintained identifying assumption throughout is that trends in any such unobservables must be parallel across the DMA borders.

4.4 Features and Limitations

This approach has both features and limitations as compared with other approaches that a researcher might take to identify advertising effects on demand. Perhaps the largest feature of this approach is that the observed advertising levels at the border are out of equilibrium. That is, we see variation that is driven by the equilibrium in other markets as opposed to adding a little bit of random noise to an observed equilibrium. By doing this, it is possible to see advertising levels that are likely to be both well above and well below what would be optimal if firms targeted each county individually. This makes the estimated treatment effect approximate an average treatment effect across the advertising response curve. In an experiment or an IV approach where the researcher injects some noise into a pre-existing equilibrium or targeting rule, the estimated effect will only be local to levels of advertising that are near that equilibrium or targeting rule. As such, if the firm is already optimally allocating advertising spending, the incremental effect from a small amount of noise being added to the equilibrium might well be hard to pin down and smaller than the average effect. Second, because this approach does not require the use 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.

In terms of limitations, this 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. Table III shows how consumers in the border sample are systematically different from consumers outside of the border sample. Most notably, the average population in a border county is considerably smaller than in a county outside of the border sample. The border sample also has a higher percentage black population, a lower percentage Asian and Hispanic population and a higher percentage Medicare eligible population. If anything, intuition suggests the larger percentage of Medicare eligible in the population (though a small difference) would lead to a higher estimated advertising effectiveness for a product that only Medicare eligible 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.

4.5 Computing Return on Investment

Using the above approach produces the average treatment effect of advertising GRPs on enrollments, or the average lift of advertising. A more managerially relevant metric that also helps to frame the economic significance of the estimates is return on investment, or ROI. To convert the estimates from average lift to ROI, I use the cost of observed advertising from the data, computed at the average cost of a health insurance GRP in the sample. For the marginal revenue from advertising, the Government Accountability Office report on the expected profits of Medicare advantage plans is used.[9] ROI estimates are provided for both ends of the 95% confidence interval of average lift.

4.6 Mechanisms

In addition to finding main effects, this study attempts to pin down exactly how advertising might work. In particular, three hypotheses are tested. First, advertising might be a prisoner's dilemma whereby firms only advertise to 'cancel out' the advertising of their competitors, leaving equilibrium shares unchanged. Second, advertising might only work over a long period of time as a kind of brand-building. Finally, advertising might work differentially on different populations, making opportunity for managers to target their advertising disproportionately to those for whom it works better.

Given that the intuition of the border strategy is that it provides 'out-of-equilibrium' advertising

[9]The 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 non-medical 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% when computing ROI to avoid double-counting ad costs. This amounts to an average profit of $504 per incremental enrollee. More details are available at http://www.gao.gov/products/GAO-14-148

levels, the brand-level analysis in principle 'partials out' the prisoner's dilemma problem to find the partial effect of rival advertising. To further and more directly test the prisoner's dilemma hypothesis, I leverage the fact that United Healthcare unilaterally stopped advertising during 2008 and 2009. In 2006, a *Wall Street Journal* article detailed a scandal involving backdated stock options.[10] This led to an SEC investigation over the course of many years and the eventual resignation of CEO William W. McGuire. In the midst of the commotion and regulatory attention, United spent almost nothing on health insurance advertising between 2008 and 2009, as can be seen in Figure VI. If the prisoner's dilemma hypothesis holds, then rival advertising should deteriorate United's share over the years it does not advertise. This cessation interacted with the border strategy can provide additional insight on the rival effect when the own brand ceases to advertise.

To address the hypothesis of long-run effects of advertising, a goodwill stock conception of advertising is used instead of yearly GRP levels. Additionally, the use of one-year lags allow for a less parametric test of long-run effects. Because the border strategy does not rely on rare transient shocks, it provides variation in this long-run stock measure of advertising, allowing for the test of the long-run hypothesis.

Finally, to address treatment effect heterogeneity, the GRP measures are interacted with observable demographic, health, plan and competition characteristics that are of interest to both firms and regulators. 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.

[10] http://www.wsj.com/articles/SB114265075068802118

5 Results

5.1 Main Effects

5.1.1 Category Expansion

Table IV presents the estimation of the category-level effect from equation 1. To provide economic intuition for the size of the advertising effects, ROI estimates and confidence intervals are presented below the advertising effects, computed as outlined above.

Column (1) presents the naive regression, where all counties are included, and no fixed effects or controls are used. It suggests a positive and significant effect of advertising on the percentage of seniors who choose an MA plan rather than a TM plan, with an increase of one GRP associated with an increase in the percent of seniors enrolled in MA by 0.0756 percentage points, implying an average ROI of 415%. Column (2) adds observable control variables for demographics, plan characteristics, and year fixed effects as well as the number and identity of the competitors in the market. The effect size is almost unchanged and implies a 310% ROI from category expansion. 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 ad effect. In column (3), 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 long run average advertising in a market. Controlling for these factors makes the ad effect disappear. The average lift from category expansion is very close to zero and the average ROI is -114%. Meanwhile, the estimated effect is not especially precise, as the 95% confidence interval of the ROI in this specification is [-195%, -32.6%]. If firms target markets of not only historic strength but also of reccent strength or target unobserved concurrent demand shocks, column (3) will also over-estimate the advertising effect. In column (4), the border approach is employed to control for all remaining endogeneity issues. The effect of advertising on MA share is now more precisely estimated and not distinguishable from zero. The 95% ROI confidence interval ranges from -126% to -45.5% with an average ROI of -85.8%. While a simple correlation would lead the researcher to conclude that advertising is a highly profitable way to shift the elderly out of TM and into MA, a more careful analysis using plausibly exogenous variation shows zero average movement from TM to MA from advertising.

To provide a graphical illustration of these results, a bin scatter, using 100 bins, of MA percent and total GRP is presented in figure VII. These amounts are residualized by a county fixed effect and a border-year fixed effect, so they reflect variation across TV market borders, as in column (4).

5.1.2 Business Stealing

While the above results indicate that advertising is not effective in moving seniors out of TM into MA, they do not imply that there is no firm-specific positive ROI associated with advertising. There might simply be MA 'types' of consumers that do not include TM in their consideration sets. While a negative industry-wide ROI might be troubling for the health insurance industry, it might still be individually rational for each firm to advertise if by doing so it can steal customers from a rival's plan to its own plan. This zero-sum type game would provide incentive for large advertising expenditures, even if those expenditures did not expand the market.

A brand-level analysis estimating equation 2 is presented in Table V. Column (1) presents the naive regression including all counties and plans with no control variables. The estimate from this regression suggests that advertising moves customers into the advertised brand in a significant way and rival advertising has a negative effect on brand share. The average ROI implied by these estimates is about 139%. Column (2) adds in control variables for plan premiums, risk status of consumers and demographics as well as brand-year fixed effects. The controls leave the estimates mostly unchanged, and the implied average ROI is 158%. Column (3) adds in brand-county fixed effects to control for time invariant factors that make a brand strong in a particular county. While the fixed effects shrink the estimated effect considerably and move the ROI confidence interval to all negative values, the average lift remains positive, though not significant. Column (4) employs the border strategy with a brand-border-year fixed effect. Moving to the border strategy makes the effect of advertising disappear altogether, with an average implied ROI of -102% and a 95% confidence interval of [-135%, -68.5%], ruling out economically meaningful advertising effects due to business stealing. It appears that in the simple fixed effects model, there is a small spurious effect due to firms targeting either unobserved demand shocks or recent strength of these particular markets, making the border strategy necessary to control for endogeneity. The simple fixed effects is also much less precise than the border strategy, with a 95% confidence range of about 100%, while the border strategy has a confidence range of about 67%.

To provide a graphical illustreation, a bin scatter, using 100 bins, of brand share and own GRP (and rival GRP) are presented in Figure VIII (and Figure IX). Again, these variables are residualized by brand-county and brand-border-year fixed effects to reflect variation within brand across the borders of TV markets.

5.1.3 Discussion

For both category-expansion and business-stealing, naive regressions give a large over-estimate of advertising effectiveness on both margins. Using plausibly exogenous variation, the advertising effect disappears and the point estimates become more precise. For identifying the business-stealing effect of advertising, a county fixed effect is not enough to fully control for the endogeneity of advertising. Of course, the finding of a null result raises further questions as to the reason for the failure of advertising to be effective and the explanation of how advertising persists in equilibrium if it does not work.

5.2 Mechanisms

In this section, three explanations for the null effect of advertising are explored. First, it could be the case that in equilibrium, advertising provides no effect because everyone is already advertising in order to cancel out each others' advertising. Since the argument of the border strategy is that it provides out-of-equilibrium variation in brand advertising, it in principle controls for this possibility. Here, further evidence will be provided by United Healthcare's unilateral decision not to advertise for two years. Second, it could be the case that advertising only provides long-run benefits in the form of brand building. To test this, a goodwill stock conception of advertising will be employed as well as advertising lags. Finally, it could be the case that the null average effect of advertising is masking interesting heterogeneity in the advertising effects. If this were the case, there would be opportunities for firms to more appropriately target advertising to improve profitability.

5.2.1 Prisoner's Dilemma

To evaluate whether rival advertising would be effective if one competitor unilaterally ceased advertising, as would be predicted by the prisoner's dilemma hypothesis, United Healthcare's two-year cessation of advertising is leveraged. Before getting to regression results, it is notable that United Healthcare does reasonably well in terms of brand share over the time in which it is not advertising, as can be seen in Figure X. Table VI presents the results of the regression analysis. Column (1) presents the naive regression when not using the border strategy, controls or fixed effects. While the point estimate implies a negative effect of rival advertising on United's market share, it is small and gives way to an average ROI of -6% and a 95% ROI confidence interval that ranges from -58.1% to 46.0%. To address concerns that United may have simply lowered price in response to the advertising cessation, column (2) includes control variables, including premium, and year fixed effects. The effect size shrinks, with average ROI of -38.1%. The sign of the estimate flips once county fixed effects are included and the border approach is used in columns (3) and (4), respectively. In both specifications, the ROI confidence interval is strictly negative. These results also provide further evidence against the prisoner's dilemma explanation of the estimated null effect of advertising. It also means that the other firms in the market did not gain positive ROI by stealing from United while it was in its advertising hiatus.[11]

5.2.2 Long Run Effects

While all of the above analysis provides evidence against advertising working in the year it airs, it could still be the case that advertising works over a longer-term horizon. If this is the case, we might not expect to see an advertising effect unless advertising is turned off over a considerable amount of time. One benefit of the border approach is that it provides clear predictions in the case where one side of the border has persistently higher advertising than the other side. In such a case, shares of MA on the high advertising side of the border should be expected to grow at a faster rate than shares on the low advertising side of the border in a divergent way. The behavior of such borders is shown graphically in Figure XI. It does not appear that the high advertising sides of the border are diverging away from the low advertising sides of the border.

To more systematically think about the question of long-run effects in both category expansion and business stealing, both a stock conception of advertising and lagged values of advertising are employed. In the case of the stock conception of advertising, advertising is assumed to decay at a geometric rate δ per year for three candidate values of δ: 0.6, 0.8 and 1. In the case where δ=1, there is no decay in advertising's effectiveness. In using lags, one year of lagged advertising is included in addition to current-year advertising. Table VII provides these results for the potential category expansion effect of advertising. Columns (1)-(3) provide estimates of the effect of accumulated advertising stock. The effect is very small and insignificant in all cases. Column (4) presents the effect of current year GRP as well as one year of lagged GRPs. Both the current year effect and the lagged effect are very small and statistically insignificant. Put together, these results provide evidence against the long-run effectiveness of advertising for category expansion.

[11] In fact, if we think that there was some direct effect of the SEC investigation, these results might wrongly attribute some of United's losses to advertising rather than to the scandal. This would make the advertising ROI from stealing from United even worse than the estimates imply.

Table VIII provides the analogous results at the brand level. Columns (1)-(3) again show no statistically significant effect of advertising on brand shares for any level of depreciation of advertising stock. The estimated effects are small, but they are a bit noisier than the estimates of the short-run effect. Column (4) shows that the effect of advertising in the current year is small and insignificant while the effect of lagged advertising is positive and significant, though is near zero and provides -67% return on investment. To the extent that the previous year's advertising proxies for all past advertising, the effect is likely to be over-estimated. Together with the long run category-expansive effects, these results provide evidence against the story that advertising works well in the long run, even if it does not work in the short run. While there is some evidence of an effect of advertising in the previous year, the magnitude is very small.

5.2.3 Heterogeneity

A final explanation of the zero estimated effect of advertising is that the estimated zero masks interesting heterogeneity in advertising effects. It could be that advertising works well on some subset of the population and is ineffective on other subsets. As the estimated ROI is negative, even at the right edge of the 95% confidence interval, it should be noted that even if advertising is useful on some subset of the population, it appears as though firms target the subsets on which it was ineffective too frequently. It also implies that the firm may have scope to improve its targeting scheme. To assess heterogeneity in the treatment effect, advertising GRPs are interacted with variables about the health status, competitive environment and demographics of a county in addition to plan characteristics. In addition, a squared advertising term is included to capture non-linear effects. These interactions are then included in the estimation of equations 1 and 2. The results are presented in Table IX. Column (1) shows heterogeneity in the treatment effect of total GRP on the percentage of customers who choose MA over TM. None of the interaction terms are significant either statistically or economically. In particular, advertising is not disproportionately moving lower-health-risk patients into MA, which was the main basis for regulatory concern. Meanwhile, column (2) shows potentially more interesting heterogeneity at the brand level. For the purposes of increasing brand share, advertising is significantly more effective in the markets that have higher competition. The estimates do not detect concavity in the advertising returns function, nor does advertising appear to work any better when premiums are low or during the open enrollment period. The point estimate is consistent with advertising working better for business-stealing in markets with a larger elderly population, though the estimate is not statistically significant.[12]

Advertising being more effective in more competitive markets is consistent with a reminder theory of advertising. If consumers can only remember some brands at a time and many brands are available to choose from, advertising may help consumers to have the advertised brand at the top of mind. The estimate of advertising effectiveness remains small, however, and implies positive ROI only when all of the major brands are present plus at least one regional insurer. The implied average ROI in those markets is 13.6%. There are only 151 county-years (in a total of 10 DMAs) in the data fitting this description out of 15,508 total county-years in the data (less than 1% of county-years).

Given these results, firms should only target the most competitive of markets. However, this is not the case. Table X shows how firms target with respect to competition levels in these data. Overall, firms tend to target broadly, but generally more so to less concentrated markets, as if they believe the primary effect of advertising is moving consumers out of TM into MA. If that were the case, advertisers would be wise to target those markets without competition, as then the monopoly firm would get all of the benefit as opposed to some of the benefit spilling over onto rivals, as in Shapiro (2016).

5.2.4 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. Looking carefully at observed heterogeneity in the advertising effect, I find that advertising shows some effectiveness in the most competitive markets, indicating that firms would benefit from more careful targeting of their advertising. However, all of the results together 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 carefully employing quasi-exogenous variation that the ad effect disappears. 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, Nosko and Tadelis (2015), has found instances where managers struggled to make correct decisions with regards to advertising due to poor measurement of causal effects.

Additionally, managers in health insurance could give little attention to television advertising

[12] Table IX suppresses race and income interactions with advertising, which were included in the regressions, though none of them were significant.

by outsourcing decision making to ad agencies. The market for health care is a three trillion dollar market, and advertising spend is only around five hundred million. This implies a relatively small advertising to sales ratio compared with other industries. Figure XII shows that relative to the size of the industry, health advertising is reasonably small. Advertising agencies might well have very different incentives from the firms contracting their services, which could lead to poorly executed advertising strategy.

6 Conclusion

In this paper, the effect of television advertising for Medicare Advantage plans is explored. Television advertising is a strategy that policy makers are very concerned with in the MA space, 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 statistically insignificant and economically inconsequential. Indeed, 95% confidence intervals exclude positive ROI from advertising either from stealing brand share or from expanding the category. Meanwhile, a more naive approach, even including a large number of control variables would lead the researcher to conclude there was a large advertising effect on both margins where none exists. Further, this study provides evidence that this null effect cannot be explained by advertising working very well over the long run or by equilibrium effects whereby all firms cancel out each others' advertising efforts. However, advertising is significantly more useful in markets with many competitors for the purposes of stealing brand share. This result stands in contrast to current firm targeting practices, which target a broad range of markets in general and highly concentrated markets in particular. Most importantly from a regulatory perspective, this study provides evidence against the theory that advertising works primarily by cream skimming healthy consumers out of TM and into MA. This result should make regulators potentially feel better that further regulation of advertising is not needed.

Puzzles remain for future research. First, the reason for the failure of advertising 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, then managers should scale back advertising on television considerably and focus on other areas of marketing strategy. If poor ad copy can explain the failure, then managers might be able to improve the effectiveness of television ads. Second, given that advertising is estimated to be ineffective 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. Given that estimating advertising effects could be costly, it might have simply been worth taking the gamble that advertising might work given the industry's large revenues. Conversely, agency issues could cause firms extra difficulty in setting optimal advertising budgets.

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 strategy. These mistakes may be caused by organizational or agency issues that this study does not have the data to identify explicitly, but other a priori rational reasons to advertise in the presence of a zero marginal effect are inconsistent with the data.

References

Abaluck, Jason, and Jonathan Gruber. 2011. "Heterogeneity in choice inconsistencies among the elderly: evidence from prescription drug plan choice." *The American Economic Review*, 101(3): 377.

Aizawa, Naoki, and You Suk Kim. 2015. "Advertising competition and risk selection in health insurance markets: Evidence from Medicare Advantage." *University of Minnesota working paper*.

Bayer, Patrick, Fernando Ferreira, and Robert McMillan. 2007. "A unified framework for measuring preferences for schools and neighborhoods." *Journal of Political Economy*, 115(4): 588–638.

Black, Sandra E. 1999. "Do better schools matter? Parental valuation of elementary education." *Quarterly Journal of Economics*, 114(2): 577–599.

Blake, Thomas, Chris Nosko, and Steven Tadelis. 2015. "Consumer heterogeneity and paid search effectiveness: a large-scale field experiment." *Econometrica*, 83(1): 155–174.

Bundorf, M.K., J. Levin, and N. Mahoney. 2011. "Pricing and welfare in health plan choice." *American Economic Review*, 102(7): 3214–3248.

Cabral, Marika, Neale Mahoney, and Mike Geruso. 2014. "Does privatized health insurance benefit patients or producers? Evidence from Medicare Advantage." *NBER Working Paper*, w20470.

Cooper, Alicia L, and Amal N Trivedi. 2012. "Fitness memberships and favorable selection in Medicare Advantage plans." *New England Journal of Medicine*, 366(2): 150–157.

Curto, Vilsa, Liran Einav, Jonathan Levin, and Jay Bhattacharya. 2014. "Can managed competition work? The US Medicare Advantage program." *Mimeo*.

Dafny, Leemore S. 2010. "Are health insurance markets competitive?" *The American Economic Review*, 100(4): 1399–1431.

Dube, Arindrajit, T William Lester, and Michael Reich. 2010. "Minimum wage effects across state borders: estimates using contiguous counties." *The Review of Economics and Statistics*, 92(4): 945–964.

Duggan, Mark, Amanda Starc, and Boris Vabson. 2016. "Who benefits when the government pays more? Pass-through in the Medicare Advantage program." *Journal of Public Economics*, 141: 50–67.

Ericson, Keith M Marzilli. 2014. "When Consumers Do Not Make an Active Decision: Dynamic Default Rules and Their Equilibrium Effects." *NBER Working Paper*, w20127.

Handel, Benjamin R. 2013. "Adverse selection and inertia in health insurance markets: when nudging hurts." *The American Economic Review*, 103(7): 2643–2682.

Hartmann, Wesley R, and Daniel Klapper. 2014. "Super bowl ads." *Available at SSRN: http://ssrn.com/abstract=2385058*.

Holmes, Thomas J. 1998. "The effect of state policies on the location of manufacturing: evidence from state borders." *Journal of Political Economy*, 106(4): 667–705.

Ito, Koichiro. 2014. "Do consumers respond to marginal or average price? Evidence from nonlinear electricity pricing." *The American Economic Review*, 104(2): 537–563.

Johnson, Garrett A, Randall A Lewis, and David Reiley. 2015. "When Less is More: Data and Power in Advertising Experiments." *Available at SSRN 2683621*.

Lewis, Randall A, and Justin M Rao. 2015. "The unfavorable economics of measuring the returns to advertising." *The Quarterly Journal of Economics*, 130(4): 1941–1973.

Lewis, Randall, and Dan Nguyen. 2015. "Display advertising competitive spillovers to consumer search." *Quantitative Marketing and Economics*, 13(2): 93–115.

Lodish, Leonard M, Magid Abraham, Stuart Kalmenson, Jeanne Livelsberger, Beth Lubetkin, Bruce Richardson, and Mary Ellen Stevens. 1995. "How TV advertising works: a meta-analysis of 389 real world split cable TV advertising experiments." *Journal of Marketing Research*, 32(2): 125–139.

McGuire, Thomas G., Joseph P. Newhouse, and Anna D. Sinaiko. 2011. "An economic history of Medicare Part C." *Milbank Quarterly*, 89(2): 289–332.

Mehrotra, Ateev, Sonya Grier, and R Adams Dudley. 2006. "The relationship between health plan advertising and market incentives: evidence of risk-selective behavior." *Health Affairs*, 25(3): 759–765.

Sahni, Navdeep S. 2015*a*. "Advertising spillovers: field-experiment evidence and implications for returns from advertising." *Journal of Marketing Research*, 53(4): 459–478.

Sahni, Navdeep S. 2015*b*. "Effect of temporal spacing between advertising exposures: evidence from online field experiments." *Quantitative Marketing and Economics*, 13(3): 203–247.

Shapiro, Bradley. 2016. "Positive spillovers and free riding in advertising of prescription pharmaceuticals: the case of antidepressants." *Journal of Political Economy, forthcoming*.

Sinkinson, Michael, and Amanda Starc. 2015. "Ask your doctor? Direct-to-consumer advertising of pharmaceuticals." *N*, w21045.

Spenkuch, Jörg L, and David Toniatti. 2015. "Political advertising and election outcomes." *Available at SSRN 2613987*.

Town, Robert, and Su Liu. 2003. "The Welfare Impact of Medicare HMOs." *RAND Journal of Economics*, 34(4): 719–736.

Tuchman, Anna E. 2015. "Advertising and Demand for Addictive Goods: The Effects of E-Cigarette Advertising." *Mimeo*.

Table I: Advertising and Shares by Brand per Year

Brand	Average GRP	% Local	Brand Share (%)	Number of Counties
Aetna	5.051	80.60	11.53	171.7
BCBS	23.24	99.92	30.06	962.9
Cigna	0.173	96.43	8.793	296.8
Humana	24.43	38.01	45.11	1989
Kaiser	20.24	99.96	30.69	97.61
United	13.02	50.63	28.13	1024
Other Insurers	38.49	100	44.08	1770
Total	101.7	69.44	100	2412

Table II: Placebo - "Balance" Test

	Est.	Standard Error	P-Value	Mean
Avg. Premium	-0.0138	0.0246	0.5753	48.23
HCC Score	-0.0013	0.0054	0.8147	95.18
% White	-0.6200	0.8400	0.4566	85.27
% Black	0.0055	0.0075	0.4650	10.73
% Hispanic	-0.0104	0.0064	0.1070	4.777
% Asian	-0.0010	0.0006	0.1228	0.887
Avg Income	-3.494	4.394	0.4283	31,000
% Medicare Eligible	0.0025	0.0029	0.3901	16.53
Population	-17.90	69.76	0.7982	65,500
Brands Present	0.0013	0.0008	0.1088	3.18

Regressions reflect a regression with the listed variable as the dependent variable and total GRP as the independent variable.

Table III: Selection into the Border Sample

	Est.	Standard Error	P-Value	Mean
Avg. Premium	1.083	1.194	0.3643	47.43
HCC Score	0.6400	0.3500	0.0667	94.75
% White	-0.6100	0.6900	0.3789	85.68
% Black	1.630*	0.6700	0.0152	9.620
% Hispanic	-3.350***	0.4000	< 0.001	7.060
% Asian	-0.4800***	0.0800	< 0.001	1.220
Avg Income	-2,960***	296.5	< 0.001	33,000
% Medicare Eligible	0.600***	0.1600	< 0.001	16.12
Population	-68,800**	10,600	< 0.001	112,000
Brands Present	-0.0579	0.0455	0.2033	3.2174

Regressions reflect a regression with the listed variable as the dependent variable and an indicator for whether the county is in the border sample is the main independent variable.

Table IV: Category Expansion (MA%)

	(1)	(2)	(3)	(4)
TotalGRP	0.0756***	0.0601***	-0.002	0.0021
	(0.0143)	(0.0181)	(0.0061)	(0.0030)
Controls		x	x	x
County FEs			x	x
Border Approach				x
ROI CI	[225%, 606%]	[68.6%, 551%]	[-195%, -32.6%]	[-126%,-45.5%]
Mean Enrollment %	12.45	13.46	13.47	11.40
R-squared	0.081	0.281	0.955	0.967
Observations	17728	16100	16075	7665

*** $p<0.001$, ** $p<0.01$, * $p<0.05$

DMA clustered standard errors in parentheses.

Table V: Brand Share (%)

	(1)	(2)	(3)	(4)
GRP	0.3333***	0.3597***	0.0339	-0.0025
	(0.0453)	(0.0538)	(0.0353)	(0.0236)
RivalGRP	-0.1517***	-0.1583***	-0.0275	0.0030
	(0.0182)	(0.0237)	(0.0189)	(0.0143)
Controls		x	x	x
County-Brand FEs			x	x
Border Approach				x
ROI CI	[75.6%, 203%]	[82.6%,234%]	[-125%, -26.0%]	[-135%,-68.5%]
Mean Brand %	39.289	39.2337	39.5131	42.7229
R-squared	0.085	0.247	0.790	0.858
Observations	37748	37519	36924	15550

*** $p<0.001$, ** $p<0.01$, * $p<0.05$

Brand-DMA clustered standard errors in parentheses.

Table VI: United Share (%)

	(1)	(2)	(3)	(4)
RivalGRP	-0.1308***	-0.0862**	0.0496	-0.0088
	(0.0370)	(0.0273)	(0.0334)	(0.0304)
Controls		x	x	x
County FEs			x	x
Border Approach				x
ROI CI	[-58.1%, -46.0%]	[-76.5%,0.404%]	[-183%,-88.6%]	[-136.4%,-51.0%]
Mean United %	25.66	25.59	25.90	23.69
R-squared	0.0510	0.217	0.947	0.928
Observations	1926	1913	1451	397

*** $p<0.001$, ** $p<0.01$, * $p<0.05$

DMA clustered standard errors in parentheses.

Table VII: Category Expansion - Stock (MA %)

	(1) δ=0.6	(2) δ=0.8	(3) δ=1	(4) Non-parametric
$GRPStock$	-0.0021	-0.0012	-0.0005	
	(0.0012)	(0.0006)	(0.0003)	
GRP				0.0021
				(0.0030)
GRP_{t-1}				-0.0002
				(0.0013)
R-squared	0.993	0.993	0.993	0.967
Observations	2581	2581	2581	7665

*** p<0.001, ** p<0.01, * p<0.05

DMA clustered standard errors in parentheses. All specifications use the border approach with a full set of control variables and county fixed-effects.

Table VIII: Brand Share - Stock (%)

	(1) δ=0.6	(2) δ=0.8	(3) δ=1	(4) Non-parametric
$GRPStock$	-0.0642	-0.0475	-0.0188	
	(0.0333)	(0.0278)	(0.0181)	
GRP				-0.0139
				(0.0223)
GRP_{t-1}				0.0456*
				(0.0186)
R-squared	0.938	0.938	0.938	0.858
Observations	5417	5417	5417	15550

*** p<0.001, ** p<0.01, * p<0.05

Brand-DMA clustered standard errors in parentheses. All specifications use the border approach with a full set of control variables and county fixed-effects. All specifications include rival GRP stock.

Table IX: Heterogeneity and Targeting

	(1) MA %	(2) Brand %
GRP	0.0050	-0.0055
	(0.0075)	(0.0583)
xGRP	0.0000	0.0002
	(0.0001)	(0.0004)
x$Premium$	-0.0023	0.0068
	(0.0022)	(0.0185)
x$Competition$	-0.0022	0.0583*
	(0.0022)	(0.0257)
x$Risk$	-0.0001	0.0064
	(0.0015)	(0.0161)
x$OpenEnrollment$	-0.0193	-0.0527
	(0.0071)	(0.0614)
x$Elderly$	0.0049	0.0326
	(0.0035)	(0.0197)
R-squared	0.967	0.858
Observations	7654	15529

*** $p<0.001$, ** $p<0.01$, * $p<0.05$

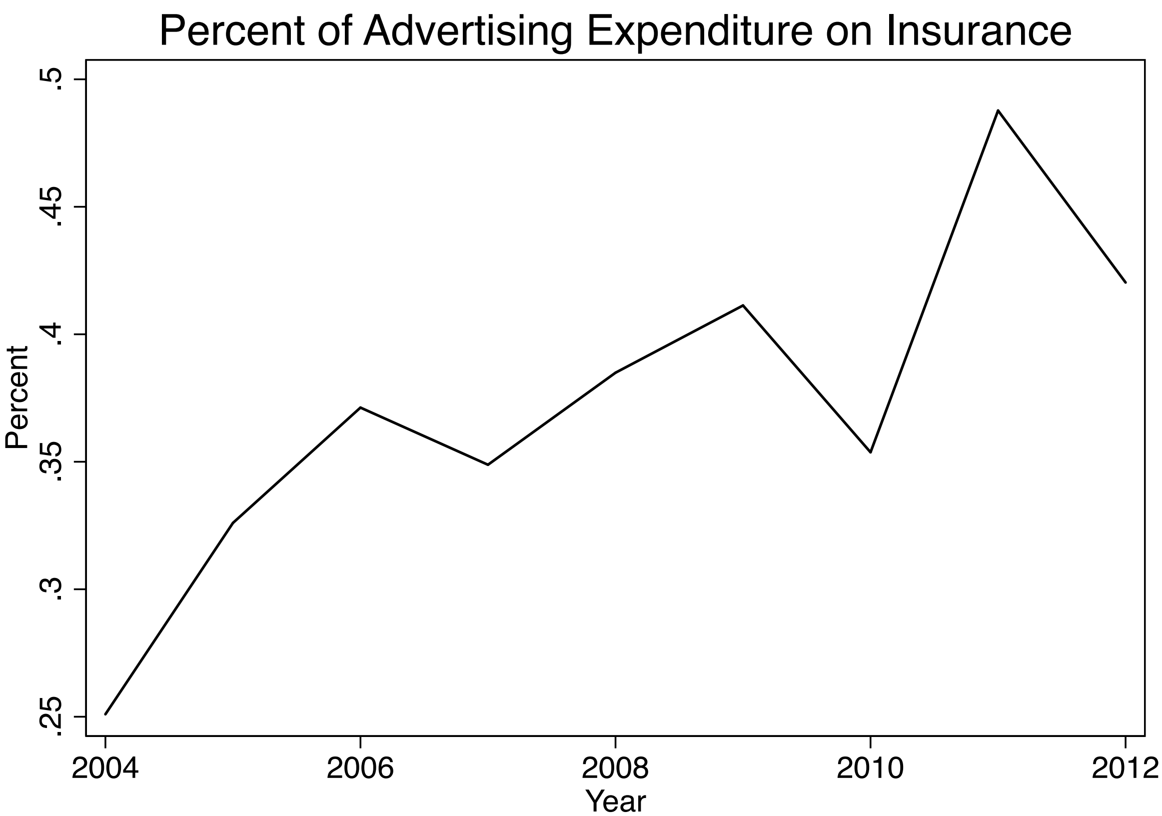
Brand-DMA clustered standard errors in parentheses. All specifications use the border approach with a full set of control variables and county fixed-effects. All specifications include interactions with demographics, none of which come up significant and are suppressed for readability.

Table X: Targeting By Competitive Structure

Brands Present	1	2	3	4	5	6	≥ 7
Across Brands	17.57	18.09	17.68	18.51	17.11	15.13	14.39
	(14.80)	(18.61)	(18.22)	(19.66)	(17.55)	(14.21)	(14.19)
Aetna	2.151	2.403	2.413	2.162	2.738	3.712	2.636
	(4.300)	(3.752)	(3.743)	(3.443)	(3.651)	(4.941)	(3.562)
BCBS	14.21	16.82	19.18	23.45	21.14	16.16	13.01
	(10.71)	(16.25)	(17.64)	(22.24)	(15.92)	(11.62)	(7.707)
Cigna	2.589	2.256	2.066	2.061	2.107	2.277	2.082
	(3.556)	(3.443)	(3.320)	(3.333)	(3.309)	(3.413)	(3.167)
Humana	19.18	23.23	22.79	25.32	24.20	19.60	17.02
	(8.587)	(12.18)	(11.86)	(12.40)	(11.84)	(9.375)	(8.480)
Kaiser		8.560	20.98	15.76	21.53	19.92	20.58
		(12.25)	(12.17)	(12.50)	(12.37)	(11.34)	(12.31)
United	13.91	14.30	14.71	15.30	15.53	14.84	14.04
	(13.21)	(13.61)	(13.81)	(14.38)	(14.02)	(13.07)	(11.97)
Other	29.78	37.22	36.37	37.06	35.46	30.95	31.33
	(17.21)	(21.12)	(20.20)	(21.59)	(20.06)	(15.63)	(18.41)
Number of County Years	992	2,380	6,101	4,749	1,680	355	151

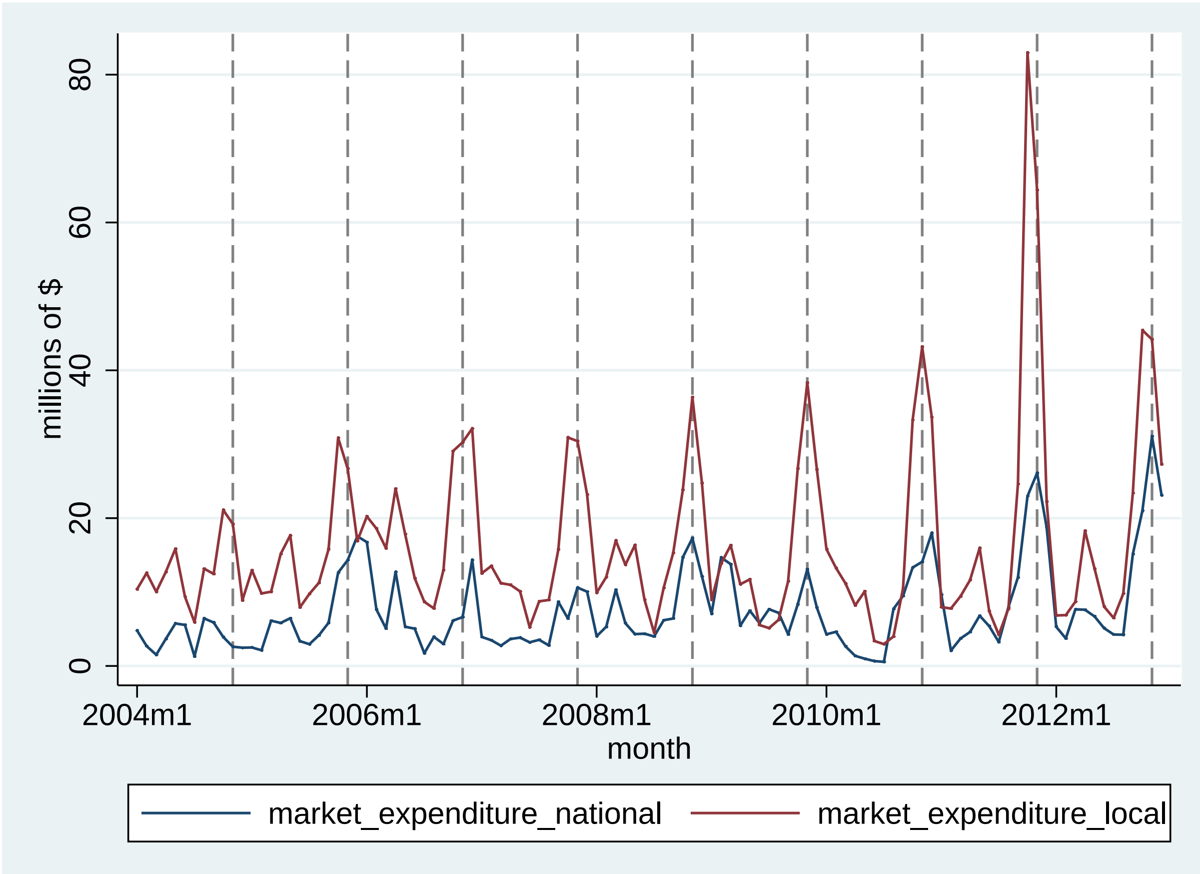
Mean Brand GRP presented with standard deviations below in parentheses. All non-national brands are grouped together as one "other" brand, so ≥ 7 indicates that all national brands plus at least one regional brand is present in the market

Figure I: Health Insurance Advertising by Year



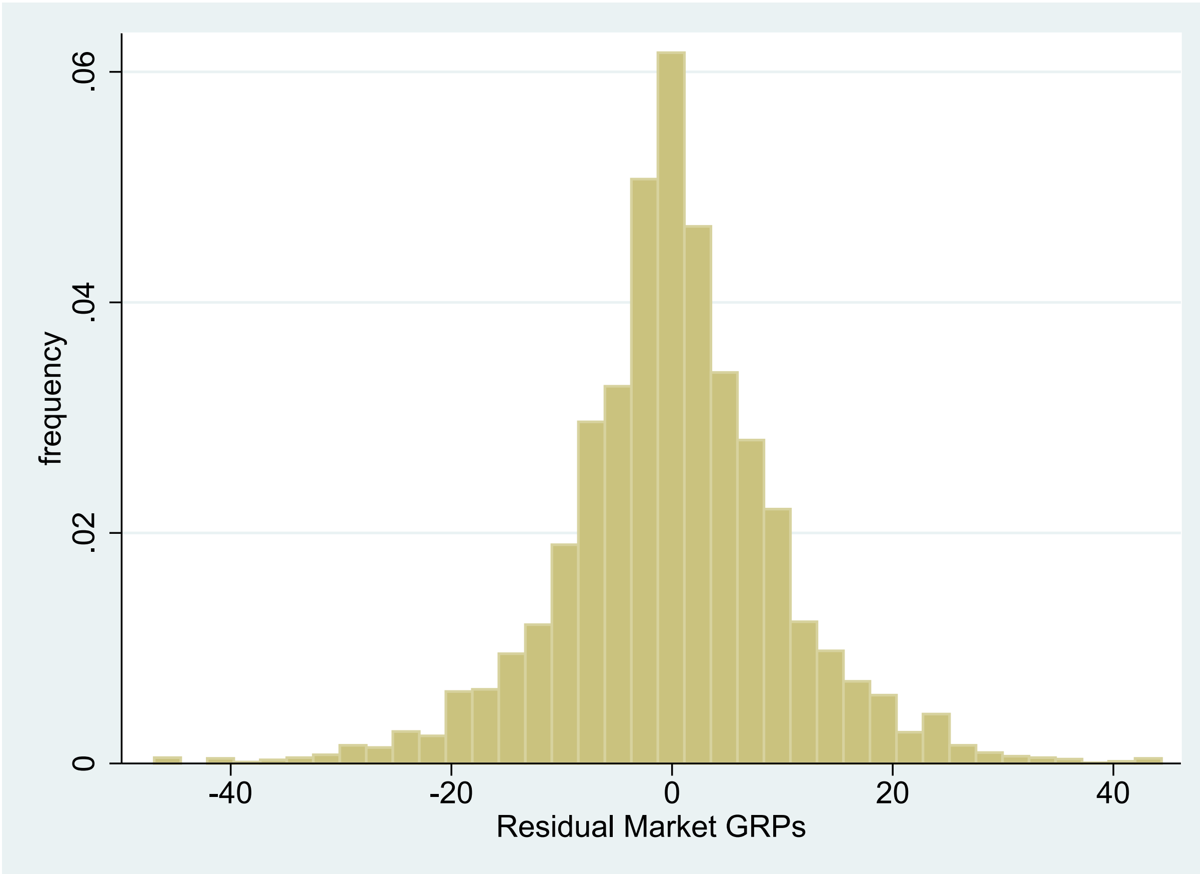
Note: Figure shows spending on health insurance television advertising as a percentage of total television advertising by year. Total television advertising is approximately $100 billion. The source is the AC Nielsen's Media Database.

Figure II: Health Insurance by Calendar Month



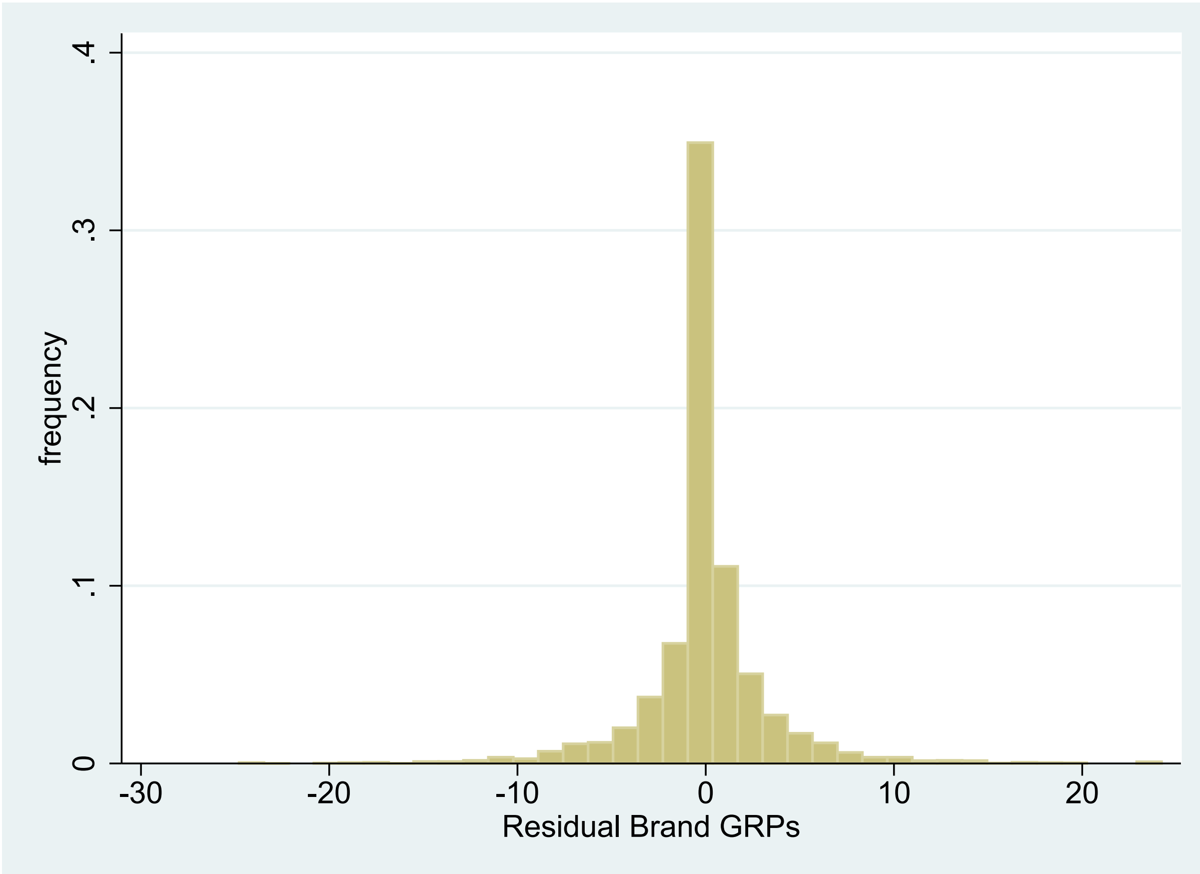
Note: Figure shows that significant amounts of advertising is local and concentrated in the open enrollment period from October 15-December 7. The source is the AC Nielsen's Media Database.

Figure III: Variation in Total GRP Changes across DMA Borders



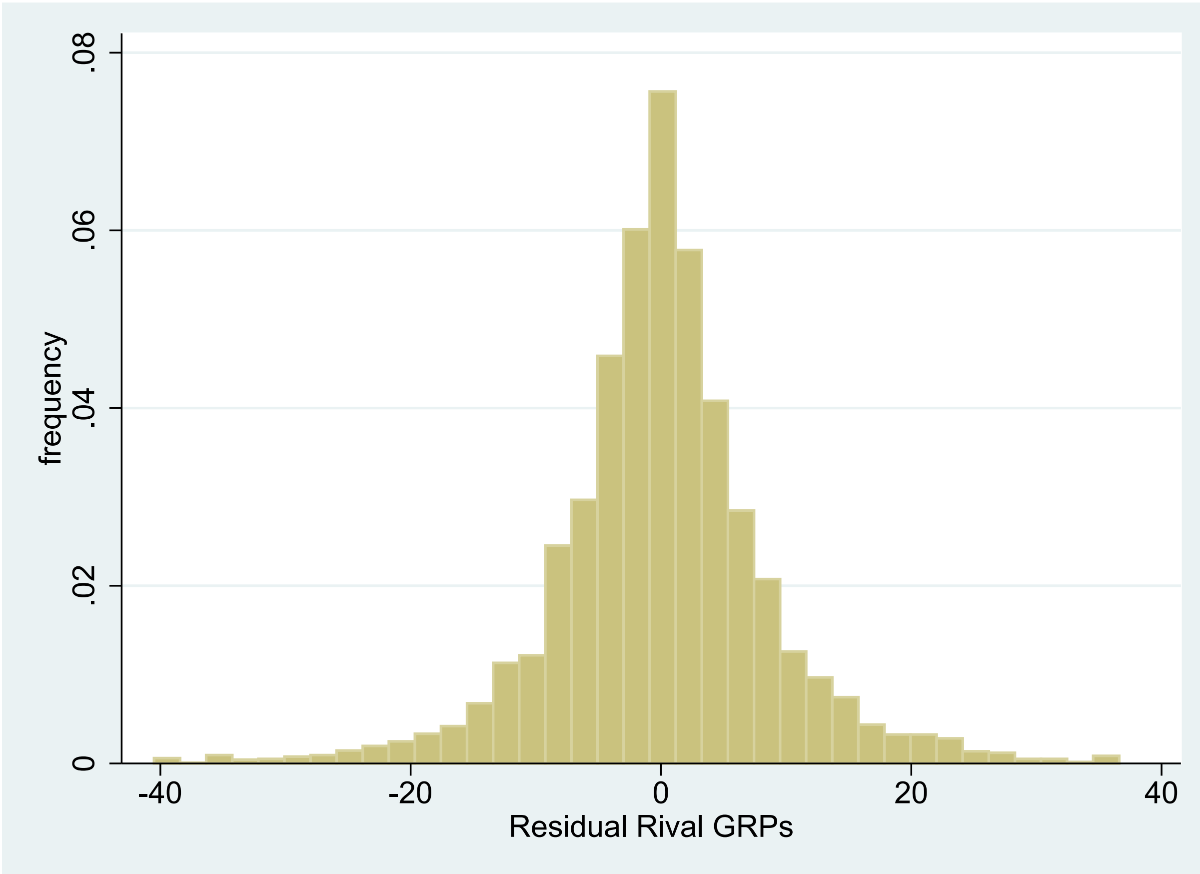
Note: Figure shows that there is significant variation in advertising across DMA borders, conditional on county and border-year fixed effects.

Figure IV: Variation in Brand GRP Changes across DMA Borders



Note: Figure shows that there is significant variation in brand advertising across DMA borders, conditional on brand-county and brand-border-year fixed effects.

Figure V: Variation in Rival GRP Changes across DMA Borders



Note: Figure shows that there is significant variation in rival advertising across DMA borders, conditional on brand-county and brand-border-year fixed.

Figure VI: United Advertising Expenditure Over Time

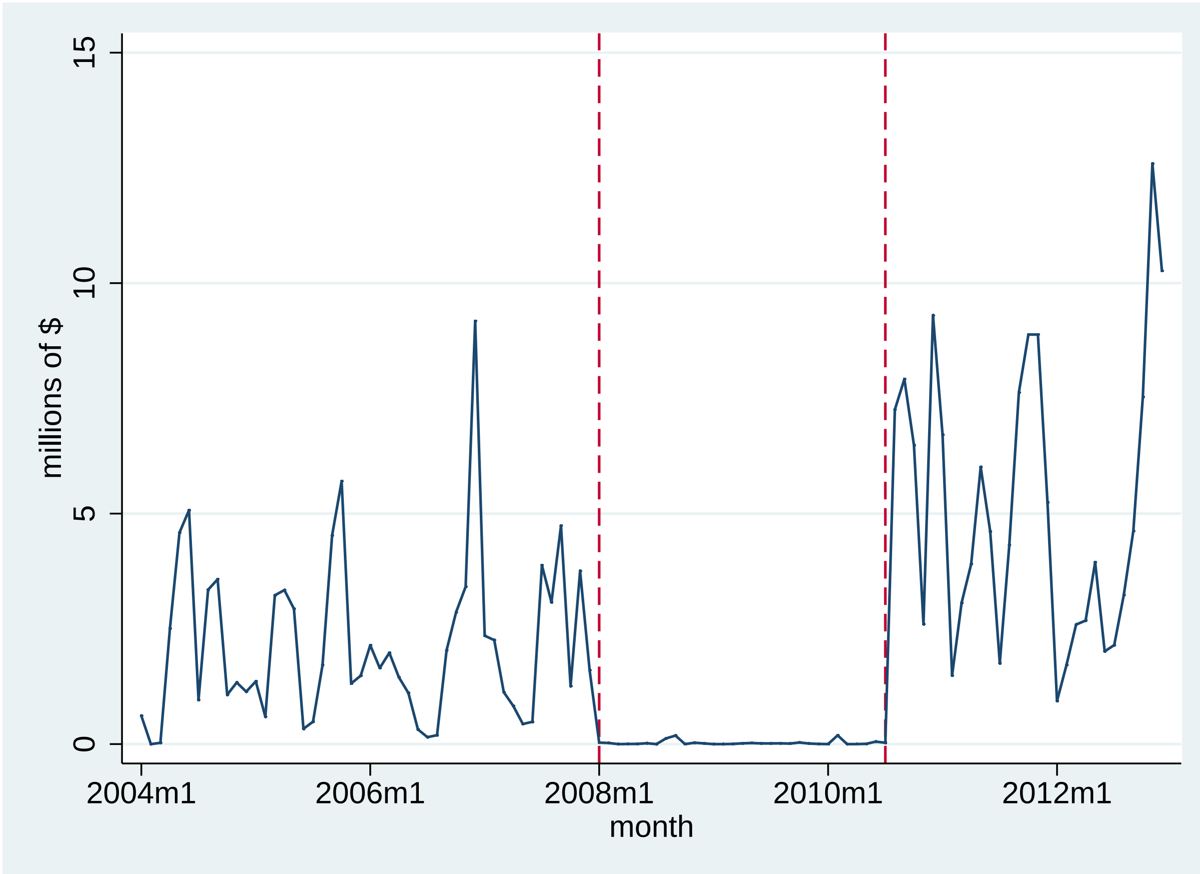
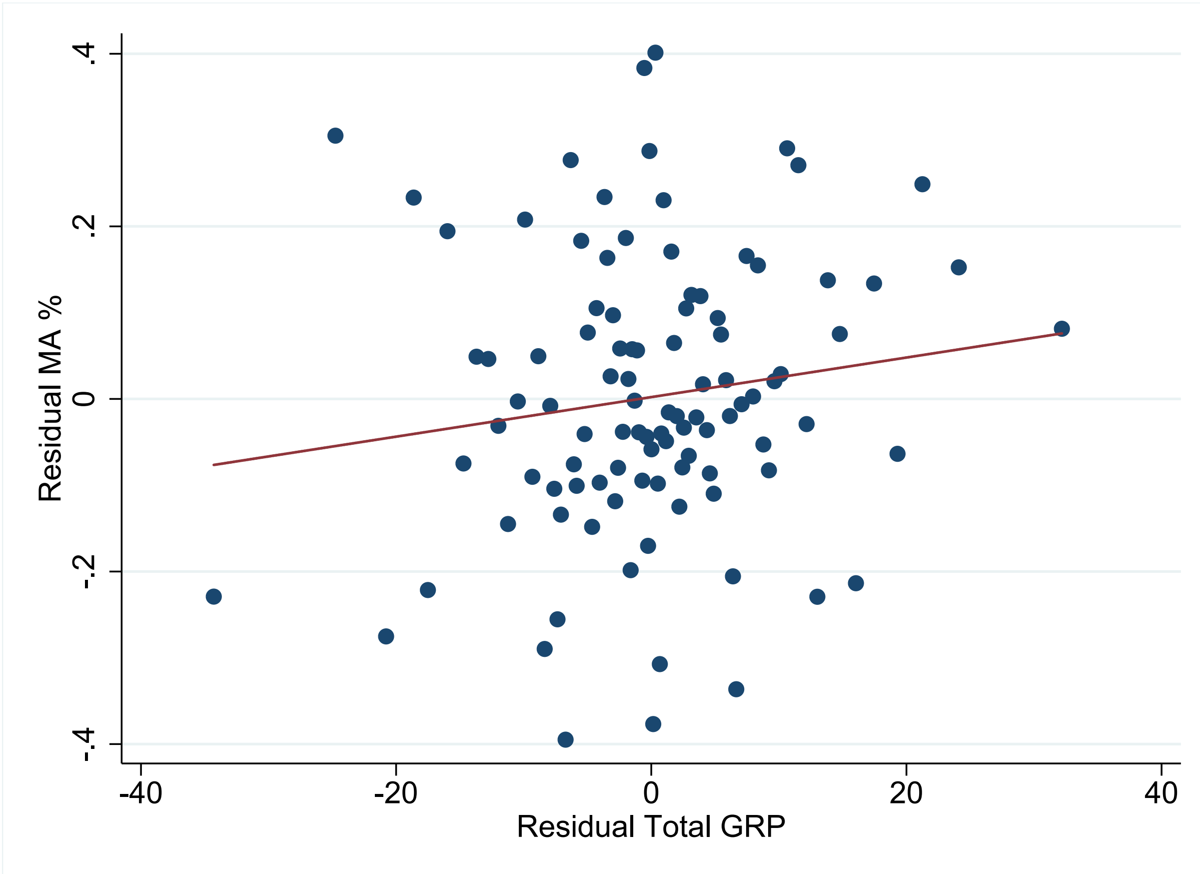
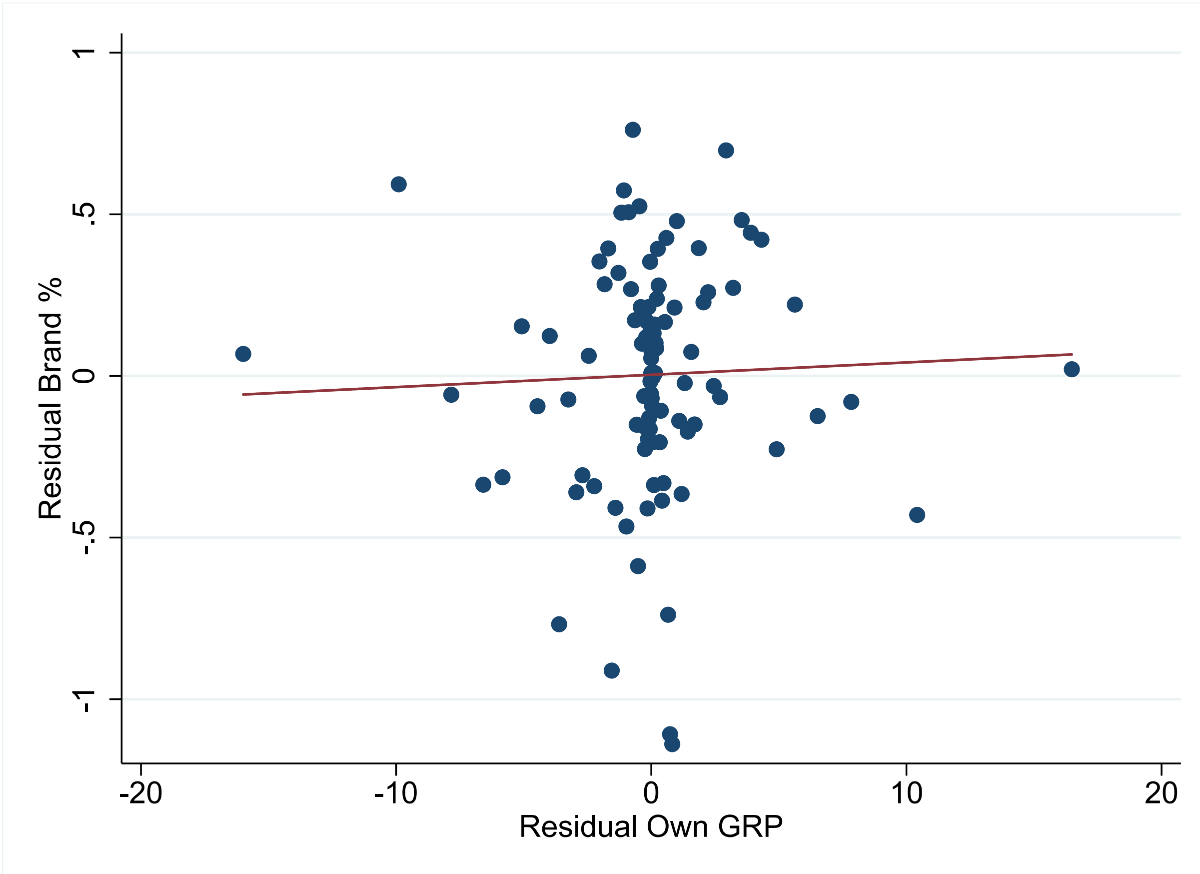


Figure VII: Category Bin Scatter



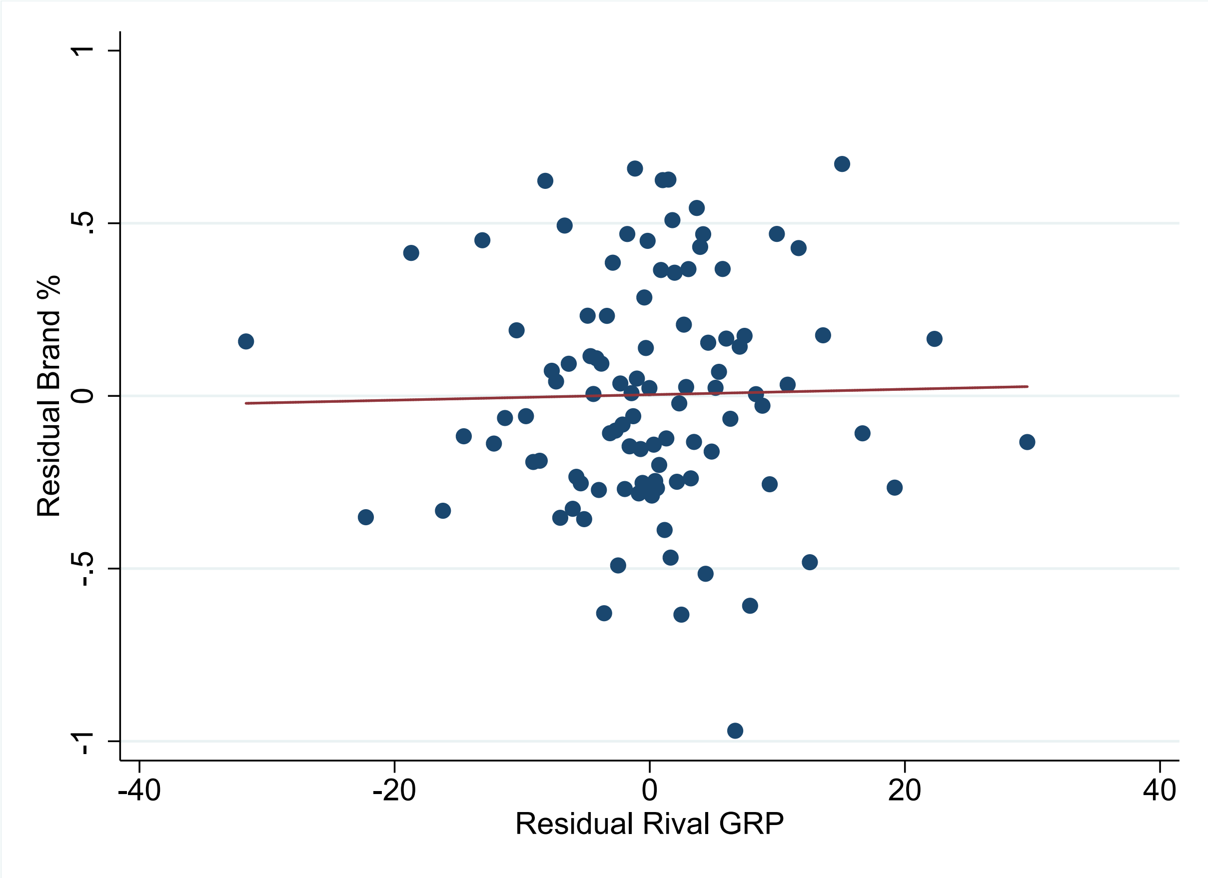
Note: This bin scatter has 100 bins and reflects MA % on the vertical axis and Category GRP on the horizontal axis, net of county and border-year fixed effects.

Figure VIII: Brand Bin Scatter - Own GRP



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

Figure IX: Brand Bin Scatter - Rival GRP

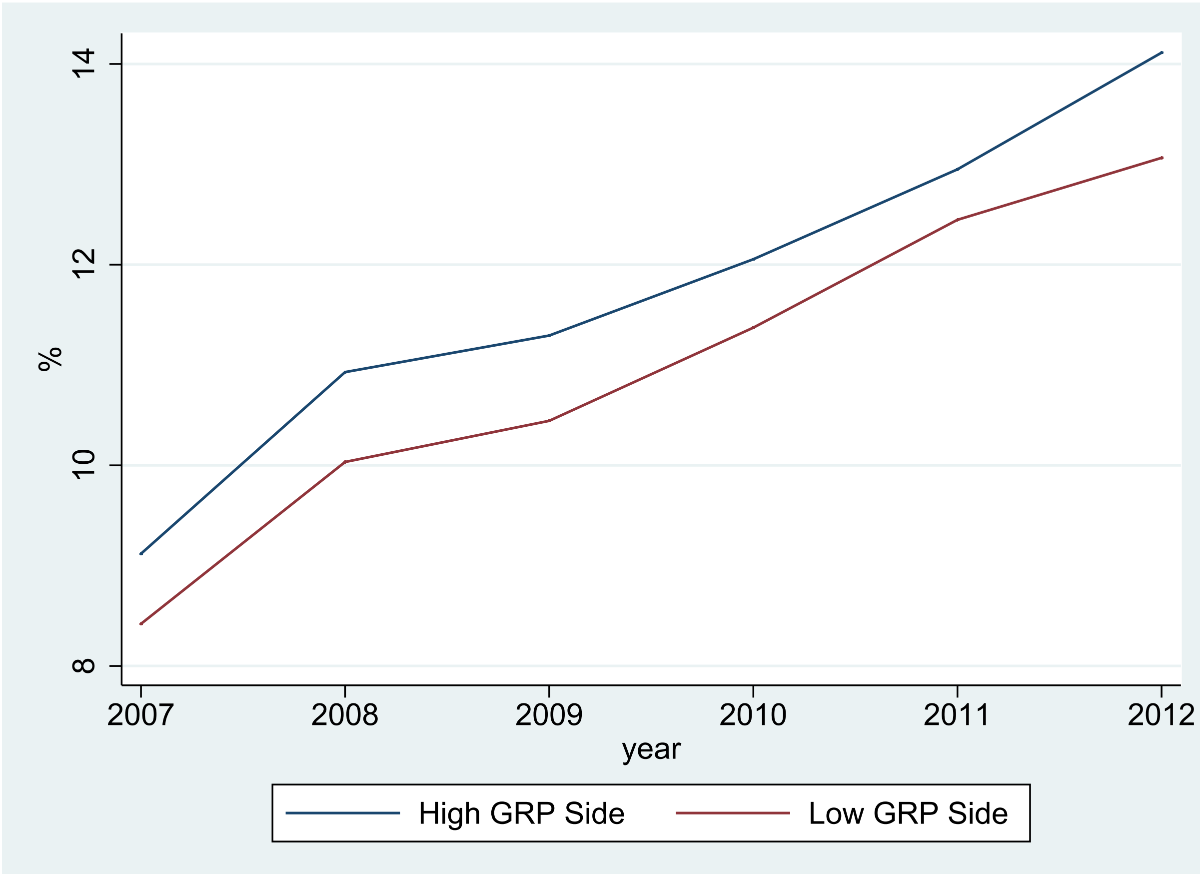


Note: This bin scatter has 100 bins and reflects brand % on the vertical axis and rival GRP on the horizontal axis, net of brand-county and brand-border-year fixed effects.

Figure X: United Shares Around Cessation

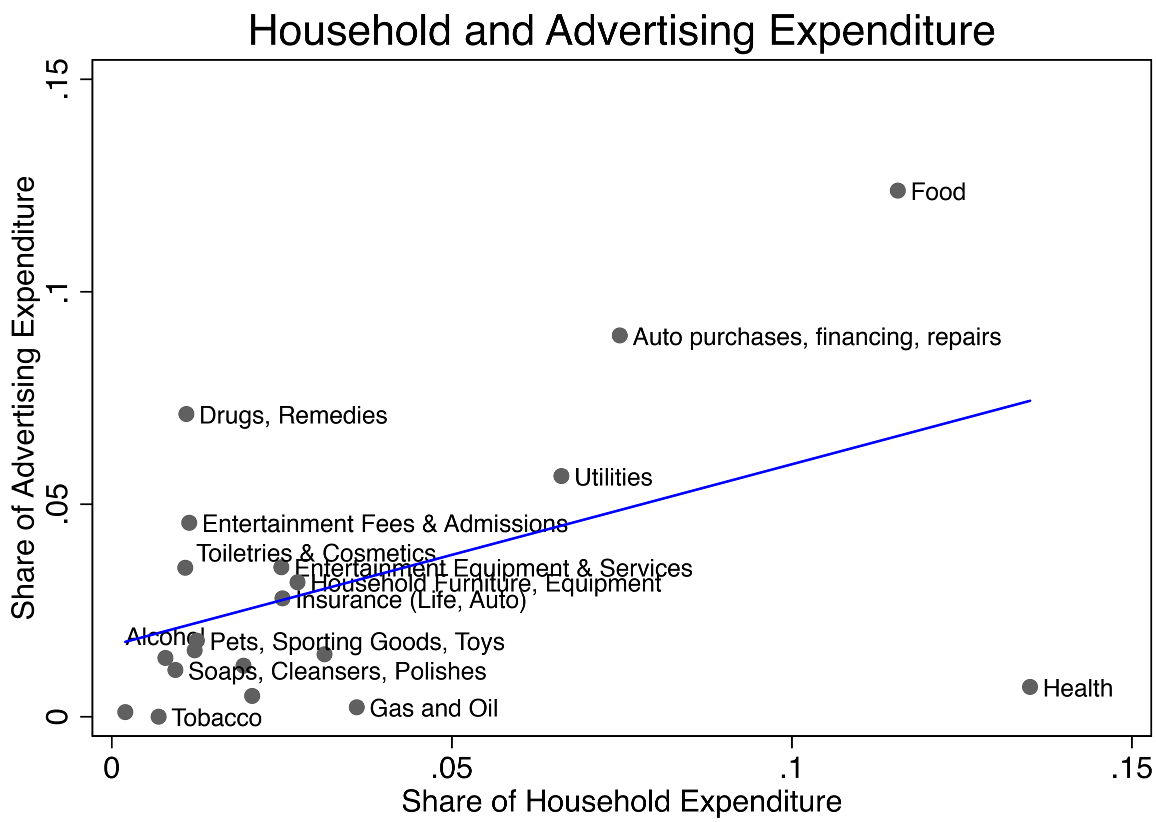
Note: This shows the average United brand share over time. Vertical dashed lines represent the years with advertising turned off.

Figure XI: MA % Over Time on High vs. Low GRP Side of Border



Note: This shows the average MA % over time separately on the persistently high versus the persistently low advertising sides of the TV market boundary.

Figure XII: Health Insurance Advertising in Context



Appendix A - Sensitiivity to Border Size

In the main analysis, the sample is restricted to those borders which make up no more than 35% of the DMA as a whole. In this section, I test sensitivity to this cutoff in both the category expansion and the business stealing analysis. Results are in Table A.1 and Table A.2. 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. Positive ROI from both category expansion and business stealing remain outside of the 95% confidence interval.

Table A.1: Category Expansion (MA%) - Sensitivity to Border Size

	(1) 10%	(2) 20%	(3) 30%	(4) 40%	(5) 50%
TotalGRP	-0.0075	-0.0022	0.0012	0.0017	0.0004
	(0.0061)	(0.0034)	(0.0030)	(0.0029)	(0.0029)
ROI CI	[-233%, -69.8%]	[-161%, -68.7%]	[-132%, -51.5%]	[-128%,-49.5%]	[-137%, -57.9%]
Mean Enrollment %	8.67	9.94	10.90	11.40	11.80
R-squared	0.944	0.957	0.963	0.968	0.971
Observations	1257	4283	7030	8202	9076

*** $p<0.001$, ** $p<0.01$, * $p<0.05$

DMA clustered standard errors in parentheses. All specifications use the border approach with controls.

Table A.2: Brand Share (%) - Sensitivity to Border Size

	(1) 10%	(2) 20%	(3) 30%	(4) 40%	(5) 50%
GRP	-0.0967	-0.0481	-0.0042	0.0061	0.0181
	(0.0632)	(0.0337)	(0.0242)	(0.0242)	(0.0226)
RivalGRP	0.0926*	0.0218	0.0028	-0.0005	-0.0014
	(0.0442)	(0.0192)	(0.0151)	(0.0139)	(0.0135)
ROI CI	[-258%,-80.5%]	[-182%, -87.1%]	[-137%, -68.9%]	[-130%, -61.6%]	[-119%, -55.1%]
Mean Brand %	48.4	44.8	43.1	42.7	42.1
R-squared	0.832	0.856	0.853	0.855	0.855
Observations	2073	8122	14129	16702	18830

*** $p<0.001$, ** $p<0.01$, * $p<0.05$

DMA clustered standard errors in parentheses. All specifications use the border approach with controls.

Appendix B - Using Log Enrollments Rather than MA% and Brand %

In the main analysis, the dependent variable in the category expansion analysis is the percentage of seniors who choose MA over TM, while in the brand share analysis, the dependent variable throughout is the percentage of MA enrollees who choose that particular brand. In this section, the dependent variable is changed to the natural log of total MA enrollments and the natural log of brand enrollments, respectively. The analysis of main effects is provided in Table B.1 and Table B.2. All results are consistent with the main analysis.

Table B.1: Category Expansion (Log MA Enrollment)

	(1)	(2)	(3)	(4)
TotalGRP	0.0094***	0.0073***	0.0001	0.0007
	(0.0024)	(0.0024)	(0.0005)	(0.0004)
Controls		x	x	x
County FEs			x	x
Border Approach				x
Mean Log Enrollment	6.3344	6.4398	6.4464	5.9783
R-squared	0.043	0.540	0.981	0.980
Observations	16726	15979	15950	7593

*** p<0.001, ** p<0.01, * p<0.05

DMA clustered standard errors in parentheses.

Table B.2: Brand Share (Log Enrollments)

	(1)	(2)	(3)	(4)
GRP	0.0093***	0.0169***	0.0029*	0.0013
	(0.0028)	(0.0036)	(0.0011)	(0.0008)
RivalGRP	0.0025	-0.0010	-0.0005	0.0012*
	(0.0014)	(0.0011)	(0.0008)	(0.0005)
Controls		x	x	x
County-Brand FEs			x	x
Border Approach				x
Mean Brand Log Enrollments	5.634	5.6393	5.6646	5.201
R-squared	0.015	0.400	0.932	0.951
Observations	37111	36887	36279	15236

*** p<0.001, ** p<0.01, * p<0.05

Brand-DMA clustered standard errors in parentheses.

Appendix C - Using Occurrences Instead of GRP

As noted in the data section, the main analysis focuses on gross rating points (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 mechancially 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 are not in this paper but available from the author upon 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 thirty seconds of ad air time is coded as one occurrence, regardless of how many people see it. As such, an ad on a midnight re-run 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 and Table C.2 present 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 the main analysis. Results are very similar to the main analysis, though with more noise. Without using the border strategy, advertising appears useful in lifting MA over TM. When the border strategy is used, the average lift from category expansion goes to zero. While the 95% confidence interval is almost entirely negative, the right edge does just cross zero. Results are similar in the brand share analysis, but with considerably more noise in all specifications other than the border strategy. The border strategy provides results consistent with the main analysis, and while the ROI confidence interval is wider, it continues to exclude positive ROI.

Table C.1: Category Expansion (MA %) - Using Occurrences

	(1)	(2)	(3)	(4)
Total Ads	0.0811***	0.2722**	0.0851*	0.0253
	(0.0089)	(0.1032)	(0.0303)	(0.0150)
Controls		x	x	x
County FEs			x	x
Border Approach				x
ROI CI	[20.2%, 85.7%]	[32.0%, 794%]	[-51.5%, 172%]	[-107.9%, 3.20%]
Mean Enrollment %	12.447	13.457	13.4747	11.4027
R-squared	0.023	0.259	0.956	0.967
Observations	17728	16100	16075	7665

*** $p<0.001$, ** $p<0.01$, * $p<0.05$

DMA clustered standard errors in parentheses. Ads are measured in 1000s of 30 second spots.

Table C.2: Brand Share (%) - Using Occurrences

	(1)	(2)	(3)	(4)
Ads	0.1343***	1.4616***	0.0619	-0.0662
	(0.0300)	(0.3014)	(0.1428)	(0.1441)
Rival Ads	-0.1803***	-0.8459***	-0.2526***	0.0187
	(0.0202)	(0.1298)	(0.1059)	(0.0814)
Controls		x	x	x
County-Brand FEs			x	x
Border Approach				x
ROI CI	[-85.0%, -61.6%]	[72.9%,308%]	[-143%, -32.1%]	[-169%,-57.0%]
Mean Brand %	37.19	37.09	37.40	40.62
R-squared	0.053	0.224	0.819	0.854
Observations	20209	20112	19811	8926

*** $p<0.001$, ** $p<0.01$, * $p<0.05$

Brand-DMA clustered standard errors in parentheses. Ads are measured in 1000s of 30 second spots.