NBER WORKING PAPER SERIES

HOW AND WHEN TO USE THE POLITICAL CYCLE TO IDENTIFY ADVERTISING EFFECTS

Sarah Moshary
Bradley T. Shapiro
Jihong Song

Working Paper 27349
http://www.nber.org/papers/w27349

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
June 2020

The authors would like to thank Sylvia Hristakeva, Mitch Lovett, Sanjog Misra and participants at the 2019 Marketing Science Conference and the 12th Workshop on the Economics of Advertising and Marketing for valuable feedback. Results calculated (or derived) based on data from The Nielsen Company (US), LLC and marketing databases provided by the Kilts Center for Marketing Data Center at The University of Chicago Booth School of Business. The conclusions drawn from the Nielsen data are those of the researchers and do not reflect the views of Nielsen. Nielsen is not responsible for, had no role in, and was not involved in analyzing and preparing the results reported herein. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

How and When to Use the Political Cycle to Identify Advertising Effects
Sarah Moshary, Bradley T. Shapiro, and Jihong Song
NBER Working Paper No. 27349
June 2020
JEL No. C01,C26,C36,L0,L62,M3,M37

ABSTRACT

A central challenge in estimating the causal effect of TV advertising on demand is isolating quasi-random variation in advertising. Political advertising, which topped $14 billion in expenditures in 2016, has been proposed as a plausible source of such variation and thus a candidate for an instrumental variable. We provide a critical evaluation of how and where this instrument is valid and useful across categories. We characterize the conditions under which political cycles theoretically identify the causal effect of TV advertising on demand, highlight threats to the exclusion restriction and monotonicity condition, and suggest a specification to address the most serious concerns. We test the strength of the first stage category-by-category for 274 product categories. For most categories, weak-instrument robust inference is recommended, as first-stage F-statistics are less than 10 for 221 of 274 product categories in our benchmark specification. The largest first-stage F-statistics occur in categories that typically advertise locally, such as automobile dealerships and restaurants. Failure to use the suggested specification leads to results that suggest violations of exclusion and monotonicity in a significant number of categories. Finally, we conduct a case study of the auto industry. Despite a very strong first stage, the IV estimate for this category is imprecise.

Sarah Moshary
University of Chicago
Booth School of Business
5807 S Woodlawn Ave
Chicago, IL 60637
Sarah.moshary@chicagobooth.edu

Bradley T. Shapiro
Booth School of Business
University of Chicago
5807 South Woodlawn Avenue
Chicago, IL 60637
and NBER
Bradley.Shapiro@chicagobooth.edu

Jihong Song
Princeton University
jihongs@princeton.edu

A data appendix is available at http://www.nber.org/data-appendix/w27349

1 Introduction

In 2016, political groups in the United States spent in excess of $4 billion on television advertising. By increasing the price of advertising time, this large disruption in the TV advertising market may generate exogenous variation in non-political commercial advertising.[1] Using this variation in an instrumental variables (IV) framework was first suggested by Sinkinson and Starc [2018], which estimates the effect of advertising for cholesterol-lowering drugs. In this paper, we provide a critical evaluation of that approach using data across all product categories that advertise on television in the United States. Our intention is to guide practitioners on how and when to use political advertising as an instrument. We also seek to provide researchers a roadmap for scrutinizing IV strategies that are meant to have broad applicability by establishing the properties and regularities of this candidate instrument. This study may be read as a cautionary one. Our findings suggest that despite the intuitive appeal of the political advertising IV strategy, the plausibility of the exclusion restriction is sensitive to the specification, and the first stage is often weak. These findings highlight pitfalls that may plague other IVs as well.

We begin by characterizing the conditions under which the political advertising IV strategy is valid. This first step is crucial for any IV strategy because the exclusion restriction and monotonicity condition are not directly testable. Thus, a sound theoretical argument is needed to pinpoint a specification that satisfies both conditions. In this case, the justification of the exclusion restriction rests on the premise that political advertising reduces commercial advertising by increasing the price of commercial advertising. Operationally, we argue that this requires the inclusion of television market fixed-effects and time fixed-effects at the periodicity of the data in order to isolate the price mechanism and prevent contamination from other sources that might violate exclusion or monotonicity. These arguments apply to any past or future election where the television advertising market continues to operate as it does in our data.

Second, we characterize where the IV strategy has a sufficiently strong first stage to limit weak instrument bias. We use data on political advertising and commercial advertising for 274 product categories from 2010-2016. This data covers two US Presidential elections (2012 and 2016) and several midterm, state, and local elections. Overall, we find that relatively few categories have a

[1] https://adage.com/article/media/2016-political-broadcast-tv-spend-20-cable-52/307346/

sufficiently strong first stage to alleviate concerns about weak instrument bias. For example, in a simple log-linear specification, we find that political advertising has a first-stage F-statistic less than 10 in 221 of 274 product categories and an F-statistic greater than 25 in only 28 categories.[2] While the exact number changes across specifications, the qualitative result is the same: most categories exhibit a weak first stage.[3] Overall, political advertising has a very concentrated effect on a few categories that advertise almost exclusively locally; for example, advertising by car dealerships, hospitals, household furnishings outlets, and appliance stores is strongly offset by political advertising, while advertising for national brands in consumer packaged goods is not.[4]

While the exact quantitative results on category-by-category first stage F-statistics do not necessarily generalize to political advertising associated with all future elections, they are useful for two reasons: first, they guide practitioners who hope to use variation in political advertising between 2010 and 2016 as an instrument for commercial advertising; and second, they guide our prior for how much and for which types of product categories a given level of political advertising is likely to disrupt the commercial advertising market in future elections.

Finally, we fully implement the strategy for one product category, automobiles, where both (1) the first stage is particularly strong, and (2) the short-run effect of advertising on sales is expected to be zero. The purpose of this exercise is twofold. First, we want to characterize the properties of the IV estimator in an instance where the F-statistic exceeds conventional thresholds that limit weak instrument bias. Second, because we believe that the true advertising effect is zero, the IV estimate provides a placebo test of the exclusion restriction. We find that the IV strategy produces a confidence interval that contains the presumed truth (zero), but the confidence interval is wide, containing improbably large negative and positive values. Thus, this case study underscores that political advertising may not identify precise advertising elasticities, even for the 28 categories with a strong first stage. For these categories, our findings suggest that reducing noise in sales is of

[2]We specify advertising as a log because researchers typically want to identify an advertising elasticity, where the independent variable of interest is the log of advertising and the dependent variable of interest is the log of quantity.

[3]This includes using machine learning as in Belloni, Chen, Chernozhukov, and Hansen [2012] to find the functional form that gets the strongest possible first stage.

[4]The distribution of first-stage F-statistics that we recover raises a question of inference with weak instruments. The F-statistic for many categories falls near the critical value of 10, proposed by Stock and Yogo [2005], although more recent work highlights that this value may be too permissive (for example, Olea and Pflueger [2013]). We suggest that researchers justify their choice of critical values on a case-by-case basis, depending on the research question. If the first-stage relationship appears weak, as measured by the F-statistic, rather than disregard the political advertising instrument, researchers should consider using weak identification robust inference methods (e.g., Andrews [2016], Anderson and Rubin [1949]).

first-order importance to obtaining a precise estimate.[5]

To demonstrate the strength of the first stage, we focus on estimating category-level advertising effects because political advertising constitutes only one source of variation. As a result, it may be used to estimate a parameter on a single endogenous variable. In oligopoly settings, own and rival advertising are potentially confounded by competitive responses, so that estimating brand-level effects using political advertising as an instrument requires either an additional source of variation or a particular functional form assumption.[6,7] As an example, in section 2.6, we describe how to use political advertising in conjunction with a logit model of demand to estimate brand-level advertising effects.

Category-level advertising effects are also interesting and important in their own right in many circumstances. For many research questions (e.g. Shapiro [2018], Sinkinson and Starc [2018]), separating category expansion from the business stealing effects of advertising is important and requires an instrument like political advertising that can identify category-expansive advertising effects. Second, several policy debates revolve around the category-level effect of advertising, such as bans on television advertising for junk food and smoking, as in Dubois et al. [2018] and Tuchman [2019].

This paper contributes to the literature that seeks to identify the effects of television advertising using observational data (including Shapiro [2018], Hartmann and Klapper [2018], Sinkinson and Starc [2018], and Thomas [2017], among others). Shapiro [2018] exploits the borders of television markets to generate quasi-random variation in advertising. Hartmann and Klapper [2018] uses the ex ante uncertain identity of the teams in the Super Bowl, which provides randomness in which households watch television during the game. Thomas [2017] combines data on how TV advertising bundles viewers of a single show with information on ideal targeting to construct an instrument for advertising. Li et al. [2019] compare IV and non-IV approaches in estimating the effect of political advertising on votes. Sinkinson and Starc [2018] pioneered the approach in this paper, using political

[5] Our results in this case study also underscore that the problem of weak instruments is one of bias rather than precision. A strong instrument circumvents bias due to weak instruments, but it need not increase precision.

[6] For example, researchers may interact political advertising with brand fixed effects. This would require an assumption that the amount of displacement on one brand is independent of the amount of displacement on a rival brand. This might be reasonable in some cases, but it is not an assumption we are willing to apply broadly across 274 categories for the purposes of this demonstration.

[7] For example, Sinkinson and Starc [2018] use political advertising in conjunction with a temporary ban on advertising for the popular product, Lipitor, to identify brand-level advertising effects.

advertising as an exogenous shifter of television advertising. Lovett et al. [2019] show that a version of this instrument provides little power in determining the effect of advertising on word-of-mouth. All of these approaches are clever ideas implemented in case-studies. This paper goes beyond a single case-study to characterize the usefulness of the instrument more generally. Thus, it also adds to the literature moving beyond case-studies to characterize empirical results in a generalizable way (Shapiro, Hitsch, and Tuchman [2019]).

Beyond the case of advertising, this paper adds to the literature on instrumental variables, in particular by providing a road map for scrutinizing instruments meant to have broad applicability. This adds to recent work on Bartik shift-share style instruments (Borusyak, Hull, and Jaravel [2018], Goldsmith-Pinkham, Sorkin, and Swift [2018]) and BLP instruments (Gandhi and Houde [2017]) by characterizing the circumstances under which IV approaches which are meant to be generally implementable are valid.[8] Finally, this paper also adds to the literature thinking carefully about the use of instrumental variables in marketing contexts (Rossi [2014]). Our results validate many of Rossi [2014]'s concerns with the implementation of IVs in marketing–in particular, the necessary conditions for validity in this case are generally not innocuous. As a result, researchers wishing to use the instrument require a strong theoretical argument that the assumptions hold in their particular settings.

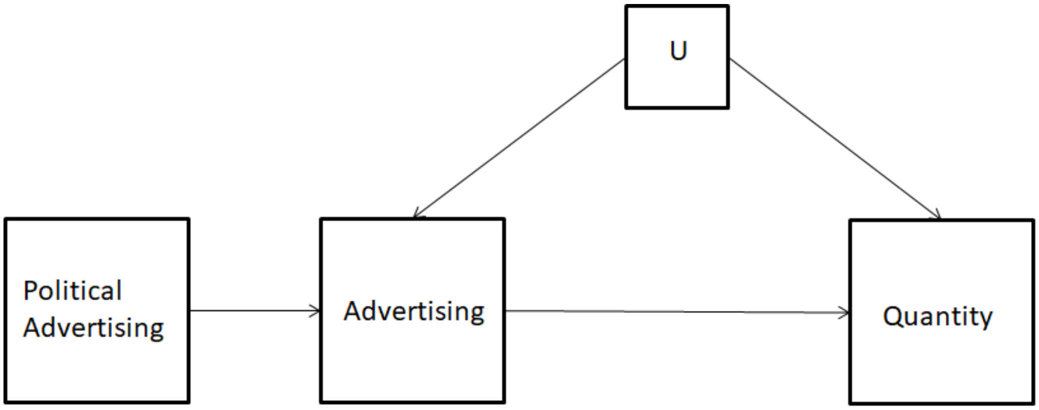
The paper proceeds as follows. Section 2 describes the political advertising IV strategy, including the theoretical justification for the instrument, exclusion restriction, and monotonicity condition. Section 3 describes the Nielsen Ad Intel data and illustrates how political advertising is broadly related to commercial advertising. Section 4 presents the results for the first stage across categories. Section 5 shows a case-study as a proof of concept using the automobile category, which has a strong first stage. Section 6 discusses limitations of the approach. Section 7 concludes.

2 Conditions for Validity

In this section, we present the conditions required for political advertising to identify the causal effect of commercial advertising either at the category level or brand level (for a monopolist). First, political advertising and commercial advertising must be correlated strongly enough to avoid weak

[8] Bartik instruments originate in (Bartik [1991]). The idea is to instrument for employment growth rate with the product of national industry growth rates with local industry shares to estimate the elasticity of labor supply. Since the endogeneity is posited to occur on the local industry growth *rates*, this instrument satisfies exclusion.

Figure 1: Political Advertising and Commercial Demand - a DAG



instrument bias. This condition is testable in the data. Second, we need any correlation between political advertising and commercial sales to operate exclusively through changes in commercial advertising, or an exclusion restriction. This condition is not directly testable, and we will use theory to justify it. Third, we require a monotonicity condition: that each treatment unit exhibits changes in commercial advertising in the same direction for a given change in political advertising. This condition is also not directly testable and will require theoretical justification.

The goal is to estimate the causal effect of category-level advertising on category-level sales. For the purposes of exposition, we will employ a log-log functional form as governing the true relationship of interest:

$$log(1 + Q_{jmt}) = \beta \cdot log(1 + A_{jmt}) + \omega_{jmt}. \tag{1}$$

where j indexes category, m indexes television market, and t indexes time in months. To be clear, we would like to estimate this regression equation category by category. An OLS regression of sales on advertising will not recover the causal parameter of interest (β) because advertising is likely a function of the unobservable ϵ_{mt}. As a result, we propose an instrumental variables approach using political advertising as an exogenous shifter of category-level commercial advertising.

To fix ideas about how the relationship between commercial advertising and political advertising maps into the necessary conditions for instrument validity, we employ a series of directed acyclic graphs (DAGs). Figure 1 illustrates the basic problem. While we want to estimate the effect of advertising on quantity, there exist unobservables, U, that simultaneously drive advertising and demand. We can think of these unobservables as demand shocks. We propose using political advertising as a source of variation in advertising that is unrelated to those unobservables. In each subsection below, we will break apart this figure to identify sources of validity and sources of potential threat.

2.1 Relevance

Figure 2 shows the relationship between the instrument, political advertising, and the endogenous variable, commercial advertising. The intuition for the political IV strategy is that there is a maximum capacity of advertising time. Thus, when candidates advertise, they reduce the residual supply of advertising facing other product categories.[9] The negative supply shock increases equilibrium prices. Because demand curves slope downward, higher prices lower the quantity of airtime demanded by commercial advertisers. Thus, we predict a negative sign on the first stage of commercial advertising on political advertising. In other words, an increase in political advertising in market m at time t should decrease commercial advertising.

It is important to note that while this argument holds directionally for each commercial advertiser, it need not hold with equal magnitude across potential commercial advertisers. The magnitude depends on exactly which ad spots are of the most interest to political advertisers as well as the willingness to pay of commercial advertisers. A main contribution of this paper is to document when and where the first stage is sufficiently strong so as to be useful.[10]

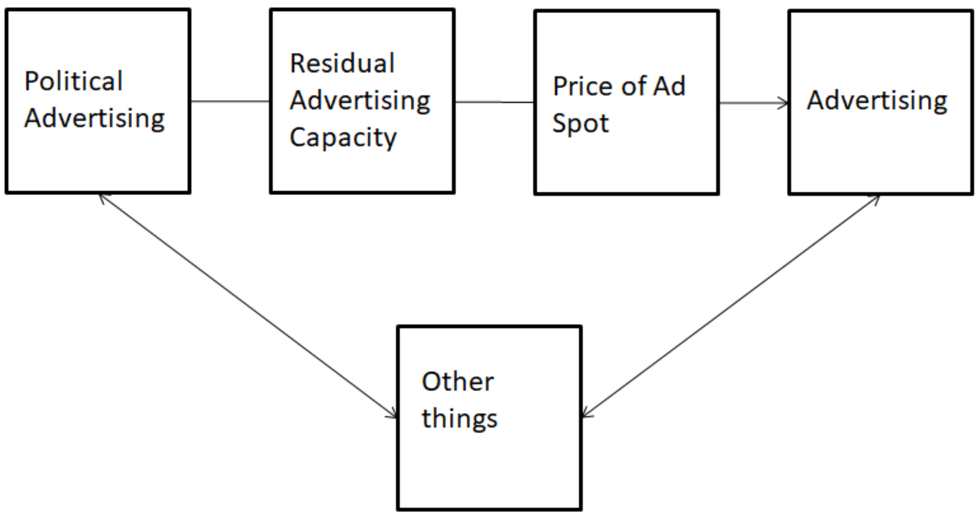
It is also important to note that there may be other sources of correlation between political

[9] Note that this condition does not require a hard capacity constraint, so long as the marginal cost of advertising time increases in the quantity of ads. In theory, local television stations have some flexibility in the airtime they devote to advertising compared to programming (e.g., run shorter or fewer local news segments). If viewers exhibit increasing disutility of advertising time, then the presence of political ads increases the marginal cost of commercial airtime. In our sample, we find very little variation in total advertising time both across stations and over time, suggesting that crowd-out of commercial advertising is nearly 1:1 with the presence of political ads.

[10] Here it is worth pointing out that Lovett et al. [2019] employs a version of this instrument in studying the effect of advertising on word-of-mouth. Instead of estimating effects category by category, it estimates a single effect of advertising on word-of-mouth with a random coefficient. The instruments employed are political advertising interacted with brand intercepts. Overall, it finds that the set of instruments has a weak first stage in that regression and produces several positive coefficients in the first stage, an indication that there might be some monotonicity violations. We will break apart the problem in detail to see exactly where the instrument is weak and where the theory might fail.

advertising and commercial advertising. These factors may threaten either the exclusion restriction or the monotonicity condition, which we discuss below. These potential threats are depicted in figure 2 below the main proposed channel of the relationship between political advertising and commercial advertising.

Figure 2: Political Advertising and Commercial Advertising



2.2 Exclusion

To recover an unbiased estimate of the causal effect of commercial advertising on sales, political advertising must be correlated with product sales exclusively through its correlation with the endogenous variable, commercial advertising. Since political advertising is not truly randomized across markets by campaigns, understanding political demand for airtime is important for justifying this exclusion restriction.

Political advertisers wish to maximize their chance of winning an election given a fixed budget. To accomplish this goal, they must advertise to viewers who meet two criteria: viewers must be persuadable and also potentially pivotal in the election of interest. A viewer is potentially pivotal if changing their vote could change the result of the election. That is, political campaigns do not want to advertise in a market where the election outcome is a foregone conclusion whether they advertise or not—advertising in these markets would waste their scarce budgets. Thus, this campaign objective suggests two determinants of political advertising: (1) the price of airtime and (2) the likelihood that the vote of the marginal ad-viewer will flip the election.

Figure 3 illustrates these sources of demand for political advertising. The potential threats to exclusion are depicted with dotted lines. In particular, we must assume that there is no path leading from the unobservable demand shocks, U, to political advertising. The first threat is a direct path from U to the pivotality of a market. This would be the case if the product market in question were especially politically relevant. For example, if the product market were health insurance on the Affordable Care Act "Obamacare" exchanges, and access to health insurance were a particularly important political issue in a market, demand shocks for health insurance could directly impact pivotality. In this case, the exclusion restriction is violated because some of the variation in political advertising would be contaminated by the unobserved demand shock U.

The second potential threat to exclusion acts through the equilibrium in the market for advertising. That is, a demand shock, U, could lead to higher (lower) commercial advertising for a product category. This increase (decrease) in advertising decreases (increases) the residual supply of advertising, raising (lowering) the price of advertising faced by political campaigns. This price change would lower (raise) the amount of political advertising. In this scenario, there is a negative correlation between political advertising and commercial advertising, but that correlation operates through U, which has a direct effect on commercial demand, violating the exclusion restriction.

We argue that making use of market and time fixed effects alleviates these concerns to an extent. Non-election time periods serve as a 'pre-treatment' period, while high and low political advertising markets serve as 'treatment' and 'control' units, respectively. Exclusion requires that political competitiveness (pivotality) does not drive the relative changes in commercial product demand conditions across markets between election and non-election time periods. This type of violation only seems plausible for very special product categories. As an example, consider a market where a manufacturing plant closed in October of an election year, leaving a large number of people without health insurance. The closure stimulates demand for individual insurance and potentially makes the political election more/less competitive. However, such stories are difficult to tell for the typical product category that advertises on television.

In terms of paths from U to prices faced by political campaigns, the difference-in-differences style variation alleviates these concerns to a large degree. Much of the variation in political advertising comes from variation in the political cycle–there is nearly zero political advertising in the pre-treatment period, not because advertising prices are high, but because there is no impending election.[11] Even during election season, many markets are completely non-competitive, and hence see almost no political advertising. The difference in commercial advertising between those markets and politically competitive markets is significantly larger than the difference in political advertising among the competitive markets. These observations combined lead us to argue that the majority of residual variation in political advertising net of market and time fixed effects is driven by the likelihood that political advertising will shift the election rather than by relative prices.

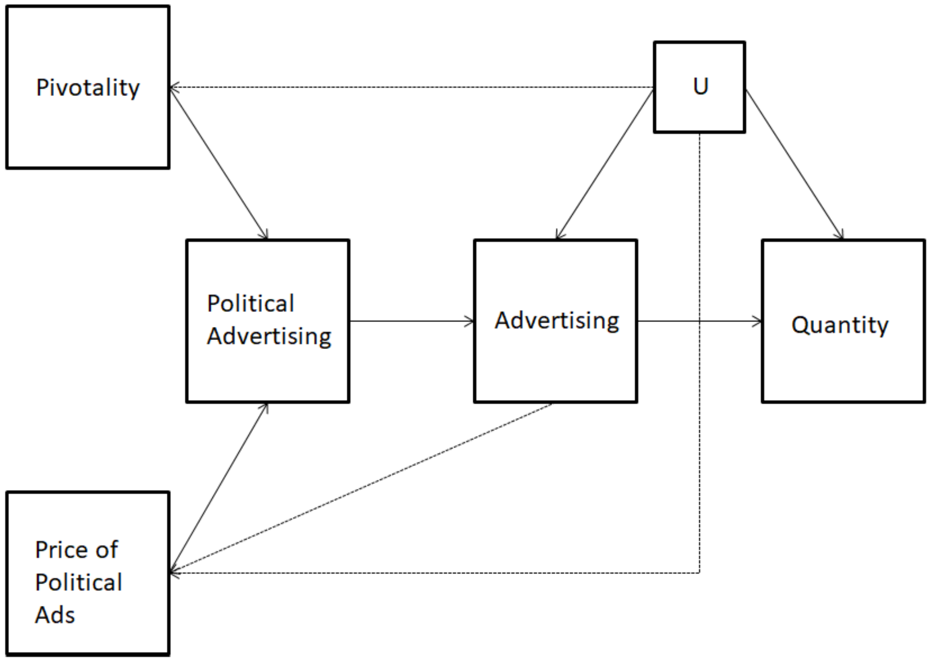
Another consideration is whether firms modify their other promotional activities in response to the political advertising shock. For example, if TV advertising becomes expensive, a firm might substitute from TV to digital advertising, confounding the relationship between TV advertising and sales. Alternatively, a firm might reduce complementary promotional activities in conjunction with a reduction in TV advertising. Fortunately, this potential confound is directly testable with data on these other promotional activities.

Finally, if demand shocks U alter the prices faced by campaigns, then these shocks must also change aggregate commercial advertising. Most commercial categories constitute a small fraction of total advertising, so that even a large change in a single category's advertising would be unlikely to disrupt the market as a whole. As a result, the type of demand shock U that would be most problematic is one correlated across all commercial product categories, such as a negative income shock. If residents of a particular DMA cut spending across all categories, then total commercial advertising demand would fall, increasing the residual supply facing political campaigns, and thus increasing political advertising. In this scenario, more political advertising is spuriously correlated with lower commercial demand, leading to an over-estimate of the effect of commercial advertising on demand. In Section 3.2, we estimate the total disruption in commercial advertising that can be attributed to political advertising, which provides some perspective on the likelihood that the exclusion restriction is violated in this fashion. However, if researchers are particularly concerned about the price mechanism in their specific case, they may wish to find a more direct measure of political pivotality to serve as an instrument for commercial advertising.

Overall, we cannot fully rule out these threats to the exclusion restriction. As a result, we instead

[11] This is shown graphically in Section 3.2 below.

Figure 3: Political Advertising and Commercial Advertising



provide a road map for the required theoretical arguments that must be made to justify exclusion. If a researcher is willing to assume a priori that an advertising effect must be greater than or equal to zero, a positive and significant reduced form coefficient from the regression of commercial demand on political advertising would indicate a violation of the exclusion restriction. Such an assumption might be considered reasonable in the context of category-level advertising effects. In Appendix D, we present reduced form regression results for a set of products in which we have demand data from AC Nielsen's RMS scanner data. For 5 out of 36 product categories, we estimate a positive and significant reduced form coefficient, indicating a possible violation of the exclusion restriction in those categories (although we would expect 1 to 2 categories to be positive and significant by chance).[12] The frequency of these results may suggest to researchers that this exclusion restriction is not entirely benign and requires careful justification on a category-by-category basis.

[12]We do not adjust our inference for multiple hypothesis testing.

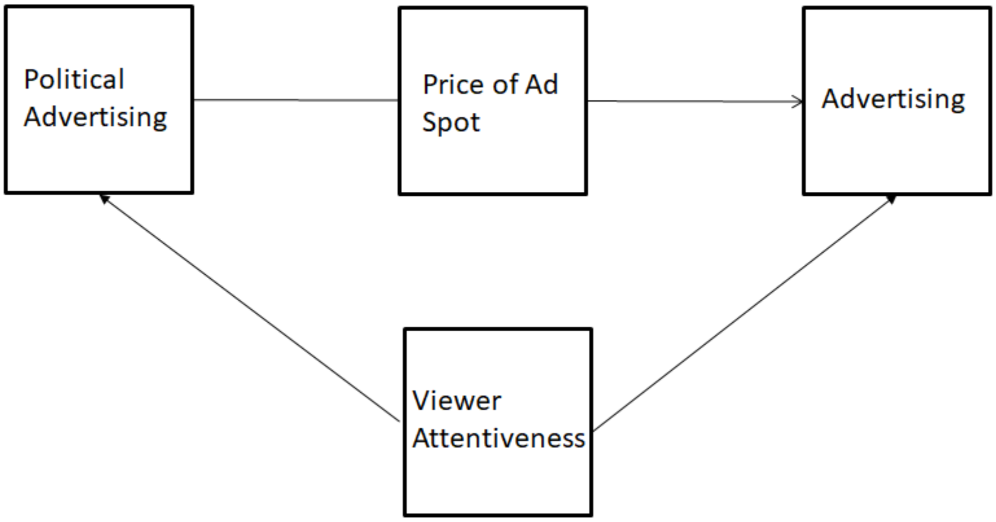
2.3 Monotonicity

Finally, the application of standard instrumental variable techniques requires a monotonicity condition, following Angrist and Imbens [1995]. Under the theory outlined above, where political advertising crowds out commercial ads, monotonicity means that an increase in political advertising must never increase commercial advertising. There are two potential threats to this assumption, and neither involves a relationship between political advertising and commercial advertising through the price mechanism (the lower path in Figure 3).

First, if there are time-varying differences across markets in viewer attention, then political advertising and commercial advertising could exhibit a positive relationship. Both political advertisers and commercial advertisers would like to reach more attentive audiences. If this effect outweighs price effects for any market-time observations, monotonicity will fail. We note that attentiveness threatens the monotonicity condition, even if it does not affect product market sales directly (if attentiveness drove sales, then it would constitute a violation of the exclusion restriction). This mechanism is illustrated in Figure 4. To deal with this potential issue, we include market fixed effects. While this does not solve any issues surrounding attentiveness shocks at the market-time level that can be anticipated by political and commercial advertisers, it does account for the fact that viewers in some markets pay more attention to their televisions. Any residual attentiveness effects on both political and commercial advertising must be assumed to be zero.

The second threat to monotonicity relates to the fact that both political advertisers and commercial advertisers operate in multiple, inter-related markets. First, political advertisers tend to advertise in ways that are correlated across months. In particular, there is more political advertising in all markets in months just before an election than in months far from an election. Figure 5 illustrates how this can be a problem for the monotonicity condition. In particular, consider advertising in market m. If our intuition holds, then political advertising in market m increases the price of a local ad spot in market m, reducing local commercial advertising in that market. At the same time, political advertising also increases in market $-m$. Again, the price of a local ad spot in market $-m$ increases, reducing the quantity demanded for local ad spots in market $-m$. The reduction in local commercial advertising in market $-m$ is not concerning by itself. However, if the viewers in market $-m$ are sufficiently important to the advertiser, it may choose to advertise *nationally*.

Figure 4: Political Advertising and Commercial Advertising



National advertising is seen by viewers in all markets. So while demand for local ad spots in market m decreases, viewers in market m may see more ads in total because more national ads are run. Monotonicity is violated if the increase in national ads dominates the decrease in local ads in market m for the category in question.

This problem is driven by the combination of two factors. First, political advertising is correlated across markets within a time period. Second, national advertising is seen by viewers in all markets. We can control for both of these factors by including a *time fixed effect at the periodicity of the data.*[13] These fixed effects allow us to partial out both national advertising and the correlation in political advertising between markets within a time period, so that only changes in local ad spots in market m identify advertising effects. Conditional on time fixed effects at the level of periodicity in the data, monotonicity violations coming from substitution from one local market to another local market are still possible, but are less obviously plausible than substitution to national advertising.

One way to check the plausibility of the monotonicity condition is to inspect the sign of the first-stage estimate. The theory highlighted in Section 2.1 indicates that political advertising and commercial advertising should be negatively correlated. A positive correlation indicates a possible violation of the monotonicity condition, either coming from shocks to viewer attentiveness or from

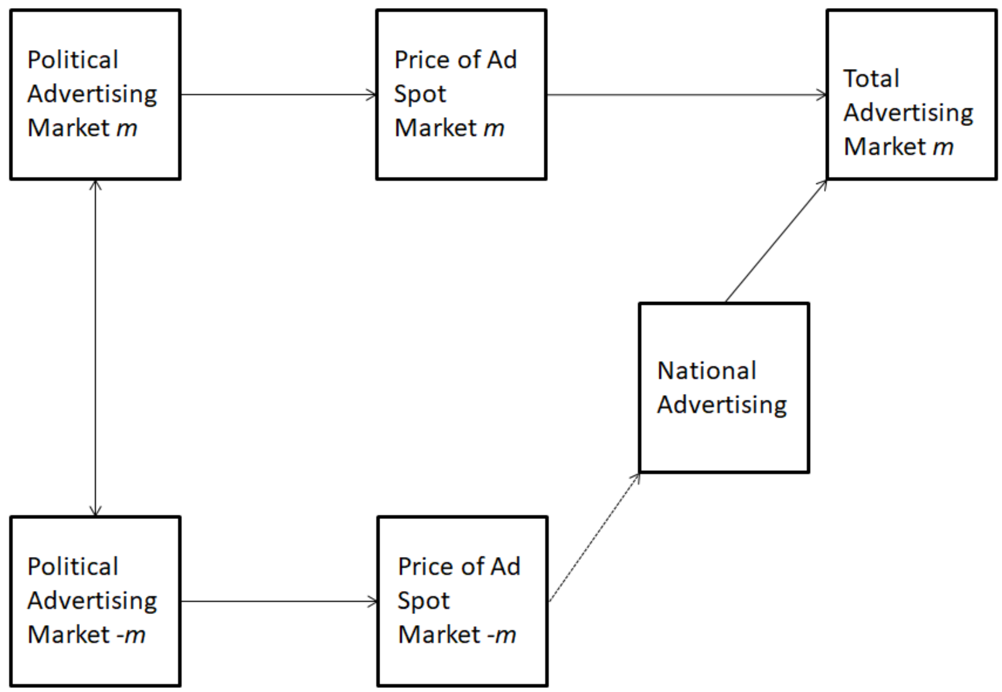
[13] Here we note that Lovett et al. [2019] find many first-stage coefficients that are the wrong sign. That paper also does not employ time fixed effects at the periodicity of the data due to limitations in its data, which we suggest could explain some of the apparent violations of the theory.

substitution across markets.[14] In Appendix F, we show that with market fixed effects and year-month fixed effects, only about 3% of first-stage estimates are positive and significant, roughly the proportion that should occur by chance under a null hypothesis of a zero first stage. However, if we remove the year-month fixed effects or use coarser time controls, nearly 50% of first-stage estimates are positive and significant, and more are positive than negative, violating the theory outlined above. These patterns highlight the importance of time fixed effects at the periodicity of the data to implement this research design.

Here we note again that a researcher may impose further structure on the problem and explicitly model heterogeneous treatment effects rather than use the LATE formulation. This could, in principle, make the monotonicity assumption unnecessary. That said, a positive relationship between political advertising and commercial advertising remains a warning sign that the instrument does not work as the theory prescribes (through offset). As this theory justifies the exclusion restriction, evidence against that theory casts doubt on the exclusion restriction, even if the monotonicity assumption is not needed.

[14] A positive correlation would indicate that a non-price mechanism was contributing to the first stage, either by itself or in conjunction with the posited price mechanism. In the former case, it is likely that more political advertising increases commercial advertising in some circumstances while reducing commercial advertising in others. This would be a violation of monotonicity. Alternatively, a positive first stage relationship could be driven entirely by a non-price mechanism that violates the exclusion restriction.

Figure 5: Political Advertising and Commercial Advertising



2.4 LATE considerations

In the presence of heterogeneous treatment effects, a valid instrumental variable identifies a local average treatment effect (LATE). That is, the treatment effect identified is the average effect among 'compliers' to the instrument, those ads which are crowded out by political advertising. To interpret the effects of crowd-out, it is important to consider how the market clears. If TV stations do not anticipate political advertising demand when selling in the upfront market to commercial advertisers (typically in May or June), then they may displace commercial ads in response to an influx of political dollars. Indeed, many contracts include provisions for the priority of the airtime: a low-priority ad is more likely to get displaced than a more expensive high-priority ad. If an ad spot is displaced, then the station issues a "make-good" and the spot is aired at another time. While make-goods supposedly offer comparable viewership, industry wisdom is that they are generally of inferior quality. Because priority is explicitly priced into commercial advertising contracts, political ads presumably displace spots for which commercial WTP is low.[15]

Such selection may be more severe if stations anticipate political demand. In this case, stations should set higher prices in the upfront market in election years, particularly in contested states. High prices reduce advertising by the lowest WTP commercial firms. Therefore, political GRPs instruct us about the treatment effect from the bottom of the distribution of advertising efficacy.

An issue related to both LATE considerations and the exclusion restriction is the potential for commercial firms to coordinate other marketing levers, such as sales force or price promotions, with its TV advertising strategy. In this case, even if political advertising does not directly affect this other lever, it may indirectly change other marketing efforts through its effect on commercial advertising. Any indirect effects would change the interpretation of the LATE. Political advertising would then identify the 'total' effect of advertising plus other marketing activities, including the direct path from advertising to sales as well as a path from advertising to other marketing levers to sales. To be clear, the LATE would not identify how TV advertising affects sales holding all else equal. This concern must be evaluated on a case-by-case basis: it is a non-issue for firms that set marketing levers independently, but a potential problem for others that link them algorithmically. If possible, we advocate directly testing the relationship using data on these other marketing activities.

2.5 Preferred Specification

Our regression of interest, described in equation 1, is a standard log-log advertising regression for category j in market m on contemporaneous sales at time t. The goal of this paper is to use political advertising as an instrument for category advertising flow, A_{jmt}.[16] In order to address the

[15]Category-level average WTP may change between different political cycles if the nature of competition in those categories changes over time. For example, in one election, it may be that a category has a new entrant firm that has a very high WTP for advertising space due to a need to inform customers of its existence. In a future election, that firm is no longer a new entrant, and as a result has a lower WTP for advertising space. In such an example, the category-level first stage could increase in a future example. These kinds of considerations are important for researchers to keep in mind when they interpret the LATE associated with the political advertising IV.

[16]We focus on estimating the effect of advertising flow even though advertising is typically thought to have accumulating stock effects. If, in August, firms anticipate higher prices of advertising in October due to political campaigns, they may increase advertising in August to take advantage of stock effects that linger into November. Such anticipation and dynamic effects would pose problems for the validity of this specification. If we specify our model with an accumulating advertising stock as the endogenous variable to account for this problem, the dynamic behavior would serve to weaken the first stage, as current period crowd-out would be partially offset by previous period anticipatory effects. In Appendix E, we specify the model using advertising stock as the endogenous variable of interest and find that the instrument is generally weaker. As a result, the main analysis should be viewed as perhaps an over-optimistic

most plausible threats to monotonicity and the exclusion restriction, our final specification includes market fixed effects and time fixed effects at the periodicity of the data. We specify our first stage equation as:

$$log(1 + A_{jmt}) = \gamma \cdot f(P_{mt}) + \alpha_m + \alpha_t + \epsilon_{jmt}. \quad (2)$$

In the equation above, A_{jmt} is the GRP for category j in market m in month t, P_{mt} is political advertising market m in month t, γ governs the relationship between political and commercial advertising, and α_m, α_t are market and month fixed effects. We specify category advertising on the left-hand side as a log because researchers typically want to identify an advertising elasticity. Here, we leave the relationship between the endogenous variable and the instrument reasonably general, as we will show robustness to many different first stage functional form specifications including a flexible one using machine learning. Next, we specify the reduced form equation as:

$$log(1 + Q_{jmt}) = \pi \cdot g(P_{jmt}) + \alpha_m + \alpha_t + u_{jmt}. \quad (3)$$

2.6 Brand-Level Advertising Effects

Our preferred specification estimates the effect of advertising on sales for an entire category and, in the case of monopoly, for a particular brand. Even in oligopoly settings, category-level advertising elasticities are of interest because they reveal the extent to which advertising is market-expanding. For a good like cigarettes, the importance of an advertising ban hinges precisely on whether advertising increases smoking. However, in other instances, it is important to quantify business-stealing; such quantification is challenging with the political advertising instrument because it affects all firms within a market simultaneously. If the researcher is willing to impose additional structure on the relationship between advertising and sales, however, then more progress can be made. As an example, one approach is to model s_{kmt}, the share of brand k in market m at time t, using the logit as follows:[17]

view of the strength of the first stage.

[17]The utility of the outside option is normalized to zero.

$$s_{kmt} = \frac{\exp\left\{X'_{kmt}\beta - \alpha \cdot p_{jkt} + \gamma \cdot Ads_{kmt} + \xi_{kmt}\right\}}{1 + \sum_{k' \in K} \exp\left\{X'_{k'mt}\beta - \alpha \cdot p_{k'mt} + \gamma \cdot Ads_{k'mt} + \xi_{k'mt}\right\}}.$$

Using the standard Berry [1994] inversion, the difference between the share of brand k and the outside option s_{omt} can be written as:

$$\ln s_{kmt} - \ln s_{omt} = X'_{kmt}\beta - \alpha \cdot p_{kmt} + \gamma \cdot Ads_{kmt} + \xi_{kmt}, \tag{4}$$

which we can estimate using linear IV with political advertising as an instrument. The key restriction is that a single parameter (γ) governs both own- and cross- advertising elasticities:

$$\begin{aligned} \epsilon_k &= \gamma \cdot Ads_{kmt} \cdot (1 - s_{kmt}) \\ \epsilon_{kk'} &= -\gamma \cdot s_{k'mt} \cdot Ads_{k'mt} \text{ for all } k' \neq k \end{aligned}$$

The idea is analogous to estimating a price coefficient in a logit demand function using a single industry-wide cost shock.

Finally, we note that in this framework, the category-level F-statistics presented below are informative about the brand-level F-statistics. To be clear, the first stage that corresponds to equation 4 is:

$$Ads_{kmt} = X'_{kmt}\pi_o + \pi_1 \cdot PGRP_{mt} + \epsilon_{kmt}. \tag{5}$$

It is important to cluster standard errors at least at the market-time ($m \times t$) level in this brand-level regression (equation 5) precisely because this is the level at which political advertising shocks occur.[18] The category-level regressions exploit exactly this market-time variation, and so the category-level F-statistics speak to the magnitude of this variation. Said differently, disaggregating the data to the brand-level rather than the category-level does not buy the researcher any additional variation

[18] In practice, due to repeated observations in the outcome, practitioners will typically want to cluster standard errors at the market level. For that reason, we cluster at the market level throughout. In either case, the relevant clustering for the brand-level formulation and the category-level formulation is the same, as the relevant quasi-exogenous variation is at the same level.

in political advertising, so will not typically buy the researcher additional first stage strength.

3 Data and Aggregate Analysis

3.1 Ad-Intel Data

Our main source of data is the Nielsen Ad-Intel database that records all TV advertisements in 130 DMAs in the contiguous United States from 2010 to 2016.[19] The ads are recorded at the occurrence level, where an occurrence is the placement of an ad by a specific advertiser on a given channel, in a specific market, for a specific duration, on a given date and time. Nielsen uses its propriety technology to collect TV programming information and identify advertising occurrences based on their unique audio and video content.[20] Occurrences are therefore measured in the same way in all of the 130 DMAs used in our analysis.

Because occurrences can differ substantially in their reach, estimated impressions (the number of eyeballs viewing an ad) can serve as a scaling factor to help in comparisons between ad occurrences. For each ad occurrence, we calculate gross rating points (GRPs), a frequently used measure of advertising intensity, as the number of impressions for the ad as a percentage of all TV-viewing households in a DMA.

The Ad-Intel database provides impressions estimates across markets and media types. For local media types the database provides impressions at the "station–month–day of week" level in 5-15 minute time intervals.[21] In the top 25 DMAs, Nielsen measures impressions using "Local People Meters" that capture all TV-viewing activities of Nielsen households, so the data is available in all months.[22] In all other DMAs, Nielsen measures impressions using diaries filled out by Nielsen households, and the data is only available in February, May, July, and November ("sweep months").[23] For non-sweep months in those DMAs, we impute impressions by taking the average between the

[19]The data is missing for North Platte, NE DMA in June 2011, and for Yakima-Pasco-Richland-Kennewick, WA DMA in December 2014.

[20]"Nielsen Monitor-Plus Methodology by Medium", page 19.
http://en-us.nielsen.com/sitelets/cls/documents/adviews/AdViews-Methodology-by-Medium-InfoKit.pdf

[21]These include spot TV, network TV broadcast locally, and syndicated TV broadcast locally.

[22]In 2016, Nielsen started providing impressions data in all months for 70 markets, including the 25 LPM markets and 45 "Set Meter" or "Code Reader" markets.

[23]The impressions data is missing for Birmingham (Ann and Tusc), AL DMA in May 2011.

two closest sweep months.[24] We show in Appendix C that our results are robust to a more flexible imputation method. For national media types (we only use cable TV), the impressions are measured at the program level but are only available nationally; we thus compute GRPs for those ads at the national level and assume that they are the same in every market.[25]

Nielsen groups advertisers into 343 categories. Political campaigns and unions are classified as "B181: Organization Advertising: Political, Union," and Super PACs are found in a separate category, titled "B189: Miscellaneous Organization Advertising." We identified 616 advertisers in the latter category as super PACs by manually matching the advertiser names with the list of super PACs created by OpenSecrets.[26] We then calculate total political advertising at the DMA-month level by summing up the GRPs for all category B181 ads and super PAC ads. Table **1** provides summary statistics for political advertising at the DMA-month level. Of particular note is that political advertising makes up a larger share of total GRPs than of total duration. This means that, on average, political advertisers buy spots with high viewership. While political advertising accounts for as much as 9.39% of total advertising views in a market-month, it never makes up more than 1.53% of the total advertising airtime.

Among the 342 non-political categories,[27] we use the top 274 that constitute 99.9% of total GRPs for analysis. We aggregate the data to the category-DMA-month level by summing up the GRPs of individual ads.

Table 1: Summary Statistics: Political Advertising at DMA-Month Level ($N = 10917$)

			Quantiles								
	Mean	SD	Min	5%	10%	25%	50%	75%	90%	95%	Max
Political GRP	4484.5	9017.1	0.0	5.4	19.3	85.8	723.0	4421.7	13698	22446	102791
Political/Total GRP (%)	0.53	1.03	0.00	0.00	0.00	0.01	0.08	0.54	1.65	2.64	9.39
Political Duration (Hours)	14.92	27.39	0.00	0.09	0.27	1.17	4.03	15.26	42.74	68.59	362.37
Political/Total Duration (%)	0.08	0.14	0.00	0.00	0.00	0.01	0.02	0.08	0.23	0.36	1.53

[24] We weight the data from two closest sweep months by time difference. For example, for March we use 2/3 February and 1/3 May.

[25] For more description of processing the Nielsen Ad-Intel database and the data computed from it, please see Shapiro et al. [2019].

[26] https://www.opensecrets.org/pacs/superpacs.php

[27] We re-define the "B189: Miscellaneous Organization Advertising" category after taking out the 616 identified super PACs.

3.2 Political Advertising and the Commercial Advertising Market

In this section, we document how the commercial advertising market as a whole fluctuates with movements in political advertising. We begin by overlaying the time series of political advertising and commercial advertising for the Columbus, OH DMA. Figure 6, panel (a) shows these time series in levels. It plots the time series of average daily GRP in each month; both series are de-meaned (the means are shown in the legend). Here, it appears that political advertising offsets commercial advertising almost 1-1. Figure 6, panel (b) shows the time series in log-scale for both commercial and political advertising. Here, we see that large percentage changes in political advertising lead to hardly any change in the percentage of commercial ads. Changes in log commercial advertising are nearly imperceptible in the picture. Panels (a) and (b) can be reconciled by the fact that each political ad GRP crowds out roughly one commercial ad GRP, but because political advertising is a relatively small share of total advertising, the total disruption of the commercial ad market is small in percentage terms.

Next, we analyze the relationship between political and commercial advertising systematically across all markets in regression form. In particular, we estimate regressions of the form:

$$\sum_{j \in J} A_{jmt} = \gamma \cdot P_{mt} + \alpha_m + \alpha_t + \epsilon_{mt}, \tag{6}$$

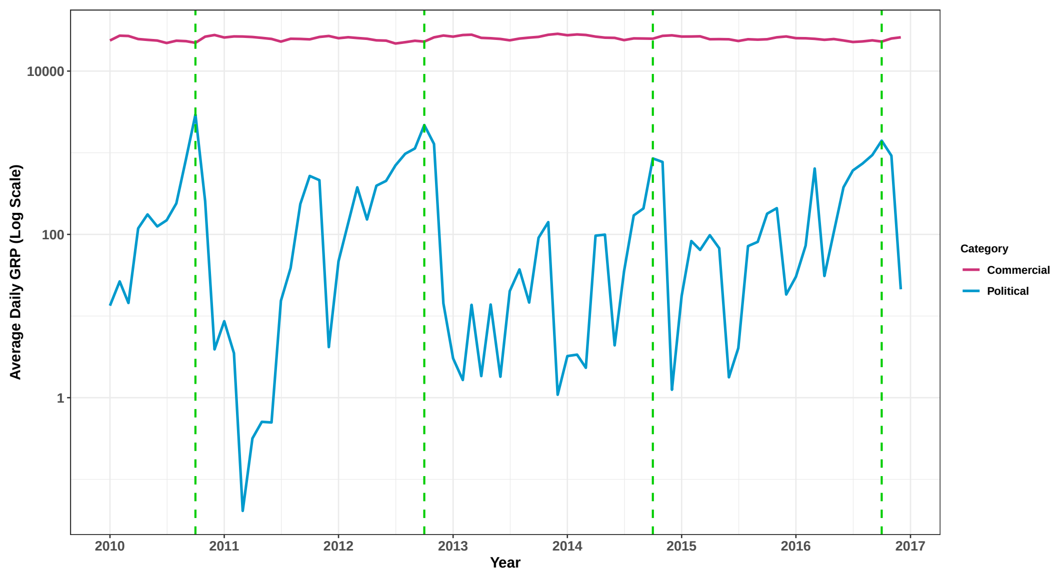
where J denotes a set of categories; A_{jmt} is the amount of advertising in commercial category j, in market m, in month t; and P_{mt} is the amount of political advertising in market m, and in month t. We include market and time fixed effects. We consider A and P measured both in duration (hours) and in GRPs. The coefficient γ measures the crowd-out effect in levels.

Panel (A) of Table 2 describes the results where A and P are measured in GRPs. Column (1) presents the results for all commercial advertising–that is, all non-political advertising, except for ads that promote the television station's own programming. Our estimates suggest that each political GRP offsets 0.85 commercial GRPs. Column (2) shows the effect of political advertising on advertising for the television station's own programming. It is negative and significant, but small–each political GRP offsets about 0.045 programming GRPs. Column (3) shows the effect of political advertising on all non-political advertising. Each political GRP offsets about 0.86 non-political GRPs. Notably, this number has a confidence interval that excludes 1, which means that

Figure 6: Time Series of Political and Commercial Advertising in Columbus, OH DMA

(a) Average Daily GRP in Each Month (Time Series Mean Taken Out)

(b) Average Daily GRP in Each Month (Log Scale)



Notes: Time-series means for panel (a) are taken out and listed in the legend.

political GRPs lead to an expansion of total advertising GRPs (shown in column (4)). Each political GRP expands total GRPs by roughly 0.14. This suggests that while political advertising offsets commercial advertising, it also expands the total amount of advertising viewed on television. Panel (B) of Table 2 describes the analogous results where A and P are measured in duration rather than in GRPs. The findings are qualitatively and quantitatively similar for each column. To give a sense of how the results look in percentage terms, we also report the approximate percentage effects of an 1% increase of political advertising from its mean in Table 2. We observe that the magnitudes in percentage terms are considerably smaller. For example, column (1) of Panel (A) shows that a 10% increase in political advertising corresponds to a 0.06% decrease in commercial advertising. In other words, although political ads crowd out commercial ads almost one to one, the effect is still small in percentage terms because there are far fewer political ads than there are commercial ads.[28]

Overall, it appears that while political advertising offsets commercial advertising, it has a relatively minimal impact (percentage wise) on the advertising market as a whole.

[28]Calculation based on Figure 6 and the point estimates in Table 2 shows that about 7% of total commercial advertising was offset in Columbus, OH in the month with the most political advertising (October 2010). In all other months and in most other markets, the offset is considerably lower.

Table 2: Crowd-out Effect for Aggregate Categories

Ad Category:	Commercial (1)	Programming (2)	Non-Political (3)	Total (4)
Panel A: Ads Measured in GRP				
Political	-0.850^{***}	-0.045^{***}	-0.859^{***}	0.141^{**}
	(0.054)	(0.013)	(0.063)	(0.063)
Partial F	247.6	11.5	184.1	4.9
Partial R^2	0.062	0.002	0.043	0.001
Percentage Effect (%)	-0.0056	-0.0014	-0.0046	0.0008
Panel B: Ads Measured in Duration				
Political	-0.798^{***}	-0.021^{*}	-0.770^{***}	0.230^{***}
	(0.049)	(0.011)	(0.050)	(0.050)
Partial F	265.2	3.8	233	20.8
Partial R^2	0.05	0.001	0.036	0.003
Percentage Effect (%)	-0.0008	-0.0001	-0.0006	0.0002

"Programming" includes all categories for TV programs and TV networks/stations. "Commercial" includes all categories except "Programming" and "Religious, Charitable, and Humanitarian Organizations". "Non-political" includes all categories except political ads.
"Percentage Effect" is the coefficient multiplied by 1/100 of the mean of independent variable and then divided by the mean of dependent variable. It approximates the crowd-out effect for an 1% increase of political advertising in percentage terms, and is close to the estimate from a log-log specification.
Standard errors in parentheses, clustered at the DMA level. DMA and month $\times$ year fixed effects included.
$N = 10917$. *p<0.1; **p<0.05; ***p<0.01

4 Political Advertising as an Instrument

In this section, we test the strength of the political advertising instrument across 274 categories. We first employ equation (2) and set $f(P_{mt}) = P_{mt}$ in our benchmark "log-linear" specification. We specify category advertising as a log in our benchmark model because researchers often want to identify an advertising elasticity, which is most natural with the endogenous variable in log form. We specify political advertising in levels in our benchmark model for simplicity.[29]

To ensure that our qualitative results are not driven by our somewhat arbitrary choice of functional form, we take two additional approaches. First, we estimate a second specification where both category advertising and political advertising are measured in levels. Second, we allow the first stage

[29]Ex post, the linear term is the most frequently selected term in the LASSO specification, suggesting that it is typically the best single first stage predictor in terms of functional form.

shape to be guided by the data using a machine learning specification to pick the functional form of political advertising. We do so using the LASSO method in Belloni, Chen, Chernozhukov, and Hansen [2012] to select the optimal $f(P_{mt})$ by combining 40 non-linear functions. Overall, our results are robust across alternative specifications and measurements, indicating that the qualitative results are not driven by the choice of functional form. Finally, we document several category-level characteristics that are linked to a strong first stage.

We cluster standard errors at the television market level to account for the potential serial correlation of errors within a market over time. We assess the instrument strength using the effective F-statistic of Olea and Pflueger [2013], which is identical to the Kleibergen-Paap partial F-statistic in the case with one endogenous variable and one instrument. For each specification, we classify categories into three bins: $F > 25$ (strong), $F \in [10, 25]$ (semi-strong), and $F < 10$ (weak). We use these bins because researchers employ a variety of rules to determine whether an F-statistic is "large enough" to allow for standard inference.[30] One common rule of thumb suggests that $F > 10$ is sufficient. However, Olea and Pflueger [2013] suggest a critical value of 23.11 for rejecting H_0 : Nager Bias $< 10\%$ with one instrument. As it is generally accepted that $F < 10$ constitutes a weak instrument, weak-instrument robust inference (e.g., Anderson and Rubin [1949], Andrews [2016]) should be used for the $F < 10$ categories. For categories where $F \in [10, 25]$, whether the instrument is sufficiently strong depends on the relevant critical value to the researcher on a case-by-case basis, but for $F \in [10, 25]$ we suggest that researchers consider using weak-instrument robust inference. While these two breakpoints are arbitrary, they correspond roughly to the critical values for rejecting the null of weak-instruments at different tolerance levels. For our benchmark specification, the Olea and Pflueger [2013] critical value is 23.11 for rejecting H_0 : Nager Bias $< 10\%$ and 12.05 for rejecting H_0 : Nager Bias $< 30\%$ at 5% level; the Stock and Yogo [2005] critical value is 16.38 for rejecting H_0 : Maximal Size $< 10\%$ at 5% level.

[30] It might be suggested that we adjust our F-statistics to account for the fact that we are testing 274 hypotheses at one time. On one hand, we are indeed testing 274 categories at the same time in this paper, so the adjustment may seem necessary. On the other hand, other researchers using this instrument will likely focus on one category at a time, so the unadjusted F-statistics should be the objective of interest. To the extent that the reader would prefer adjusted F-statistics, our F-statistics reported here would represent an upper bound on the adjusted F-statistics.

4.1 First Stage Results

We begin by evaluating the first stage employing the level of political advertising (in GRPs) as a single instrument for the log of category advertising, which we call the "log-linear" specification. We will discuss these results in some detail and then present the results for alternative functional form choices.

The distribution of partial F-statistic is reported in Figure 7, in which the vertical line is at the Stock and Yogo [2005] critical value of 16.38. The partial-F ranges from 0.0004 to 327.5, and it is greater than 25 for 28 categories, between 10 and 25 for 25 categories, and below 10 for 221 categories. The categories with $F > 25$ are listed in Panel A of Table 3, the categories with $F \in [10, 25]$ are listed in Panel B of Table 3, and the categories with $F < 10$ are listed in Table 7 in Appendix A.

The coefficient β_j in the log-linear first stage represents the crowd-out effect in log points of one extra political GRP. To make the coefficient easier to interpret, we scale it by the median political GRP across markets in October 2016 (which is 23,438). The scaled coefficient thus represents the median crowd-out effect in log points across markets in the most heated month before election. Figure 8 plots the distribution of scaled coefficients, which range from -0.264 to 0.088 and have a median of -0.008 across all categories. The median effect is -0.064 for the $F > 25$ categories, -0.028 for the $F \in [10, 25]$ categories, and -0.005 for the $F < 10$ categories.

The signs of the first stage coefficients can also be used as a check on the theory used to justify the exclusion restriction and monotonicity condition. For the set of 28 categories with $F > 25$, the first-stage sign is negative for 27—consistent with crowd out—with the sole exception of "Miscellaneous Organization Advertising". Although not easily identified by our hand-matching, many of these organizations may be political, so that a positive first stage reflects the importance of political pivotality for these advertisers. Overall, for the categories where the instrument appears strongest, we find no indication of a violation of our theory in the first-stage coefficient signs. For the set of 25 categories with $F \in [10, 25]$, only one has a positive sign, and that is "Cellular Radio Systems & Accessories." For the set of 221 categories with $F < 10$, 64 have a positive point estimate on the first-stage coefficient. Many of these may be due to chance, as the first stage is quite weak. For our specifications, in the categories for which we have a strong first stage, we only see a 'wrong sign' first stage for one commercial advertising category. It appears that the market and time fixed effects do a reasonable job of correcting for the potential monotonicity concerns. While these patterns do not prove that there are no monotonicity violations (this assumption is fundamentally untestable), it does speak to perhaps the most plausible ones.

Figure 7: Distribution of Partial F-Statistic, Log-Linear

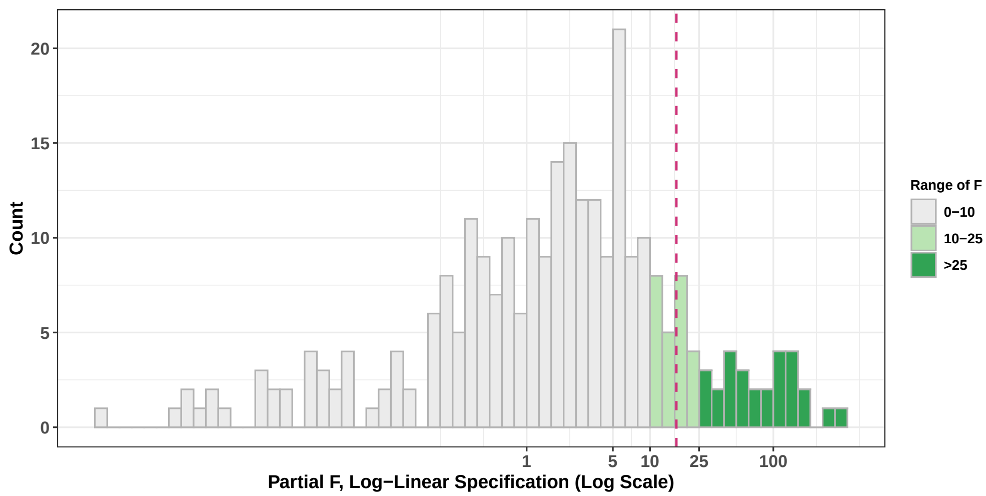


Figure 8: Distribution of Scaled First-Stage Coefficients, Colored by Range of Partial-F

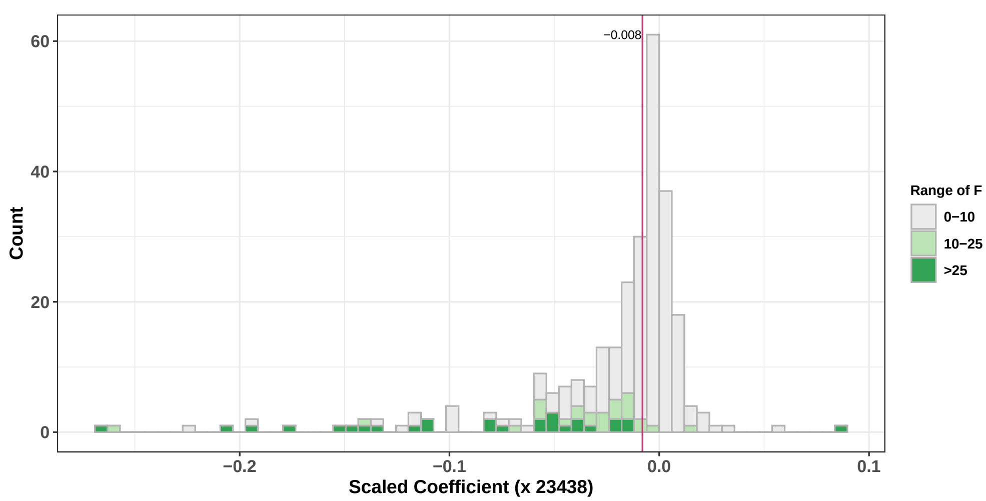


Table 3: List of Categories with Strong or Semi-Strong Log-Linear First Stage

Category	Partial F	Partial R^2	Scaled Coef	Size Rank	Spot Share	GRP/OCC
Panel A: 28 Categories With $F > 25$						
Household Furnishings & Appliance Stores	327.5	0.058	-0.134	13	0.74	2.8
Autos & Light Truck - Dealerships	280.2	0.035	-0.197	16	0.96	2.9
Hospitals, Physicians & Misc. Physical Culture	197.2	0.031	-0.117	25	0.68	2.6
Optical Goods and Services	184.6	0.018	-0.139	70	0.59	2.6
Misc. Entertainment & Combination Copy	144.6	0.020	-0.264	78	0.80	2.3
Passenger Cars-Factory: New & CPO	143.0	0.026	-0.040	6	0.24	2.9
Miscellaneous Professional Services	141.2	0.028	-0.051	3	0.36	1.8
Restaurants, Hotel Dining & Nightclubs	135.4	0.026	-0.031	2	0.22	2.7
Construction, Engineering & Architecture	122.2	0.018	-0.154	61	0.84	2.7
Miscellaneous Retail	114.4	0.023	-0.043	7	0.25	2.5
Passenger Cars-Dealer Assn: New & CPO	106.8	0.018	-0.206	55	1.00	3.1
Light Trucks & Vans-Factory: New & CPO	105.9	0.024	-0.051	15	0.29	3.0
Misc. Organization Advertising	96.5	0.028	0.088	26	0.31	1.9
Dance, Theater, Concerts, Opera	91.3	0.014	-0.149	76	0.54	2.6
Automotive	72.2	0.011	-0.083	58	0.41	2.2
Schools & Camps, Seminars	71.9	0.014	-0.054	18	0.37	1.5
Light Trucks & Vans-Dealer Assn: New & CPO	63.0	0.011	-0.176	77	1.00	3.2
Data Communications Networks	53.5	0.010	-0.050	28	0.29	2.0
Misc Financial Inst. Services & Products	52.4	0.009	-0.042	33	0.25	1.6
Banks	45.0	0.006	-0.073	65	0.43	3.3
Cakes, Pies, Pastries & Donuts	43.1	0.005	-0.022	116	0.10	1.6
Cable Television Stations	42.9	0.007	-0.054	42	0.36	2.1
Direct Response Products	40.3	0.011	-0.012	8	0.05	0.9
Apparel, Footwear & Accessory Stores	36.8	0.004	-0.020	31	0.14	2.5
TV Station	35.4	0.006	-0.113	51	0.99	2.1
Fix-it Supplies	31.3	0.002	-0.113	203	0.13	1.8
Cereals	30.1	0.007	-0.013	20	0.06	1.8
Real Estate, R.E. Brokers & Developers	27.2	0.004	-0.083	123	0.43	2.5
Panel B: 25 Categories With $F \in [10, 25]$						
Plumbing & Sanitary Equipment	23.6	0.003	-0.026	113	0.11	2.4
Amusement Parks & Sporting Events	23.0	0.002	-0.059	71	0.48	2.2
Appetizers, Snacks & Nuts	21.4	0.003	-0.009	35	0.04	1.7
Automobile Insurance	20.0	0.004	-0.010	11	0.07	1.9
Medical & Dental Insurance	18.9	0.002	-0.055	69	0.43	2.0
Life Insurance	18.3	0.004	-0.023	97	0.07	1.2
Golf Equipment	17.9	0.004	-0.028	216	0.00	1.9
Breads, Rolls, Waffles & Pancakes	17.6	0.001	-0.033	95	0.18	1.6
Hotels & Resorts	17.2	0.003	-0.039	45	0.32	2.7
Food & Liquor Stores	16.8	0.003	-0.072	57	0.77	3.2
Religious, Charitable & Humanitarian Org.	16.1	0.003	-0.040	49	0.33	1.3
Lotteries	15.9	0.002	-0.262	142	1.00	3.2
Magazines, Newspapers, Misc Media	15.1	0.002	-0.034	107	0.15	2.1
Other Insurance & Combination Copy	14.4	0.003	-0.015	19	0.13	2.0
Medical Appliances & Equipment	14.3	0.001	-0.029	159	0.07	2.4
Cold, Cough & Sinus Remedies	12.7	0.001	-0.012	21	0.04	1.8
Misc Accessories, Supplies & Hardware	12.6	0.001	-0.054	175	0.11	2.2
Coffee, Tea, Cocoa & Derivatives	12.2	0.002	-0.016	85	0.06	2.5
Cellular Radio Systems & Accessories	11.8	0.003	0.013	46	0.03	2.8
Jewelry, Gift Stores & Galleries	11.2	0.001	-0.042	56	0.28	2.2
TV Program: Late Night News	11.0	0.002	-0.139	155	0.92	2.9
Comb Copy & Misc Major Appliances	10.4	0.001	-0.023	162	0.05	2.0
Cheese Products	10.2	0.002	-0.017	101	0.09	2.2
Computerized Games, Accessories & Software	10.2	0.001	-0.006	27	0.02	2.0
Drugs, Toiletries & Salons	10.2	0.001	-0.021	72	0.16	2.6

Table Notes: "Size Rank" is the category rank by total GRP across all markets and years. "Spot Share" is the share of ads bought in spot markets. "GRP/OCC" is the average GRP per occurrence. See Section 4.3 for further discussions about those category characteristics. "Scaled Coefficient" is the log-linear regression coefficient times 23438.

4.2 Robustness to Specification and Measurement

In this section, we examine the robustness of the first-stage results above to alternative specifications and measurements.

First, we show that the first-stage results are stable if we specify the category ads (left hand side of equation 2) in levels instead of logs. Next, we use a machine-learning method to capture the potential non-linear effect of political advertising. Finally, we show that our results are robust to measuring advertising in total duration of advertising, which is measured consistently across markets and time (in lieu of GRP).

We start by specifying category advertising in levels rather than in logs in the first stage. Column (3) of Table 4 summarizes the results for the "linear-linear" specification. Comparing columns (3) and (1), changing the category GRPs from logs to levels adds 5 categories to the $F > 25$ bucket and 5 categories to the $10 < F < 25$ bucket. Figure 10 compares the partial-F for all categories between the two specifications, and shows that the identities of the strongest categories are not changed.

Next, we employ a LASSO to try and get additional first stage strength through a better functional form fit. Technical details of the LASSO formulation are presented in Appendices B.1 and B.2.[31] The basic statistics of the "Log-Lasso" and "Linear-Lasso" specifications are listed in columns (2) and (4) of Table 4, respectively. Comparing with the "Log-Linear" specification, 3 categories are "downgraded" from $F > 25$ to $10 < F \leq 25$, but 28 categories are "upgraded" from $F < 10$ to $10 < F \leq 25$.[32] This is potentially a notable improvement, depending on which critical values are relevant to the researcher. However, zero categories are "upgraded" from $F < 25$ to $F > 25$. The identities of the strong instrument categories remain stable. The "Linear-Lasso" compares similarly with the "Linear-Linear" model. Overall, the results of the first stage analyses employing a LASSO are not qualitatively different from the benchmark models with a single instrument and do not

[31]Intuitively, we employ this approach to make sure we choose a functional form of political advertising to get every last bit of available strength out of the first stage. If we specify a linear model and the first stage is highly non-linear, we may find a weak instrument but there exists a specification where that instrument might be strong. Conducting this LASSO exercise comes at a cost, as it requires further assumptions and may make final estimates difficult to interpret in the presence of heterogeneous treatment effects. For the purposes of our exercise, the LASSO is simply meant to demonstrate how much additional strength could be added to the first stage. Authors should carefully consider the additional required assumptions of the LASSO if they choose to employ such a formulation in order to increase first stage strength.

[32]Further analysis in Appendix B.3 shows that the LASSO is indeed picking up crowd-out relationships that are very different from log-linear for those 28 "upgraded" categories, though for some categories the first stage improvement comes at the cost of monotonicity concerns.

produce any additional categories where political advertising is clearly a strong instrument.[33]

Finally, we show robustness to using advertising measured in duration rather than in GRP. This exercise is helpful for two reasons. First, as we discuss in Section 3.1, the impressions data for the 105 non-LPM DMAs are only measured by Nielsen in 4 sweep months (February, May, July, and November), so we must impute the impressions for non-sweep months. Because the vast majority of political ads occur from August to October in election years, one might worry that this induces measurement error in our analyses based on GRPs. In contrast, advertising duration is measured by Nielsen using the same technology across all 130 DMAs. Second, the intuition of crowd-out is that a political ad takes the place of a commercial ad, a mechanism which operates through ad duration rather than viewership. Specifying the model in duration as such may increase first stage strength.[34]

Columns (5) to (8) in Table 4 summarizes the first-stage results when both category ads and political ads are measured in duration (minutes). By comparing panels (A) and (B) for columns (5)-(8) to columns (1)-(4), we see that the first-stage results using durations are generally weaker. However, panel (C) suggests that the identities of the strongest categories are not changed, though the partial-F stats become smaller. Figure 9 confirms this finding by showing that the strong categories in the GRP-log-linear specification are also relatively strong in the duration-log-linear specification (the correlation is strong and positive in the right tail of the scatterplot). In total, these results suggest that weak first stages in the benchmark model are not driven by the interpolation of viewership that comprises part of GRP measurement.

[33] The number of optimal instruments selected is 1 for 49 categories, 2 for 25 categories, 3 for 11 categories, 4 for one category, and 5 for one category. The linear term $f(P_{mt}) = P_{mt}$ is most likely to be selected: it is the optimal instrument for 19 categories, and numbers among the terms selected for another 31 categories. This result supports our choice of the log-linear specification as the benchmark, and suggests that the crowd-out is stronger when the level of political advertising is high.

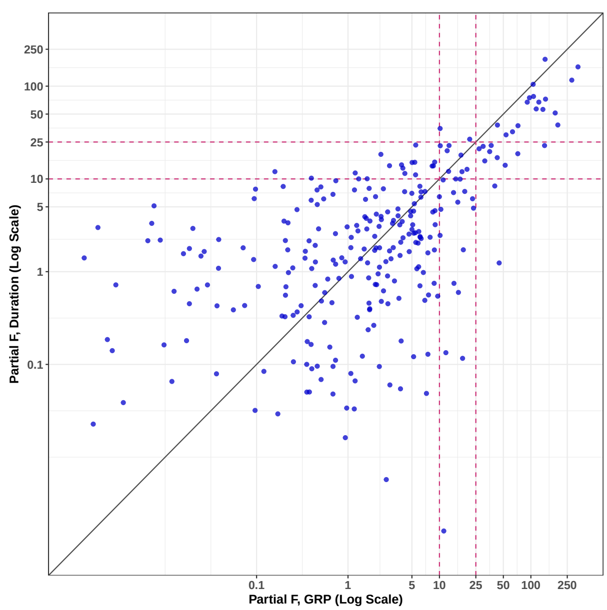
[34] A final reason to show the duration analysis, is some researchers may have access to advertising occurrence data, but not viewership data, and this would provide a guide in such cases.

Table 4: Comparison Across Different Specifications

Measure	GRP				Duration			
LHS	Log		Linear		Log		Linear	
RHS	Linear	Lasso	Linear	Lasso	Linear	Lasso	Linear	Lasso
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Number of Categories in Partial-F Bins								
> 25	28	25	33	32	19	18	19	18
$10 - 25$	25	49	30	56	33	45	40	56
< 10	221	13	211	14	222	10	215	12
NS		187		172		201		188
Panel B: Top 3 Categories for Each Specification								
1	FA (327)	HP (197)	FA (253)	FA (253)	Car (195)	Car (195)	AD (152)	Car (115)
2	AD (280)	FA (163)	AD (209)	AD (209)	FA (161)	FA (118)	Car (115)	AD (87)
3	HP (197)	Car (143)	Ent. (174)	Opt. (130)	AD (116)	LT (105)	FA (97)	LT (72)

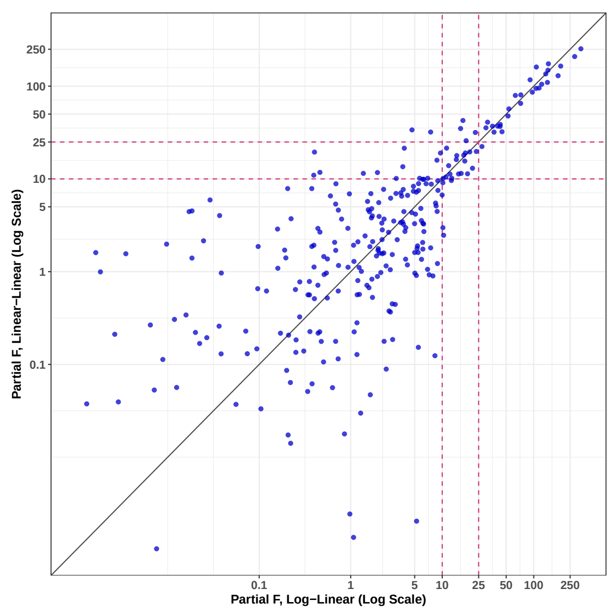
Table Notes: The left hand side is category advertising, and the right hand side is political advertising. Both sides are either measured in GRP or in duration. "NS" means that the Lasso algorithm does not select any of the 40 non-linear transformations. The top 3 categories for each specification are listed in abbreviations, with their partial-F statistics in parentheses. The abbreviations are: "FA" for "Households Furnishings & Appliance Stores", "HP" for "Hospitals, Physicians, & Misc. Physical Culture", "AD" for "Autos & Light Truck: Dealerships" ,"Car" for "Passenger Cars–New & CPO", "LT" for "Light Trucks & Vans–New & CPO", "Ent." for "Misc. Entertainment & Combination Copy", and "Opt." for "Optical Goods and Services".

Figure 9: Partial F: Duration / GRP



Note: This plot compares columns (5) and (1) of Table 4.

Figure 10: Partial F: Linear / Log-Linear



Note: This plot compares columns (3) and (1) of Table 4.

4.3 Category Characteristics Related to First-Stage Strength

In this section, we examine the characteristics associated with first-stage strength. Figure 11 plots the distributions of four characteristics for each F-statistic bin. Panel (A) shows that larger categories tend to have stronger crowd-out effects. Per Section 2.4, these may be categories with lower willingness-to-pay on the margin due to an already built-up advertising stock; they may be on the relatively flat part of the advertising response curve. Panels (B) and (C) show that categories that rely more on spot markets (higher spot GRP shares) and advertise more in high-impression periods (higher GRP per occurrence) tend to have stronger first stages. For the $F > 25$, $10 < F < 25$, and $F < 10$ categories, the median spot GRP share is 0.37, 0.11, and 0.04, respectively, and the median average GRP per occurrence is 2.46, 2.16, and 1.99. Since political ads are primarily bought on spot markets (spot share 0.94) and are very eyeball-heavy (3.54 GRPs per occurrence), the associations between these two characteristics and first-stage strength are intuitive. The categories that advertise primarily locally likely have less flexibility to substitute to national television due to their inherently local nature (e.g. car dealerships, appliance stores, and hospitals). The association with high-impression dayparts may reflect a preference for reach over frequency, as might be the case for durable products. For these types of goods, frequent reminders may be less useful than broadly distributed information. Because there are relatively few ad slots with high reach, these types of categories likely have less flexibility to substitute their ad dollars elsewhere. Finally, Panel (D) plots the share of spot GRPs in the "morning" and "early fringe" (which corresponds to the late afternoon/early evening) dayparts, in which political ads are the most concentrated. The shares in those two dayparts are slightly larger for the $F > 25$ categories (median 0.41 v.s. 0.38), but its association with first-stage strength is weak. This pattern might indicate that some categories substitute to other dayparts, offsetting low willingness-to-pay advertisers in those dayparts indirectly. The first three category characteristics are also listed along with the first-stage results in Tables 3 and 7.

Figure 11: Distributions of Four Category Characteristics, Colored by Partial-F

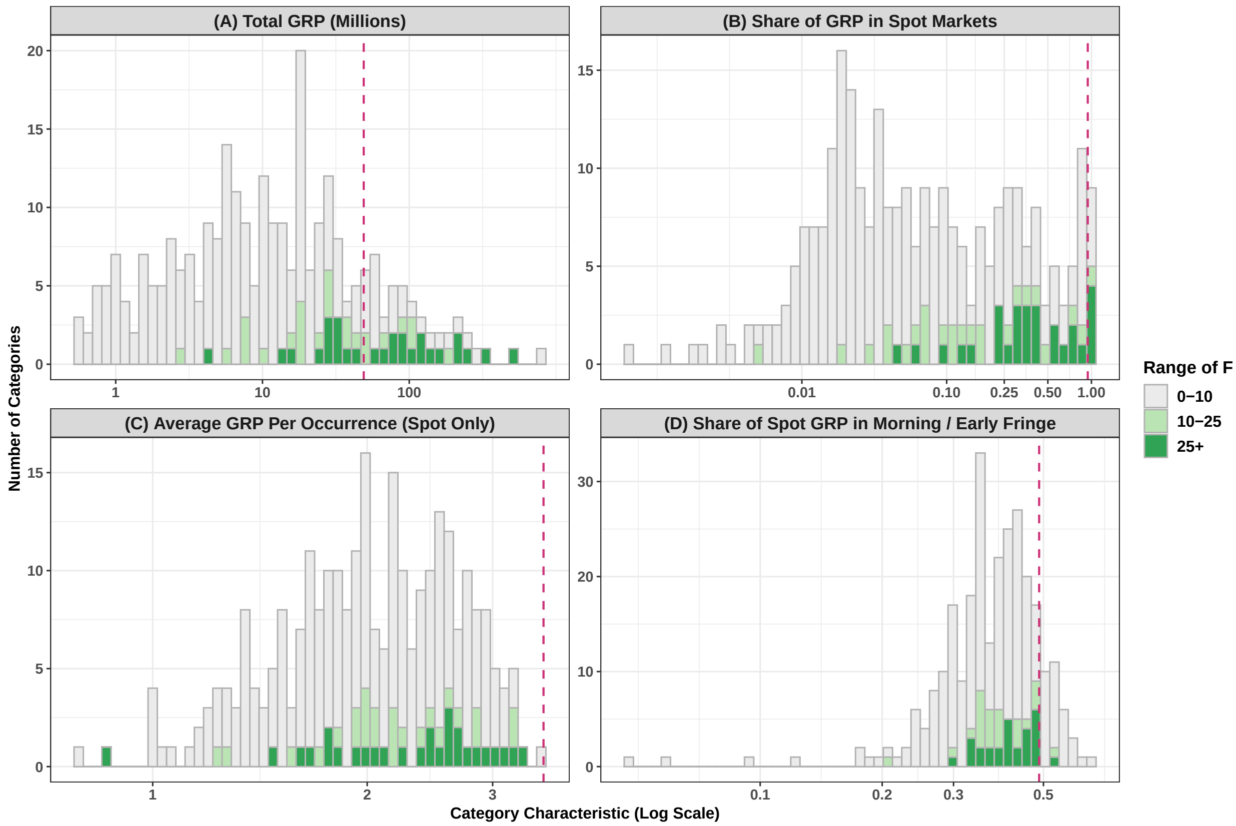


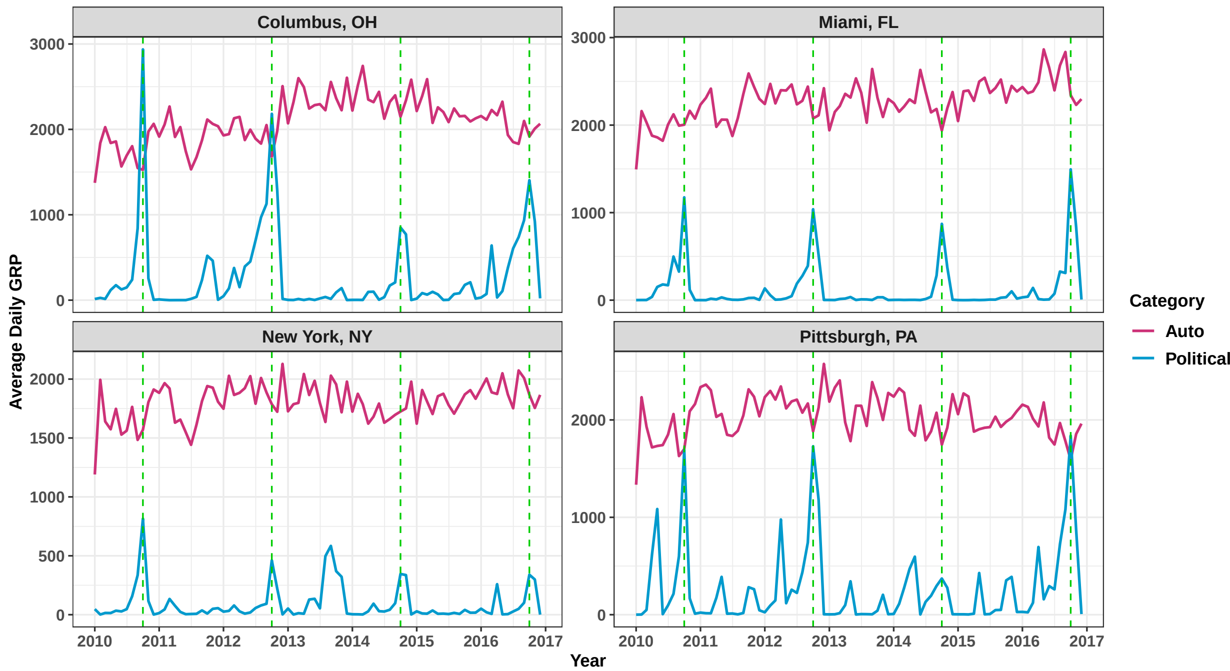
Figure Notes: Each observation is a category. Panels C and D use data from local spot markets only. The vertical dotted lines show the characteristics for political advertising. The "morning" daypart is from 6 to 9 AM, and the "early fringe" daypart is from 4:30 to 7 PM.

5 Proof of Concept: Auto Advertising

We implement the political advertising IV to understand ad effects for motor vehicles. The purpose of this exercise is twofold. First, we want to characterize the properties of the IV estimator conditional on having a sufficiently strong first stage to limit weak instrument bias. Second, our prior is that the short-run category-level auto advertising elasticity is likely zero. That is, we believe auto advertising is unlikely to change an individual's need for a new car (although it may change an individual's choice of car, so that advertising is primarily business-stealing). Thus, we interpret the IV estimate of the short run category level ad effect as a placebo test of the exclusion restriction.

Automotive advertising is large. Dealers and manufacturers spent an estimated $51 billion

Figure 12: Auto and Political GRPs Over Time in 4 Example Markets



worldwide in 2017.[35] Automotive advertising has also been studied in the economics and marketing literature (e.g. Murry [2017]). Anecdotal evidence from the popular press suggests that political advertising is particularly relevant for car dealerships, so we interpret this case study as a near ideal case for the theory of the identification strategy.[36] We sum GRPs across manufacturers, dealers, and dealer associations for motorcycles, light trucks, passenger cars, and vans.[37] Figure 12 shows the co-movement of GRPs aired by political groups and the automotive industry for 4 media markets: Columbus, Miami, New York, and Pittsburgh. It is easy to see the ebb and flow of the political cycle across all media markets, where GRPs spike in the fall of even years. Advertising cycles appear to be high frequency for auto advertisers, but low amplitude. Our sales data comes from RL Polk, and includes the monthly quantity of motor vehicle purchases by buyer ZIP code from 2010-2016.[38]

An OLS regression of sales on GRP that includes month×year fixed effects and ZIP code fixed effects suggests an ad elasticity of 0.108 (Table 6), which is likely spurious due to the usual endo-

[35] https://adage.com/article/cmo-strategy/world-s-largest-advertisers-2017/311484

[36] Tim Higgins. "No One Hates Political Ads More than Car Dealers," Bloomberg News, January 8, 2016 [https://www.bloomberg.com/news/articles/2016-01-08/no-one-hates-political-ads-more-than-car-delaers].

[37] The 7 categories are: T111 "Passenger Cars–Factory, New & CPO", T112 "Light Trucks & Vans–Factory, New & CPO", T115 "Motorcycles and Misc. Vehicles–Factory, New", T121 "Passenger Cars–Dealer Association, New & CPO", T122 "Light Trucks & Vans–Dealer Association, New & CPO", T161 "Autos and Light Truck–Dealerships", and T163 "Motorcycles and Misc. Vehicles–Dealerships".

[38] The data comes from vehicle registrations.

Table 5: Summary Statistics on Auto Sales

	count	mean	sd	min	max
sales	1646659	37.38261	287.9025	0	76646
grps	1681092	55012.33	11941.77	18757.73	108048.8

Summary statistics on automobile purchases from 24608 ZIP codes 2010-2016.

geneity concerns that dealers and manufacturers strategically target advertising at areas with high latent demand for motor vehicles. That association is depicted graphically in Figure 13, panel (a).

To obtain a causal estimate of the effect of category advertising on category sales for automobiles, we turn to the political advertising instrument. Based on our results in Section 4, the auto industry appears amenable to the political ad strategy; six auto categories number among the categories where political advertising has a strong first-stage (see table 3).

Panel (b) of Figure 13 reproduces that first stage relationship graphically. As predicted by our theory, there is a negative relationship between log political GRPs and log auto GRPs. Panel (c) of the figure shows the reduced form relationship between auto sales and political GRPs that appears nearly flat and is not statistically significant. Table 6 displays the corresponding regression results, which include month×year fixed effects and ZIP code fixed effects (standard errors are clustered at the DMA level).[39][40] The preferred specification is in column (4), where we estimate a strong first stage, with a partial F-statistic of 402, which is far above conventional thresholds for limiting weak instrument bias. The IV estimate of the ad elasticity is 0.07 and is not statistically significant (column 4 of table 6). From the perspective of the exclusion restriction, it is reassuring that zero is in the confidence interval, as we expect the true short-run effect of car advertising on category car sales to be zero. However the 95% confidence interval of [-0.024, 0.16] is reasonably wide, and spans a considerable portion of the distribution found in Shapiro et al. [2019]. It also contains point estimates obtained from estimating the model with automobile advertising stock (rather than flows) as the explanatory variable. The results using stock are reported in columns 5 and 6 of table 6, which differ in how political advertising is measured (as a flow and stock, respectively).

[39] We use ZIP code fixed effects instead of DMA fixed-effects due to the fact that we have outcome data at the ZIP code level. ZIP code fixed effects reduce more noise in the dependent variable than do DMA fixed-effects.

[40] Note that we omit price in our regression specification (equation 1) because we aim to estimate the total effect of advertising, which is potentially be mediated through price. That is, if advertising affects pricing, then price is a "bad control" (Angrist and Pischke [2009]). If our goal was to estimate a "partial equilibrium" effect of advertising that held prices fixed, we would instead include price as a covariate, requiring a second instrument for price.

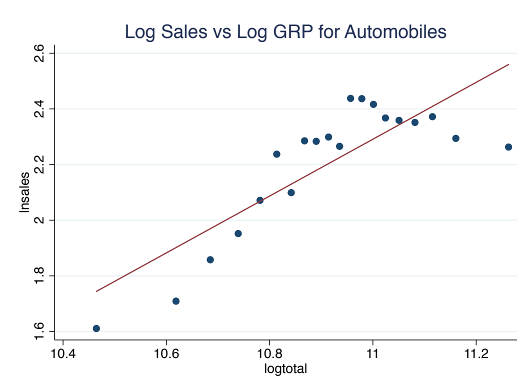
Table 6: Automotive Advertising Effects

	OLS		IV				Optimal IV
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Log(Auto GRP)	0.779***	0.108**	2.463*	0.068			0.054
	(0.163)	(0.048)	(1.447)	(0.047)			(0.049)
Log(Stock Auto GRP)					0.022	0.100	
					(0.077)	(0.083)	
N	1,646,415	1,646,415	1,646,415	1,646,415	1,361,372	1,361,372	1,646,415
Instrument			flow	flow	flow	stock	flow
FE	No	Yes	No	Yes	Yes	Yes	Yes
First-Stage F	.	.	15.05	402.85	136.48	142.37	159.59

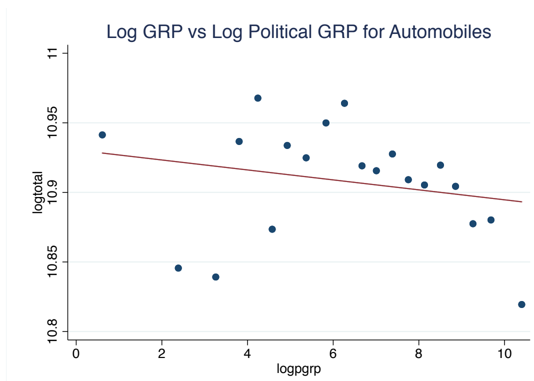
Notes: Standard errors clustered at the DMA level. FE indicate DMA and month $\times$ year fixed effects. *Instrument* refers to whether political GRPs are measured as a flow or stock variable. The first-stage MOP F-statistic is reported in the last row.

Figure 13: Political Advertising and Auto Sales

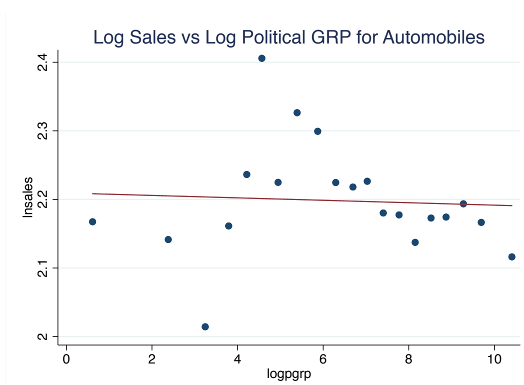
(a) OLS



(b) First Stage



(c) Reduced Form



Notes: All panels of this figure show bin-scatter plots with 20 equi-percentile bins. Zipcode and Year $\times$ Month fixed effects are taken out.

We take a few main lessons from this exercise. First, it demonstrates that a very strong first stage does not guarantee a precise estimate in the two-stage least squares regression. In this case, the dependent variable is noisy, even with a battery of fixed-effects. Individuals do not buy cars frequently, and the exact timing of their purchases is likely driven by many omitted factors. Our estimated confidence interval includes both zero and meaningfully positive advertising elasticities.[41] Second, that we find a statistically significant short-run effect of advertising on category-level car purchases in the OLS (when the true effect is likely zero) reinforces our belief that the endogeneity of advertising in this setting is a first order problem. Further, our confidence in the exclusion restriction for political advertising is incrementally increased by the fact that using the IV puts the true effect (zero) into the confidence interval.

6 Conclusions

This paper investigates when and how the political cycle can be used to estimate the causal effect of commercial advertising on sales. Absent a source of quasi-random variation, observational data is likely to yield biased estimates on the return to commercial advertising. We carefully enumerate the necessary assumptions for political advertising to identify causal effects, and discuss considerations of which LATE it identifies. Using Ad Intel data from 2010-2016, we present descriptive evidence that political advertising increases sharply during election seasons and offsets commercial advertising almost 1-1. We also show that the offset constitutes only a small overall change in the aggregate commercial advertising market. We find that political advertising moves commercial advertising levels (first stage $F > 25$) for 28 out of 274 product categories that advertise on television. We show how to use LASSO to obtain optimal instruments and document that it marginally improves first-stage power for several particularly weak categories.

An important consideration in employing the political advertising instrument is whether the exclusion restriction holds. We highlight two potential threats: first, through the price mechanism, as negative product-market demand shocks might reduce commercial advertising demand, lowering prices for airtime and luring political advertisers; and second, through fluctuations in viewer attentiveness across markets and over time, which might attract both political and commercial

[41] This further highlights that the point of testing for weak instruments is limiting finite sample *bias* and not about ensuring precision in the IV estimate.

advertisers. We argue that integral to the implementation of the political advertising instrument is the inclusion of time and market fixed effects. Using this preferred specification, we estimate the advertising elasticity for the automotive industry using monthly vehicle sales data. While the confidence interval of the IV estimate contains the presumed truth of zero short-run advertising effect on category-level demand of automobiles, the estimate is imprecise.

Given our results, our subjective recommendation to researchers interested in employing political advertising as an instrument is three-fold. First, given that political advertising is a weak instrument in the vast majority of cases, researchers should consider using weak identification robust inference. Second, researchers may want to avoid the instrument in cases where the first stage is the "incorrect" sign because it suggests that the theory underlying the validity of the IV approach is violated. In these cases, political advertising and commercial advertising are related in some way other than crowd out. It could be that there is substitution between markets or positively correlated preferences for particularly attentive viewers occurring in conjunction with crowd-out, violating monotonicity. Alternatively, there could be some path from unobservable demand shocks to both political advertising and commercial advertising, violating the exclusion restriction. While we stop short of recommending complete avoidance of the instrument for the categories where LASSO returns zero terms, it signals to the researcher that the instrument is particularly weak. While Belloni, Chen, Chernozhukov, and Hansen [2012] provide a way forward in such cases, it is unlikely that the political cycle will produce an informative IV. Researchers may wish to find other sources of variation in such cases. Finally, we want to emphasize that the exclusion restriction might be difficult to justify in some cases. Appendix D documents that for 36 categories of grocery products, 5 of them show signs of potential exclusion restriction violations. When employing this instrument, researchers and practitioners should make sure to carefully consider the possible exclusion restriction violations and decide how likely such violations are in their particular case. It should not be taken for granted that the instrument satisfies the exclusion restriction in all cases.

References

Theodore W Anderson and Herman Rubin. Estimation of the parameters of a single equation in a complete system of stochastic equations. *The Annals of Mathematical Statistics*, 20(1):46–63, 1949.

Isaiah Andrews. Conditional linear combination tests for weakly identified models. *Econometrica*, 84(6):2155–2182, 2016.

Joshua D Angrist and Guido W Imbens. Two-stage least squares estimation of average causal effects in models with variable treatment intensity. *Journal of the American statistical Association*, 90 (430):431–442, 1995.

Joshua D Angrist and Jorn-Steffen Pischke. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton University Press, 2009.

Timothy J Bartik. Who benefits from state and local economic development policies? *WE Upjohn Institute for Employment Research*, 1991.

Alexandre Belloni, Daniel Chen, Victor Chernozhukov, and Christian Hansen. Sparse models and methods for optimal instruments with an application to eminent domain. *Econometrica*, 80(6): 2369–2429, 2012.

Steven T. Berry. Estimating discrete-choice models of product differentiation. *The RAND Journal of Economics*, 25(2):242–262, 1994.

Kirill Borusyak, Peter Hull, and Xavier Jaravel. Quasi-experimental shift-share research designs. *arXiv preprint arXiv:1806.01221*, 2018.

Pierre Dubois, Rachel Griffith, and Martin O'Connell. The effects of banning advertising in junk food markets. *The Review of Economic Studies*, 85(1):396–436, 2018.

Amit Gandhi and Jean-François Houde. Measuring substitution patterns in differentiated products industries. 2017.

Paul Goldsmith-Pinkham, Isaac Sorkin, and Henry Swift. Bartik instruments: What, when, why, and how. *NBER Working Paper 24408*, 2018.

Wesley R Hartmann and Daniel Klapper. Super bowl ads. *Marketing Science*, 37(1):78–96, 2018.

Xing Li, Wesley R Hartmann, and Tomomichi Amano. Preference externality estimators: A comparison of border approaches and ivs. *Available at SSRN 3402187*, 2019.

Mitchell J. Lovett, Renana Peres, and Linli Xu. Can your advertising really buy earned impressions? the effect of brand advertising on word of mouth. *Quantitative Marketing and Economics*, Apr 2019. ISSN 1573-711X. doi: 10.1007/s11129-019-09211-9. URL `https://doi.org/10.1007/s11129-019-09211-9`.

Sarah Moshary. Price discrimination in political advertising: Evidence from 2012 us presidential election. *The RAND Journal of Economics*, Forthcoming.

Charles Murry. Advertising in vertical relationships: An equilibrium model of the automobile industry. *Working Paper.*, 2017.

José Luis Montiel Olea and Carolin Pflueger. A robust test for weak instruments. *Journal of Business & Economic Statistics*, 31(3):358–369, 2013.

Peter E Rossi. Even the rich can make themselves poor: A critical examination of iv methods in marketing applications. *Marketing Science*, 33(5):655–672, 2014.

Bradley Shapiro, Guenter J. Hitsch, and Anna Tuchman. Generalizable and robust tv advertising effects. *Working Paper, Available at SSRN: https://ssrn.com/abstract=3273476*, 2019.

Bradley T. Shapiro. Positive spillovers and free riding in advertising of prescription pharmaceuticals: The case of antidepressants. *Journal of Political Economy*, 126(1):381–437, 2018. doi: 10.1086/695475. URL `https://doi.org/10.1086/695475`.

Michael Sinkinson and Amanda Starc. Ask your doctor? Direct-to-consumer advertising of pharmaceuticals. *Review of Economic Studies, forthcoming*, 2018.

James Stock and Motohiro Yogo. *Testing for Weak Instruments in Linear IV Regression*, pages 80–108. Cambridge University Press, New York, 2005. URL `http://www.economics.harvard.edu/faculty/stock/files/TestingWeakInstr_Stock%2BYogo.pdf`.

Michael Thomas. Mass advertising as a natural experiment. *Available at SSRN: https://ssrn.com/abstract=3182092*, 2017.

Anna E Tuchman. Advertising and demand for addictive goods: The effects of e-cigarette advertising. *Marketing Science*, 38(6):994–1022, 2019.

A Additional Tables

Table 7: List of 221 Categories with Weak Log-Linear First Stage ($F < 10$)

Category	Partial F	Partial R^2	Scaled Coef	Size Rank	Spot Share	GRP/OCC
Cosmetics, Misc. Beauty Aids & Supp	10.0	0.002	-0.096	244	0.08	3.2
Radio Stations	9.6	0.002	-0.223	193	0.91	1.9
City, State & Foreign Governments	9.0	0.001	-0.100	141	0.86	2.1
TV Program-Multi-News	8.9	0.001	-0.054	48	0.96	2.5
TV Program-Cable-Talk Show	8.9	0.000	-0.018	273	0.01	1.5
Non-Carbonated	8.8	0.001	-0.012	239	0.03	1.6
Food Wraps, Foils & Bags	8.8	0.001	-0.008	125	0.03	1.7
Garden Machinery, Implements & Fixtures	8.6	0.001	-0.068	158	0.25	2.8
Accting, Auditing & Bookkeeping Services	8.5	0.002	-0.024	100	0.10	1.6
Pasta Products & Pasta Product Dinners	8.3	0.000	-0.046	218	0.06	2.8
Pre-Recorded Video	7.9	0.001	0.003	40	0.01	1.7
Telecommunication Systems & Services	7.6	0.001	-0.050	151	0.22	2.2
Car Batteries	7.5	0.001	-0.023	258	0.01	2.8
Fitness & Diet Programs & Spas	7.5	0.001	-0.024	60	0.18	2.3
TV Program-Evening-Ent	7.2	0.001	-0.136	187	0.32	1.5
All Other Prepared Dinners & Entrees & Pizzas	6.9	0.001	-0.007	32	0.06	1.9
Motion Pictures	6.9	0.001	0.004	5	0.02	2.0
Computers, Components & Accessories	6.7	0.001	0.008	59	0.01	2.4
Home Insurance	6.3	0.001	-0.023	111	0.09	2.0
Telephone Companies, Public & Private	6.3	0.001	-0.010	14	0.13	2.2
TV Program-Prime-Sports	6.3	0.001	-0.056	223	0.35	1.7
Garden Pest Controls	6.1	0.001	-0.081	169	0.22	3.1
Regular Carbonated	6.1	0.001	-0.011	93	0.04	2.2
Security Systems & Fire Safety Systems	6.1	0.001	-0.051	171	0.19	2.3
Sport & Protective Footwear	6.1	0.001	-0.011	118	0.01	1.5
Soups	5.9	0.000	-0.021	67	0.05	1.9
Motorcycles & Vhcls Misc-Factory:new	5.9	0.000	-0.012	164	0.03	2.2
Pre-Fabricated Homes & Buildings	5.8	0.001	-0.192	271	0.81	2.7
Books & Music	5.7	0.001	-0.015	219	0.05	1.5
TV Program-Synd-Ent	5.5	0.001	-0.025	38	0.88	1.3
Combination Copy & Misc Dairy Products	5.5	0.001	-0.009	44	0.06	1.9
Petroleum Companies, General Promotion	5.5	0.000	-0.035	201	0.05	3.2
Cookies & Crackers	5.4	0.002	0.008	74	0.01	1.6
Jellies, Jams, Preserves, Peanut Butter	5.4	0.001	-0.010	134	0.05	2.2
Mattresses	5.3	0.000	-0.048	135	0.07	2.3
Blouses & Shirts	5.2	0.000	-0.016	255	0.00	2.9
Paint, Varnishes & Misc Coatings	5.2	0.001	0.032	115	0.03	2.4
Miscellaneous Medication & Proprietary Remedies	5.2	0.001	-0.007	82	0.04	1.6
TV Program-Multi-Ent	5.1	0.001	-0.044	160	0.49	1.6
Refrigerators & Freezers	5.0	0.000	-0.018	198	0.01	2.0
TV Program-Synd-Talk Show	5.0	0.001	-0.039	96	0.91	1.7
Car Cleaners, Waxes & Polishes	5.0	0.001	-0.013	231	0.03	1.2
Pet Food (not Elsewhere Classified)	4.8	0.001	-0.045	265	0.06	2.8
TV Program-Morning-News	4.8	0.001	-0.101	170	0.99	2.1
Stationery & Miscellaneous Paper Goods	4.7	0.001	-0.018	192	0.02	1.7
Sauces, Gravies, Dips	4.6	0.000	-0.010	127	0.02	1.9
General Household Small Appliances(not Else Class)	4.2	0.000	0.007	130	0.01	1.8
Farm Machinery & Equipment	4.2	0.001	-0.054	245	0.43	2.7
Shampoos, Conditioners & Cream Rinses	4.0	0.000	-0.005	52	0.02	1.3
TV Program-Daytime-Sports	4.0	0.000	-0.036	213	0.33	1.6

Continued on next page

Table 7 – *Continued from previous page*

Category	Partial F	Partial R^2	Scaled Coef	Size Rank	Spot Share	GRP/OCC
Women's Scents & Fragrances	3.9	0.000	0.011	143	0.03	2.9
Mobile Homes,campers,rvs:dealerships	3.8	0.000	-0.124	204	0.96	2.7
Mutual Funds & Money Funds	3.8	0.000	0.007	194	0.01	2.9
TV Program-Prime-Ent	3.8	0.001	0.010	4	0.19	2.2
Toys	3.7	0.001	0.001	24	0.00	1.9
Home Audio Equipment & Accessories	3.7	0.000	0.023	209	0.01	2.8
Pre-Recorded Audio	3.7	0.000	0.008	185	0.04	1.2
TV Program-Daytime-Ent	3.6	0.001	-0.018	109	0.24	1.5
Shoes & Boots, Regular & Casual	3.5	0.000	-0.012	157	0.01	1.8
Disinfectants	3.5	0.001	-0.007	144	0.04	1.4
Electronics Stores (Incl Rental)	3.2	0.000	-0.018	110	0.10	2.2
Resort Promotion (Area)	3.1	0.001	-0.028	92	0.40	2.5
Horticultural Services & Misc	3.1	0.001	-0.116	199	0.28	2.3
Baking Mixes, Crusts & Baking Ingred	3.1	0.001	-0.025	189	0.06	1.5
Hair Dressings, Sprays & Restoration Prod	3.0	0.000	-0.006	147	0.01	1.3
Car & Truck Rental Svcs	2.9	0.000	0.018	161	0.01	1.3
TV Program-Cable-News	2.8	0.000	-0.001	112	0.00	1.6
Credit Cards	2.8	0.000	-0.004	30	0.03	1.8
Investments & Collectns (not Else Class)	2.7	0.000	-0.043	257	0.30	2.0
Diapers(Incl Infant & Adult)	2.7	0.000	-0.005	81	0.02	1.6
Small Kitchen Appliances & Equipment	2.7	0.000	-0.007	133	0.03	1.9
Home Phone Equipment	2.6	0.000	0.014	128	0.04	2.5
Bedding & Linens	2.6	0.000	-0.019	232	0.12	1.2
Gasoline Dealrs Services & Promotions	2.4	0.000	-0.038	208	0.39	3.0
Misc Toilet Goods	2.4	0.000	0.005	62	0.02	2.0
Personal Care/Grooming Appliances	2.3	0.000	0.007	207	0.02	1.2
Laxatives	2.3	0.000	-0.005	114	0.02	2.4
Miscellaneous Sporting Equipment	2.3	0.000	-0.009	261	0.04	1.0
Toilet Soaps	2.3	0.000	-0.005	99	0.02	1.5
TV Program-Daytime-Talk Show	2.2	0.000	0.056	196	0.08	3.0
Misc Corporate Advertising (not Else Class)	2.2	0.000	-0.012	119	0.10	2.4
US Government	2.2	0.000	-0.008	63	0.15	1.0
Gelatins & Puddings (Mixes & Prepared)	2.2	0.000	-0.012	227	0.01	2.0
Department Stores	2.1	0.000	-0.005	9	0.11	2.9
Boats, Motors & Accessories	2.1	0.000	-0.051	267	0.70	2.6
Jewelry	2.0	0.000	-0.056	241	0.26	3.0
TV Program-Overnite-News	2.0	0.000	-0.119	174	1.00	2.8
Cable Tv Network	2.0	0.000	-0.001	17	0.01	1.3
Misc Equipment, Fixtures & Systems	2.0	0.000	-0.038	243	0.29	2.2
Sunglasses, Frames & Corrective Lenses	2.0	0.000	-0.005	122	0.02	2.1
Laundry Detergents & Cleaning Preparations	2.0	0.000	-0.004	39	0.02	1.6
Gasoline & Oil	1.9	0.000	0.009	149	0.05	2.0
Other Fruits	1.7	0.000	-0.015	225	0.09	2.0
Misc Men's Toiletries & Hygienic Goods	1.7	0.000	0.002	217	0.00	0.8
Motorcycle & Vehicle Misc-Dealerships	1.7	0.000	-0.098	252	0.93	2.5
Dairy Product Substitutes	1.7	0.000	-0.009	168	0.04	2.3
Sponges	1.7	0.000	-0.003	75	0.02	1.8
Misc Building Materials	1.7	0.000	-0.029	270	0.30	2.0
Non-Computerized Games	1.7	0.000	0.002	86	0.00	1.9
Infants Foods	1.6	0.000	-0.007	190	0.02	1.5
Combination Copy Food/Food Products/Gp	1.6	0.000	-0.017	154	0.07	2.3
Computer & Data Processing Services	1.6	0.000	-0.025	264	0.20	1.8
Pharmaceutical Houses, Gp	1.6	0.000	-0.003	10	0.03	2.5
Steamship, Truck & Other Facilities	1.6	0.000	-0.008	120	0.10	2.4
Television Stations	1.5	0.000	0.022	237	0.13	2.3

Continued on next page

Table 7 – *Continued from previous page*

Category	Partial F	Partial R^2	Scaled Coef	Size Rank	Spot Share	GRP/OCC
Liquor	1.5	0.000	0.001	64	0.01	1.4
Car & Truck Tires & Tubes	1.4	0.000	-0.004	136	0.02	2.3
Hardware & Lumber Stores	1.4	0.000	-0.023	152	0.12	2.6
Feminine Hygiene Products	1.3	0.000	-0.003	84	0.02	1.1
Shaving Equipment & Supplies	1.3	0.000	0.003	91	0.01	1.7
TV Program-Multi-Sports	1.3	0.000	-0.012	102	0.29	2.1
Doors & Windows	1.2	0.000	-0.072	266	0.73	2.5
Candy & Gum	1.2	0.000	0.002	12	0.01	1.5
Deodorizers, Air Fresheners & Purifiers	1.2	0.000	-0.002	68	0.02	1.4
Floor & Furniture Polish & Wax	1.2	0.000	-0.009	205	0.02	1.7
Dental Supplies & Mouthwashes	1.2	0.000	-0.003	23	0.03	1.7
Insecticides	1.2	0.000	-0.019	226	0.08	2.3
Stoves & Ranges (Including Microwave)	1.2	0.000	-0.019	256	0.03	2.7
Railroad Travel	1.1	0.000	-0.028	249	0.16	2.7
Salad Dressings & Mayonnaise	1.1	0.000	0.006	129	0.08	2.0
Facial Cleansers & Make-Up Removers	1.1	0.000	-0.005	139	0.02	1.4
Shortening, Oil, Margar & No-Stick Prod	1.1	0.000	0.006	165	0.03	1.9
Vegetable Juices	1.0	0.000	0.004	177	0.01	2.6
Patio & Barbecue Equip & Accessor	1.0	0.000	-0.011	229	0.03	1.5
Hosiery	0.9	0.000	-0.009	259	0.01	2.6
Gas, Power, Lighting & Water Companies	0.9	0.000	-0.062	167	0.81	3.1
Online Tv Program-Promo	0.9	0.000	0.006	268	0.02	1.3
Video Equipment & Accessories	0.8	0.000	0.009	180	0.02	3.0
TV Program-Cable-Sports	0.7	0.000	0.001	41	0.01	2.3
Misc Household Furnishings	0.7	0.000	0.012	230	0.02	1.6
Employment Recruitment	0.7	0.000	-0.039	221	0.41	1.9
TV Program-Evening-Sports	0.7	0.000	-0.027	224	0.54	1.8
Beans & Grains	0.7	0.000	0.027	179	0.05	2.2
Furniture	0.7	0.000	-0.035	220	0.20	2.8
Computer Software	0.7	0.000	0.002	50	0.01	2.1
Beer	0.7	0.000	0.002	43	0.02	1.8
Heating & Cooling Equip, Fixtrs & Systems	0.7	0.000	0.023	210	0.36	2.8
Combination Copy & Misc Prepared Food	0.6	0.000	0.004	236	0.02	2.5
Reducing Aids	0.6	0.000	-0.003	131	0.02	1.2
Condiments, Pickles, Relishes	0.6	0.000	-0.034	212	0.08	1.9
Telex, Telemail & Telegram Systems	0.6	0.000	0.010	238	0.04	1.0
Cameras and Photographic Supplies	0.5	0.000	0.006	183	0.02	2.3
Meats, Poultry & Fish	0.5	0.000	-0.004	80	0.12	2.8
Pets & Pet Supplies	0.5	0.000	-0.004	103	0.08	2.0
TV Program-Morning-Ent	0.5	0.000	-0.005	94	0.25	2.2
Vacuum Cleaners & Dishwashers	0.5	0.000	-0.002	126	0.02	1.8
Misc Medical & First-Aid Supplies	0.5	0.000	0.005	188	0.02	1.3
Comb Copy & Misc Commun & Public Utilit	0.5	0.000	-0.021	247	0.18	3.2
Depilatories, Face & Body Hair Bleaches	0.5	0.000	-0.004	234	0.02	1.2
Watches	0.4	0.000	-0.008	211	0.07	2.8
Comb Copy & Misc Elec Entertnmnt Equip	0.4	0.000	0.005	121	0.01	2.0
Men's Scents & Fragrances	0.4	0.000	-0.007	181	0.02	2.5
Bottled Waters	0.4	0.000	-0.014	178	0.07	1.8
Sugars, Syrups & Artificial Sweeteners	0.4	0.000	-0.005	191	0.02	2.0
Medicated Skin Products & Liniments	0.4	0.000	-0.003	90	0.05	1.5
In-Home Tests (Including Computerized)	0.4	0.000	-0.004	182	0.07	1.3
Vegetables	0.4	0.000	-0.005	140	0.10	2.0
Shoe Care Products	0.4	0.000	-0.002	145	0.02	2.6
Sunscreens & Tanning Products	0.4	0.000	-0.005	172	0.02	1.7
Apparel Fabrics & Finishes	0.4	0.000	-0.003	248	0.00	1.8

Continued on next page

Table 7 – *Continued from previous page*

Category	Partial F	Partial R^2	Scaled Coef	Size Rank	Spot Share	GRP/OCC
Other Soaps & Detergents	0.4	0.000	0.001	98	0.02	1.4
Laundry Equipment	0.4	0.000	0.004	186	0.02	2.6
Window Coverings & Trtmnts & Hh Fabrics	0.4	0.000	-0.015	272	0.09	2.2
Misc Clnrs, Waxes & Polishes (Spec Uses)	0.3	0.000	-0.002	89	0.04	1.2
Skin Care Creams, Lotions & Oils	0.3	0.000	0.001	29	0.03	1.6
Anti-Freeze & Additives	0.3	0.000	0.005	253	0.05	1.7
Facial Make-Up Products	0.3	0.000	0.002	88	0.02	2.0
Hispanic Tv Network	0.3	0.000	-0.012	242	0.21	1.1
Home Workshop Hand & Power Tools	0.3	0.000	-0.003	163	0.02	1.7
Miscellaneous Laundry Supplies	0.3	0.000	0.008	206	0.02	1.5
Auto & Truck Parts,accessories & Kits	0.2	0.000	0.005	173	0.09	2.6
Underclothing	0.2	0.000	-0.004	184	0.00	1.9
Vinyl Flrng & Misc Hard Surf Flr Covrgs	0.2	0.000	-0.006	262	0.05	1.4
Fruit Juices & Fruit Flavored Drinks(Incl Powdrd)	0.2	0.000	0.001	37	0.02	2.0
Exercise Equipment	0.2	0.000	0.001	215	0.01	1.4
Airline Travel	0.2	0.000	0.010	153	0.23	3.2
Hair Coloring & Curling Formulae	0.2	0.000	0.001	73	0.02	1.8
Airline Freight	0.2	0.000	0.003	214	0.04	3.0
Seasonings, Spices & Extracts & Marinades	0.2	0.000	-0.007	166	0.03	2.0
Cat Food	0.2	0.000	-0.001	106	0.01	1.8
TV Program-Daytime-News	0.2	0.000	-0.024	235	0.92	2.2
Cruise Ship Travel	0.2	0.000	0.004	148	0.07	2.4
Dog Food	0.2	0.000	-0.001	53	0.01	1.7
Auto & Light Trucks-Factory:corp Advrtsng	0.2	0.000	-0.009	260	0.10	1.9
Brokerage Svcs	0.1	0.000	-0.001	36	0.03	2.5
Manicure Preparations & Implements	0.1	0.000	0.004	197	0.02	1.4
Combination Copy Ready-to-Wear	0.1	0.000	-0.001	132	0.01	1.8
Dietary Carbonated	0.1	0.000	-0.002	150	0.04	2.5
Vitamin Preparations & Tonics	0.1	0.000	-0.001	22	0.02	1.7
TV Program-Prime-News	0.1	0.000	-0.006	156	0.52	1.8
Pain Relievers, Sedatives & Sleeping Preps	0.1	0.000	0.001	34	0.02	1.4
Digestive Aids & Antacids	0.1	0.000	-0.001	54	0.02	1.6
Lip Make-Up	0.1	0.000	0.001	138	0.01	1.3
Cleaners & Cleansers (Genl Hshold Use)	0.0	0.000	-0.001	87	0.03	1.2
Craft, Hobby & Sporting Goods & Toy Stores	0.0	0.000	0.002	66	0.17	2.6
Citrus Fruits	0.0	0.000	-0.001	240	0.01	3.5
TV Program-Evening-News	0.0	0.000	-0.003	108	0.65	2.4
Typewriters & Misc Office Equipment	0.0	0.000	-0.005	250	0.01	2.4
Miscellaneous Passenger Travel	0.0	0.000	-0.015	254	0.64	2.3
Fertilizers & Seed Treatments	0.0	0.000	0.007	222	0.17	2.4
Household Paper Products	0.0	0.000	-0.000	105	0.02	1.9
Combination Copy & Misc Hygienic Goods	0.0	0.000	-0.002	274	0.04	2.6
TV Program-Latenite-Sports	0.0	0.000	-0.012	251	0.81	2.1
Wine	0.0	0.000	-0.001	200	0.02	2.2
Sportswear	0.0	0.000	-0.001	202	0.01	2.2
Eye Make-Up	0.0	0.000	0.000	79	0.01	1.4
TV Program-Cable-Ent	0.0	0.000	0.000	1	0.00	1.4
Ice Cream, Frozen Novelties & Sherbet	0.0	0.000	-0.002	137	0.11	2.6
Travel Services & Tours	0.0	0.000	-0.000	47	0.03	2.3
Deodorants & Anti-Perspirants	0.0	0.000	0.000	104	0.01	1.0
Miscellaneous Pdts/Svcs	0.0	0.000	-0.001	124	0.46	1.8
Hobbycraft	0.0	0.000	0.000	228	0.02	2.5
Pens, Pencils, Markers & Miscellaneous	0.0	0.000	0.001	195	0.01	2.2
Otr Manuf Materials & Supplies	0.0	0.000	0.001	176	0.11	2.5
Cookware & Cutlery	0.0	0.000	-0.001	263	0.00	1.3

Continued on next page

Table 7 – *Continued from previous page*

Category	Partial F	Partial R^2	Scaled Coef	Size Rank	Spot Share	GRP/OCC
Broadcast Tv Network	0.0	0.000	-0.001	146	0.25	1.8
TV Program-Latenite-Ent	0.0	0.000	-0.000	117	0.23	3.0
Lawn & Garden Seeds, Bulbs & Nursery Stock	0.0	0.000	-0.002	246	0.57	2.9
Agricultural Services & Miscellaneous	0.0	0.000	-0.002	269	0.28	2.8
Business Propositions	0.0	0.000	-0.001	233	0.05	1.0
Milk, Butter, Eggs (Including Powdered)	0.0	0.000	-0.000	83	0.05	2.5

Table 8: List of 87 Categories with Non-Empty LASSO Selection

Category	Partial F		Partial R^2		# Terms
	LASSO	Log-Linear	LASSO	Log-Linear	Selected
Hospitals, Physicians & Misc. Physical Culture	197.2	197.2	0.031	0.031	1
Household Furnishings & Appliance Stores	162.9	327.5	0.059	0.058	2
Passenger Cars-Factory:New & CPO	143.0	143.0	0.026	0.026	1
Construction, Engineering & Archit Srvcs	122.2	122.2	0.018	0.018	1
Autos & Light Truck - Dealerships	109.0	280.2	0.036	0.035	2
Light Trucks & Vans-Factory:New & CPO	105.9	105.9	0.024	0.024	1
Miscellaneous Professional Services	100.5	141.2	0.029	0.028	2
Dance, Theater, Concerts, Opera	91.3	91.3	0.014	0.014	1
Misc Entertainment & Combination Copy	80.1	144.6	0.020	0.020	2
Automotive	72.2	72.2	0.011	0.011	1
Miscellaneous Organization Advertising	68.9	96.5	0.028	0.028	2
Optical Goods and Services	64.6	184.6	0.021	0.018	3
Data Communications Networks	53.5	53.5	0.010	0.010	1
Misc Financial Inst. Services & Products	52.4	52.4	0.009	0.009	1
Miscellaneous Retail	51.1	114.4	0.025	0.023	3
Restaurants, Hotel Dining & Nightclubs	50.8	135.4	0.029	0.026	5
Passenger Cars-Dealer Assn:New & CPO	49.4	106.8	0.019	0.018	3
Cakes, Pies, Pastries & Donuts	43.1	43.1	0.005	0.005	1
Direct Response Products	40.3	40.3	0.011	0.011	1
Apparel, Footwear & Accessory Stores	36.8	36.8	0.004	0.004	1
Schools & Camps, Seminars	35.8	71.9	0.015	0.014	2
Light Trucks & Vans-Dealer Assn:New & CPO	34.8	63.0	0.011	0.011	2
Tv Station	29.5	35.4	0.006	0.006	2
Real Estate, R.e. Brokers & Developers	27.2	27.2	0.004	0.004	1
Banks	25.3	45.0	0.006	0.006	2
Fix-it Supplies	24.4	31.3	0.002	0.002	2
Medical Appliances & Equipment	22.1	14.3	0.002	0.001	1
Medicated Skin Products & Liniments	19.5	0.4	0.002	0.000	1
Cereals	19.4	30.1	0.009	0.007	3
Laxatives	19.3	2.3	0.003	0.000	1
Non-Computerized Games	19.2	1.7	0.002	0.000	1
Cable Television Stations	18.8	42.9	0.008	0.007	3
TV Program-Synd-Talk Show	18.8	5.0	0.003	0.001	1
Religious, Charitable & Humanitarian	18.2	16.1	0.005	0.003	2
all Other Prepared Dinners & Entrees & Pizzas	17.4	6.9	0.002	0.001	1
Hotels & Resorts	17.2	17.2	0.003	0.003	1
Computers, Components & Accessories	16.5	6.7	0.002	0.001	1
Jewelry, Gift Stores & Galleries	16.4	11.2	0.002	0.001	1
Home Audio Equipment & Accessories	16.4	3.7	0.000	0.000	1
Cookies & Crackers	16.1	5.4	0.004	0.002	2
Lotteries	15.9	15.9	0.002	0.002	1

Continued on next page

Table 8 – *Continued from previous page*

Category	Partial F		Partial R^2		# Terms
	LASSO	Log-Linear	LASSO	Log-Linear	Selected
Plumbing & Sanit Equip, Fixtrs & Systems	15.9	23.6	0.003	0.003	2
Vitamin Preparations & Tonics	15.5	0.1	0.002	0.000	1
Automobile Insurance	15.4	20.0	0.004	0.004	2
Combination Copy & Misc Dairy Products	14.9	5.5	0.001	0.001	1
Life Insurance	14.8	18.3	0.004	0.004	2
City,state & Foreign Governments	14.6	9.0	0.001	0.001	1
TV Program-Prime-Sports	14.6	6.3	0.002	0.001	1
Sport & Protective Footwear	14.4	6.1	0.002	0.001	1
Other Insurance & Combination Copy	14.4	14.4	0.003	0.003	1
Regular Carbonated	14.2	6.1	0.004	0.001	3
Golf Equipment	13.9	17.9	0.004	0.004	2
Magazines, Newspapers, Misc Media	13.8	15.1	0.002	0.002	2
Cheese Products	13.3	10.2	0.002	0.002	1
Pasta Products & Pasta Product Dinners	13.2	8.3	0.000	0.000	1
Women's Scents & Fragrances	13.1	3.9	0.001	0.000	1
Beer	13.1	0.7	0.005	0.000	3
Disinfectants	12.9	3.5	0.002	0.001	1
Pet Food (not Elsewhere Classified)	12.8	4.8	0.003	0.001	2
TV Program-Multi-News	12.8	8.9	0.002	0.001	1
Misc Accessories, Supplies & Hardware	12.6	12.6	0.001	0.001	1
Stationery & Miscellaneous Paper Goods	12.4	4.7	0.001	0.001	1
Cold, Cough & Sinus Remedies	12.4	12.7	0.002	0.001	2
Home Insurance	12.3	6.3	0.001	0.001	1
Sauces, Gravies, Dips	12.1	4.6	0.001	0.000	1
Appetizers, Snacks & Nuts	11.9	21.4	0.005	0.003	4
Cat Food	11.6	0.2	0.002	0.000	1
Vacuum Cleaners & Dishwashers	11.6	0.5	0.002	0.000	1
Shaving Equipment & Supplies	11.3	1.3	0.001	0.000	1
TV Program-Latenite-News	11.0	11.0	0.002	0.002	1
Fitness & Diet Programs & Spas	10.4	7.5	0.001	0.001	2
Coffee, Tea, Cocoa & Derivatives	10.2	12.2	0.002	0.002	2
Computerized Games & Accessories Includ Software	10.2	10.2	0.001	0.001	1
Comb Copy & Misc Commun & Public Utilit	10.1	0.5	0.001	0.000	1
Cosmtcs Comb Cpy,misc Beauty Aids & Supp	10.0	10.0	0.002	0.002	1
Toilet Soaps	10.0	2.3	0.001	0.000	1
Medical & Dental Insurance	9.7	18.9	0.002	0.002	3
Liquor	9.7	1.5	0.001	0.000	1
Breads, Rolls, Waffles & Pancakes	9.3	17.6	0.001	0.001	2
Accting, Auditing & Bookkeeping Services	8.7	8.5	0.002	0.002	2
Food & Liquor Stores	8.7	16.8	0.004	0.003	3
Amusement Parks & Sporting Events	8.1	23.0	0.002	0.002	3
Apparel Fabrics & Finishes	7.9	0.4	0.001	0.000	1
Drugs, Toiletries & Salons	7.9	10.2	0.002	0.001	2
Pre-Fabricated Homes & Buildings	7.9	5.8	0.002	0.001	1
Cellular Radio Systems & Accessories	7.8	11.8	0.003	0.003	2
Radio Stations	7.3	9.6	0.003	0.002	3

B Selecting the Optimal Instruments

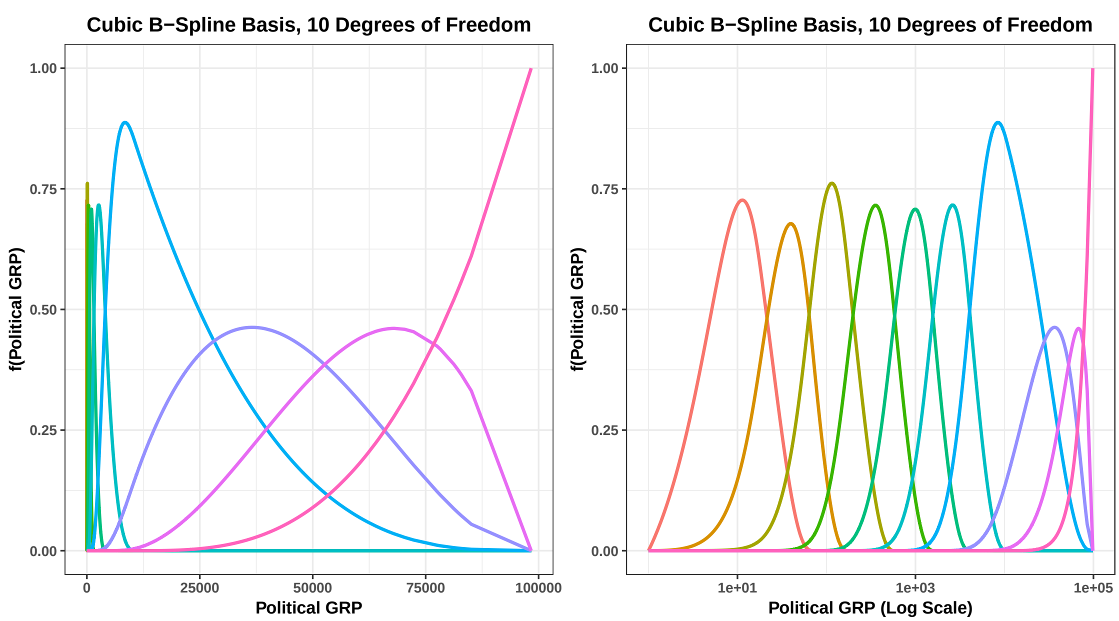
B.1 Candidates of Optimal Instruments Selection

We use the following four sets of non-linear transformations for political GRP as candidates for optimal instruments:

- A 5th degree orthogonal polynomial. [5 Terms]
- A 5th degree orthogonal polynomial of log(1+political GRP). [5 Terms]
- A cubic B-spline basis with 10 degrees of freedom. The knots are placed on equally spaced quantiles of political GRP. [10 Terms]
- A cubic B-spline basis with 20 degrees of freedom. [20 Terms]

While each set of candidates could capture any non-linearity in the data fairly well on its own, we use all of them together to account for the fact that we have no *a priori* knowledge about which set fits the data best. In addition, our broad set of candidates allows the optimal instrument to be some combination of polynomial and spline terms. Finally, we include the polynomial of log(1+political GRP) to improve the fit for a potential constant-elasticity relationship. Figure 14 illustrates the cubic spline basis with 10 terms by plotting them against both political GRP and log(1+political GRP).

Figure 14: Illustration: Cubic B-Spline Basis for Political GRP



B.2 The LASSO Method from Belloni et al (2012)

Given a first-stage regression equation $d_i = D(x_i) + v_i$ where d_i is the endogenous variable, $x_i \in \mathbb{R}^p$ are a vector of instruments, and $D(\cdot)$ is an unknown function, Belloni, Chen, Chernozhukov, and Hansen [2012] approximate the optimal instrument by $\hat{D}(x_i) = x_i'\hat{\beta}$, where $\hat{\beta}$ is the solution of a LASSO optimization problem:

$$\hat{\beta} \in \operatorname{argmin}_{b \in \mathbb{R}^p} \frac{1}{n} \sum_{i=1}^{n} (d_i - x_i'b)^2 + \frac{\lambda}{n} \sum_{k=1}^{p} |\hat{\gamma}_k \beta_k| \tag{7}$$

In the program above, λ is a tuning parameter that controls regularization, and $\hat{\gamma}_k$ is an estimate of the optimal penalty loading $\gamma_k :- \sqrt{\sum_{i=1}^{n} x_{ik}^2 v_i^2}$. Following the recommendation of Belloni et al. [2012] (Appendix A), we set the tuning parameter to $\lambda = 2.2\sqrt{n}\Phi^{-1}\left(1 - \frac{0.1}{2p\log(n)}\right)$ and estimate the optimal penalty loadings using the following iteration procedure:

1. Set the initial penalty loadings to $\hat{\gamma}_k = \sqrt{\sum_{i=1}^{n} x_{ik}^2 (d_i - \bar{d})^2}$ and solve the LASSO program above to obtain $\tilde{\beta}$. Then estimate the first-stage residual as $\tilde{v}_i = d_i - x_i'\tilde{\beta}$.
2. Set the penalty loadings to $\hat{\gamma}_k = \sqrt{\sum_{i=1}^{n} x_{ik}^2 \tilde{v}_i^2}$ and repeat.
3. Iterate step (2) 15 times.

Since the DMA and month fixed effects should always be included in the regression, we partial out these fixed effects from both the endogenous variable (category ads) and the 40 potential instruments before running the LASSO procedure.

B.3 Further LASSO Results

To further compare the LASSO optimal instruments with our benchmark log-linear specification, we plot the partial-F and partial-R^2 in scatter plots in Figures 15 and 16. We observe three main patterns from this comparison. First, among the 221 categories with log-linear partial-F below

10, LASSO selects nothing for 186 categories, but produces a higher partial-F for 33 categories. This pattern can be clearly seen from the left tail in Figure 15 above the 45-degree line, and it suggests that the LASSO optimal instruments pick up a non-linear relationship between political and category ads that is neglected in the log-linear specification. Second, for each of the 28 categories with log-linear partial-F above 25, LASSO produces a smaller or equal partial-F and a slightly larger partial-R^2, as shown in the right tail in Figure 15 below the 45-degree line, and in Figure 16 above the 45-degree line. This pattern suggests the LASSO procedure trades off instrument strength (which relates to bias of the IV estimator) for a better first-stage fit (which relates to variance of the IV estimator).

To further understand what the optimal instruments capture, we calculate the predicted effect function $g_j(P) \coloneqq \beta_j' f_j(P)$ for each category j from the LASSO selection $f_j(\cdot)$ and the first-stage coefficients β_j. We then plot $g_j(P)$ against P for the 25 categories with LASSO $F > 25$ in Figure 17, for the 21 categories with both log-linear and LASSO $F \in [10, 25]$ in Figure 18, and for the 28 "upgraded" categories with log-linear $F < 10$ and LASSO $F \in [10, 25]$ in Figure 19. These curves are helpful for considering monotonicity. In Figure 17, the curves all look reasonably flat until very high levels of political advertising, when the crowd out becomes more severe. While some early ranges show some positive effects of P on A, it does not appear significant and is drastically outweighed by the negative impact at the high levels of P. Figure 18 shows similar results. Figure 19 looks somewhat different. There are a few brands with considerable amounts of positive effect of P on A which may be a warning sign of monotonicity violations. However, most categories in this figure show a very different crowd-out curve than those categories that have a strong first-stage in the log-linear case. This illustrates that the LASSO strengthens the first stage, particularly for brands that show patterns of crowd out that do not appear log-linear. However, this improvement in relevance comes at the cost of monotonicity concerns.

Figure 15: Partial F: LASSO v.s. Linear

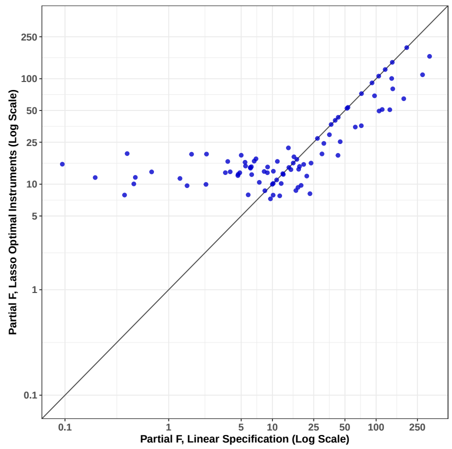


Figure 16: Partial R^2: LASSO v.s. Linear

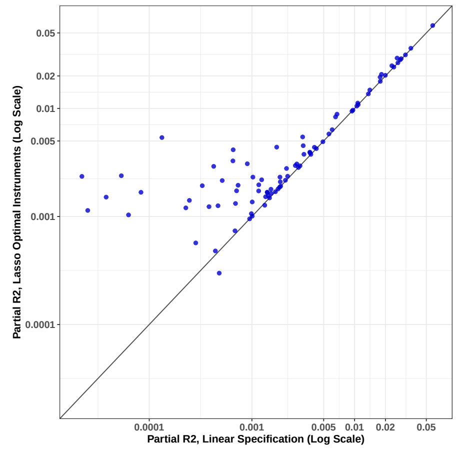
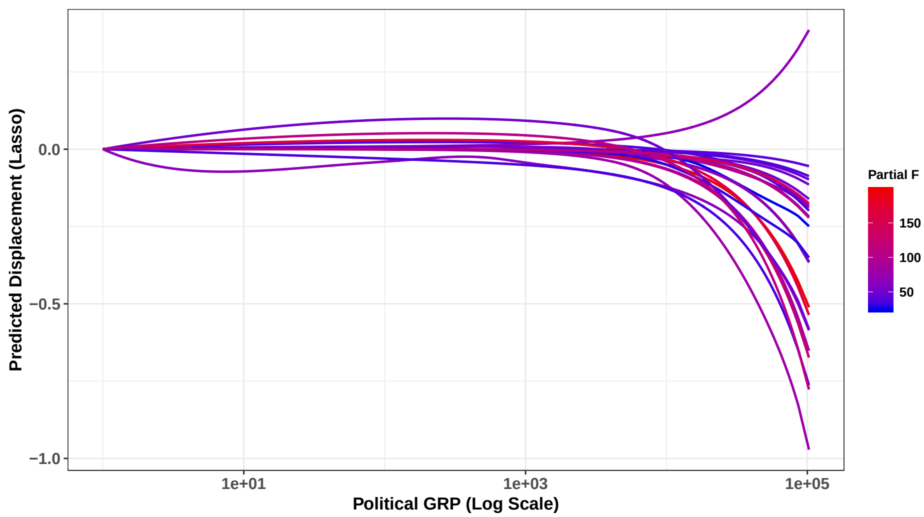


Figure 17: LASSO Predicted Displacement Effect, 25 Categories with LASSO $F > 25$



Notes: For each category, we compute and plot the predicted displacement effect $g_j(P) := \beta_j' f_j(P)$ at each level of political GRP P. Each line represents one category, and is colored by the Lasso partial-F.

Figure 18: LASSO Pred. Effect, 21 Categories with LASSO $F \in [10, 25]$ and Log-Linear $F \in [10, 25]$

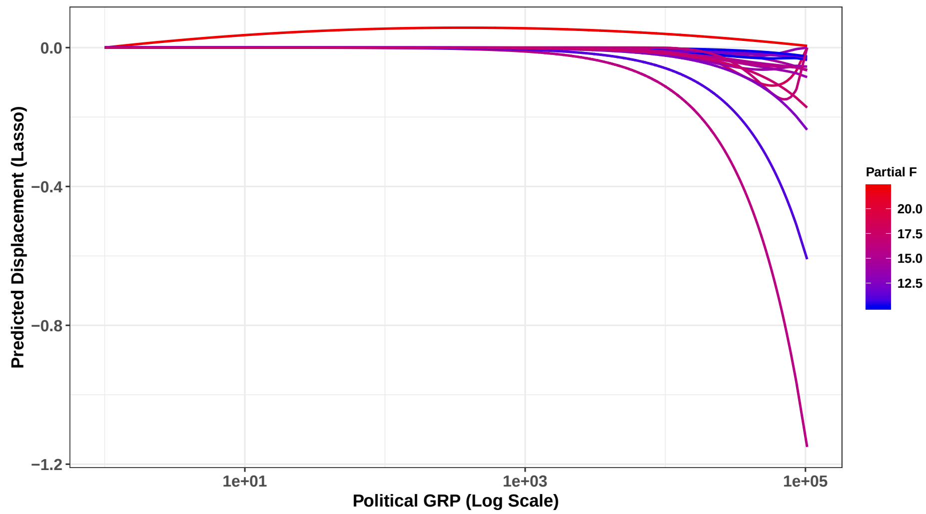
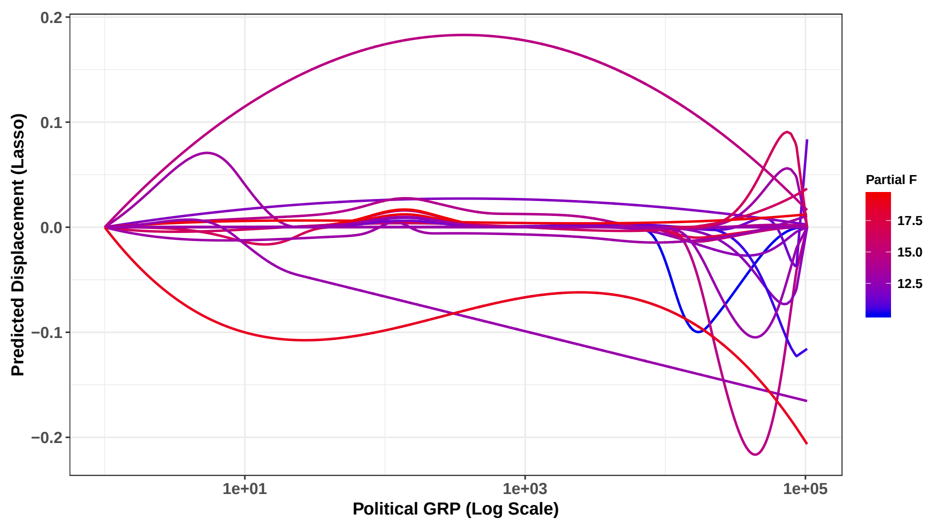


Figure 19: LASSO Predicted Effect, 28 Categories with LASSO $F \in [10, 25]$ and Log-Linear $F < 10$



C Alternative Method For Imputing Local Impressions

In this appendix, we explore a more flexible method for imputing local impressions data for non-sweep months. Let I_{ijkt} be Nielsen's impression estimate for station i, month t, day-of-week j, and 5-15 minute time interval k. Let t_1 be the last sweep month before t, and let t_2 be the first sweep month after t. When I_{ijkt} is not observed, we want to impute it by taking a weighted sum of I_{ijkt_1} and I_{ijkt_2}: $I_{ijkt} \approx \beta_1 I_{ijkt_1} + \beta_2 I_{ijkt_2}$.

Our simple imputation method in the main text uses time distance as weights: $\beta_1 = (t-t_1)/(t_2-t_1)$ and $\beta_2 = (t_2-t)/(t_2-t_1)$. The accuracy of this method on observed data in non-sweep months in LPM markets is summarized by the blue line in Figure 20, which shows that the method tends to overestimate the impressions.

Our alternative, more flexible method allows β_1 and β_2 to vary across days of week and 3-hour time intervals. Specifically, for each day of week j and each 3-hour time interval K, we take the observed impressions across all stations in LPM markets and run the following regression:

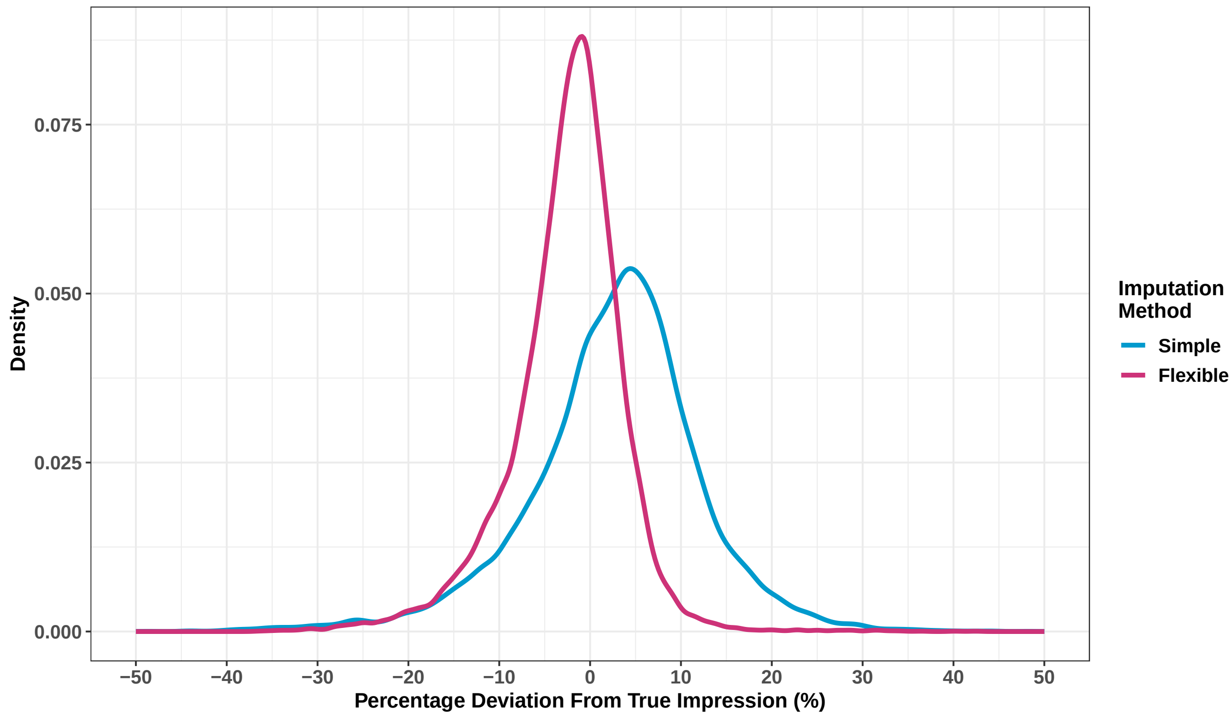
$$I_{ijkt} = \beta_{1jK} I_{ijkt_1} + \beta_{2jK} I_{ijkt_2} + \varepsilon_{ijkt} \qquad k \in K,\ j \text{ Fixed}$$

and then we use the estimated parameters to impute the unobserved impressions for all stations in non-LPM markets: $I_{ijkt} \approx \hat{\beta}_{1jK} I_{ijkt_1} + \hat{\beta}_{2jK} I_{ijkt_2}$. The accuracy of our flexible method on observed data in non-sweep months is summarized by the red line in Figure 20. The flexible method is able to correct a large fraction of overestimates.

Finally, we use the flexible imputation method to re-compute GRPs and re-run the log-linear first stage estimations as in Section 4.1. The scatterplot in Figure 21 compares the first stage partial-F based on the two imputation methods, and it shows that the results are virtually unchanged for most categories. While the partial-F from the flexibly-imputed data is slightly higher for most categories with strong or semi-strong first stage, the differences only change our classification of two categories, resulting in 29 categories with $F > 25$ and 25 categories with $10 < F < 25$.[42]

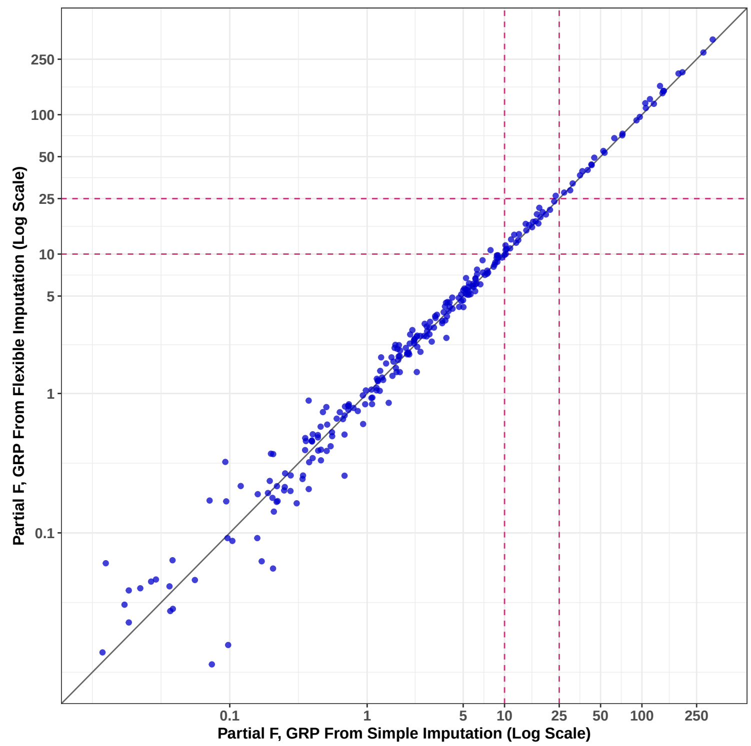
[42] Category "Plumbing & Sanitary Equipment, Fixtures, and Systems" has a partial-F of 23.6 under the original imputation and 26.2 under the flexible imputation. Category "Pre-Recorded Video" has a partial-F of 7.9 under the original imputation and 10.7 under the flexible imputation.

Figure 20: Accuracy of Simple and Flexible Impressions Imputation Methods



Note: The figure plots the distributions (kernel density estimates) of imputation errors for two imputation methods. Imputation error is calculated as the percentage deviation of the estimated impression from observed impression in sweep months. We aggregate observed and imputed impressions to [Market-Year-Month-Day of Week] level (i.e. summing the impressions across stations and time intervals) before calculating imputation errors. The densities are estimated across 12586 such observations for each imputation method.

Figure 21: First Stage Partial-F Statistics: Simple v.s. Flexible Imputation



Note: First stage partial-F statistics for 274 categories using GRPs calculated from flexibly-imputed impressions are plotted along the vertical axis. Partial-F statistics based on our simple imputation method are plotted along the horizontal axis.

D Exclusion Restriction Robustness Test

In this appendix, we conduct a robustness check under an assumption about the range of potential advertising effect sizes. In particular, we assume that the category-level advertising effect is positive ($\beta \geq 0$). Further, we posit that political advertising P crowds out commercial advertising A. Under these assumptions, the effect of P on commercial sales $log(1+Q)$ should be strictly negative. As a result, any positive and statistically significant reduced-form estimate of P on $log(1+Q)$ constitutes a violation of the exclusion restriction.

We implement this test on 36 CPG categories in our data that could be matched to one or more product modules in the Nielsen RMS database.[43] For each category j, we calculate the total revenue Q_{jmt} of market m in month t by summing up the revenue of all products in the category

[43]The RMS data contains the weekly quantity sold at the UPC-retail outlet level for several anonymous retail chains that partner with Nielsen.

across all stores in the market. We then estimate the reduced form regression as in equation (8), where the right hand side is the same as the first-stage equation (2), and the standard errors are clustered at market level. The RMS data is only available through 2014, so we have 60 months of data. For each category, we estimate the reduced-form equation:

$$\log(1 + Q_{mt}) = \Pi P_{mt} + \alpha_m + \alpha_t + \varepsilon_{mt}, \tag{8}$$

and report the reduced-form coefficient Π for each of the 36 categories in Table 9. We find that 5 of the 36 (14%) categories show positive and significant reduced-form coefficients, indicating a violation of the exclusion restriction; this is more than we should expect to occur simply due to chance. None of the 36 categories show negative and significant reduced-form coefficients. That said, only one of the 36 categories has a first stage $F > 25$ and only three more have $F \in [10, 25]$.

Overall, we think the results of this robustness test indicate that the exclusion restriction requires careful thought on a case-by-case basis. It is neither a weak nor obvious assumption. If we think that demand shocks U might be correlated across categories, as might be the case for an income shock, then the results here may themselves raise concerns for other categories.

Table 9: Exclusion Restriction Robustness Check for 36 CPG Categories

Category	RF Coef.	RF Std. Err.	RF p-value	Scaled RF Coef.	FS Partial F
Candy & Gum	9.89×10^{-7}	2.15×10^{-7}	0.000	0.023	1.2
Soups	1.04×10^{-6}	3.05×10^{-7}	0.001	0.024	5.9
Sponges	6.72×10^{-7}	2.55×10^{-7}	0.009	0.016	1.7
Deodorizers, Air Fresheners & Purifiers	5.91×10^{-7}	2.35×10^{-7}	0.013	0.014	1.2
all Other Prepared Dinners & Entrees & Pizzas	6.12×10^{-7}	2.71×10^{-7}	0.025	0.014	6.9
Jellies, Jams, Preserves, Peanut Butter	4.58×10^{-7}	2.56×10^{-7}	0.076	0.011	5.4
Cold, Cough & Sinus Remedies	5.48×10^{-7}	3.07×10^{-7}	0.076	0.013	12.7
Cheese Products	4.76×10^{-7}	2.84×10^{-7}	0.096	0.011	10.2
Cereals	3.31×10^{-7}	2.24×10^{-7}	0.142	0.008	30.1
Skin Care Creams, Lotions & Oils	-2.72×10^{-7}	1.97×10^{-7}	0.168	-0.006	0.3
Deodorants & Anti-Perspirants	-2.46×10^{-7}	1.83×10^{-7}	0.181	-0.006	0.0
Dietary Carbonated	-2.39×10^{-7}	2.04×10^{-7}	0.244	-0.006	0.1
Feminine Hygiene Products	1.84×10^{-7}	1.64×10^{-7}	0.263	0.004	1.3
Vitamin Preparations & Tonics	1.84×10^{-7}	1.65×10^{-7}	0.267	0.004	0.1
Cookies & Crackers	2.33×10^{-7}	2.11×10^{-7}	0.272	0.005	5.4
Coffee, Tea, Cocoa & Derivatives	2.09×10^{-7}	1.96×10^{-7}	0.290	0.005	12.2
Dog Food	1.89×10^{-7}	1.83×10^{-7}	0.305	0.004	0.2
Milk, Butter, Eggs (Including Powdered)	2.52×10^{-7}	2.48×10^{-7}	0.311	0.006	0.0
Food Wraps, Foils & Bags	2.14×10^{-7}	2.43×10^{-7}	0.381	0.005	8.8
Diapers(Incl Infant & Adult)	1.27×10^{-7}	1.53×10^{-7}	0.410	0.003	2.7
Regular Carbonated	-1.5×10^{-7}	1.82×10^{-7}	0.413	-0.004	6.1
Pain Relievers, Sedatives & Sleeping Preps	1.22×10^{-7}	1.63×10^{-7}	0.456	0.003	0.1
Salad Dressings & Mayonnaise	-1.94×10^{-7}	2.69×10^{-7}	0.473	-0.005	1.1
Shampoos, Conditioners & Cream Rinses	1.17×10^{-7}	1.83×10^{-7}	0.524	0.003	4.0
Beer	-4.03×10^{-7}	7.43×10^{-7}	0.588	-0.009	0.7
Fruit Juices & Fruit Flavored Drinks(Incl Powdrd)	9.39×10^{-8}	2.14×10^{-7}	0.661	0.002	0.2
Dental Supplies & Mouthwashes	6.62×10^{-8}	1.71×10^{-7}	0.700	0.002	1.2
Ice Cream, Frozen Novelties & Sherbet	-8.41×10^{-8}	2.27×10^{-7}	0.712	-0.002	0.0
Cat Food	6.17×10^{-8}	1.83×10^{-7}	0.737	0.001	0.2
Laxatives	4.92×10^{-8}	1.63×10^{-7}	0.764	0.001	2.3
Laundry Detergents & Cleaning Preparations	-5.26×10^{-8}	2.14×10^{-7}	0.806	-0.001	2.0
Liquor	-1.93×10^{-7}	9.07×10^{-7}	0.832	-0.005	1.5
Shaving Equipment & Supplies	4.03×10^{-8}	2.04×10^{-7}	0.843	0.001	1.3
Digestive Aids & Antacids	3.05×10^{-8}	1.81×10^{-7}	0.866	0.001	0.1
Hair Dressings, Sprays & Restoration Prod	-1.93×10^{-8}	1.8×10^{-7}	0.915	-0.000	3.0
Cleaners & Cleansers (Genl Hshold Use)	-1.42×10^{-8}	2.05×10^{-7}	0.945	-0.000	0.0

E Log-Linear First Stage with Advertising Stock

In this appendix, we report the log-linear first-stage results when the endogenous variable is the "ad stock" over the past 12 months, rather than the "ad flow" of the current month. Specifically, we estimate the following first-stage equation:

$$log(1 + S_{jmt}) = \beta_j Z_{mt} + \alpha_{jm} + \alpha_{jt} + \varepsilon_{jmt}, \tag{9}$$

where the ad stock S_{jmt} is the weighted sum of category GRPs A_{jmt} over the past 12 months:

$$S_{jmt} = \sum_{k=0}^{12} \delta^k A_{jm,t-k}. \tag{10}$$

The ad stock measure captures the cumulative effect of past advertising and is frequently used in the estimation of advertising elasticities. We use the monthly decay parameter $\delta = \sqrt[12]{0.9^{52}} \approx 0.6335$, which is consistent with the weekly decay parameter of 0.9 used by Shapiro, Hitsch, and Tuchman [2019]. We conduct this exercise for two versions of the instrument: the flow of political advertising ($Z_{mt} = P_{mt}$ as before) and the stock of political advertising, which we assume decays at the same rate as commercial advertising. ($Z_{mt} = \sum_{k=0}^{12} \delta^k P_{m,t-k}$) Because political advertising flow should not be able to affect the component of commercial advertising stock that comes from past commercial advertising, the political stock instrument might be a stronger instrument. However, the stock instrument requires additional assumptions.[44]

The number of categories with partial-F statistics above 25 is 28 for the ad-flow specification and 17 for the ad-stock specification with political ad flow as an instrument (the "stock-flow" specification). For the ad-stock specification with political ad-stock as the instrument (the "stock-stock" specification), there are only 14 categories with partial-F above 25. The distribution of partial-F for the three specifications are compared in Table 10, Table 11, and Figures 22--23 below. Figure 22 plots the partial-F for the "stock-flow" specification against the partial-F for the ad-flow specification, while Figure 23 plots the partial-F for the "stock-stock" specification against the partial-F for the ad-flow specification. The comparisons show that the ad-stock first stage is generally weaker than the ad-flow first stage. The weak first stages of the two ad-stock specifications are more pronounced for categories with strong ad-flow first stages, as shown by the right tails below the 45-degree line in Figures 22–23.

We also compare the scaled first-stage coefficients for the three specifications in Table 11 and Figures 24--25.[45] The negative effect of political advertising on ad stock is generally smaller than the negative effect on ad flow, which is expected because the ad stock is driven in a large part

[44] Using the stock of political advertising as an instrument requires an additional exclusion. That is, we must assume that the only structural state dependence operates through advertising carry-over. Otherwise, it is possible that increased political advertising in time $t-1$ leads to lower sales in time $t-1$, which then leads to lower sales in time t through structural state dependence. This violates the exclusion restriction that the instrument must only affect the outcome through its correlation with the endogenous variable.

[45] As before, the scaling factor is the median of instrument Z_{mt} across markets in October 2016. The value is 23438 for the political ad flow and 32597 for the political ad stock.

by past advertising that is not affected by current political advertising. The scaled coefficients are qualitatively similar in the specification with political ad stock as the instrument.

Table 10: Classification of Categories by Range of Partial-F in 3 Specifications

		Stock-Flow			Stock-Stock			Total (Ad-Flow)
		$F > 25$	$10 < F \leq 25$	$F \leq 10$	$F > 25$	$10 < F \leq 25$	$F \leq 10$	
Ad-Flow	$F > 25$	16	10	2	14	11	3	28
	$10 < F \leq 25$	1	9	15	0	10	15	25
	$F \leq 10$	0	6	215	0	3	218	221
Total (Stock-Flow / Stock-Stock)		17	25	232	14	24	236	

Table 11: Quantiles of Log-Linear First Stage Partial-F and Scaled Coefficients: 3 Specifications

Quantiles:		Min	0.05	0.10	0.25	0.50	0.75	0.90	0.95	Max
Partial-F	Ad-Flow	0.0	0.0	0.0	0.4	2.1	6.9	26.1	78.9	327.5
Partial-F	Stock-Flow	0.0	0.0	0.0	0.3	1.6	4.8	14.5	32.9	98.2
Partial-F	Stock-Stock	0.0	0.0	0.1	0.3	1.3	4.1	14.2	26.8	88.3
Scaled Coef.	Ad-Flow	-0.264	-0.121	-0.072	-0.029	-0.008	0.000	0.006	0.010	0.088
Scaled Coef.	Stock-Flow	-0.162	-0.066	-0.039	-0.016	-0.003	0.001	0.007	0.016	0.130
Scaled Coef.	Stock-Stock	-0.157	-0.063	-0.039	-0.014	-0.003	0.002	0.008	0.020	0.103

Figure 22: Partial F: "Stock-Flow" v.s. Benchmark

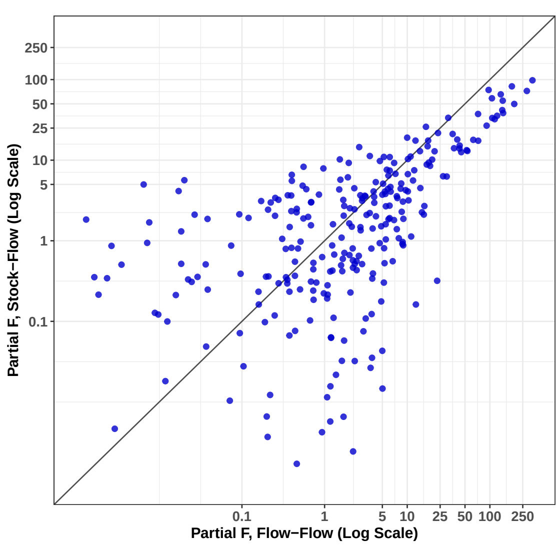


Figure 23: Partial F: "Stock-Stock" v.s. Benchmark

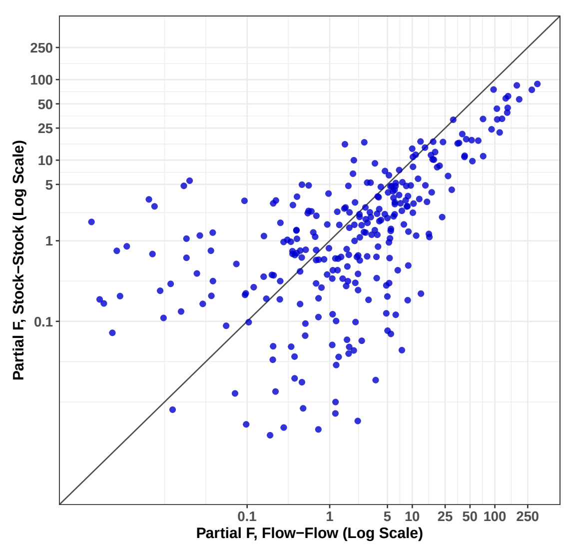


Figure 24: Scaled Coefficients: "Stock-Flow" v.s. Benchmark

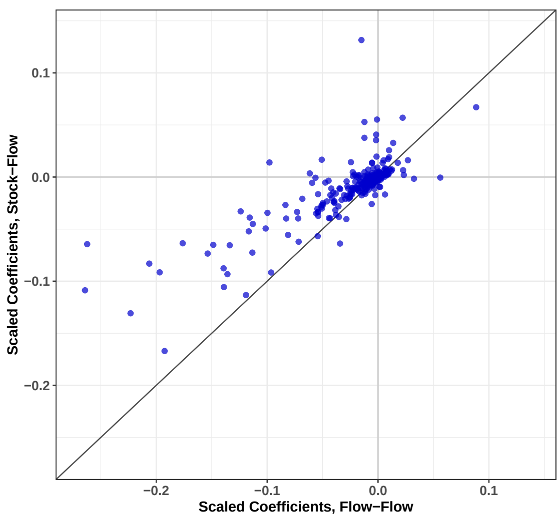
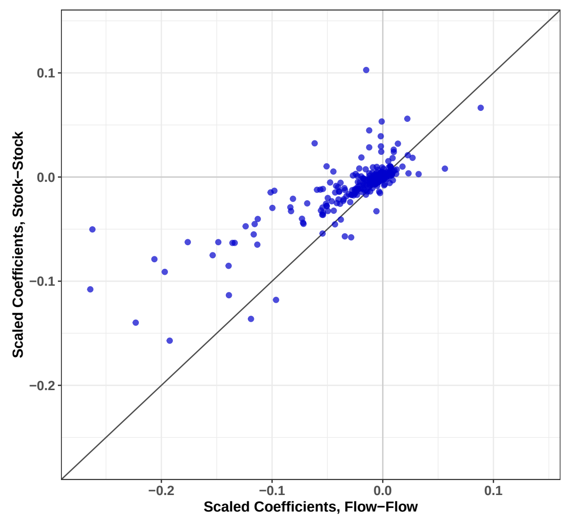


Figure 25: Scaled Coefficients: "Stock-Stock" v.s. Benchmark



F Distributions of First-Stage Coefficients With Alternative Fixed Effects

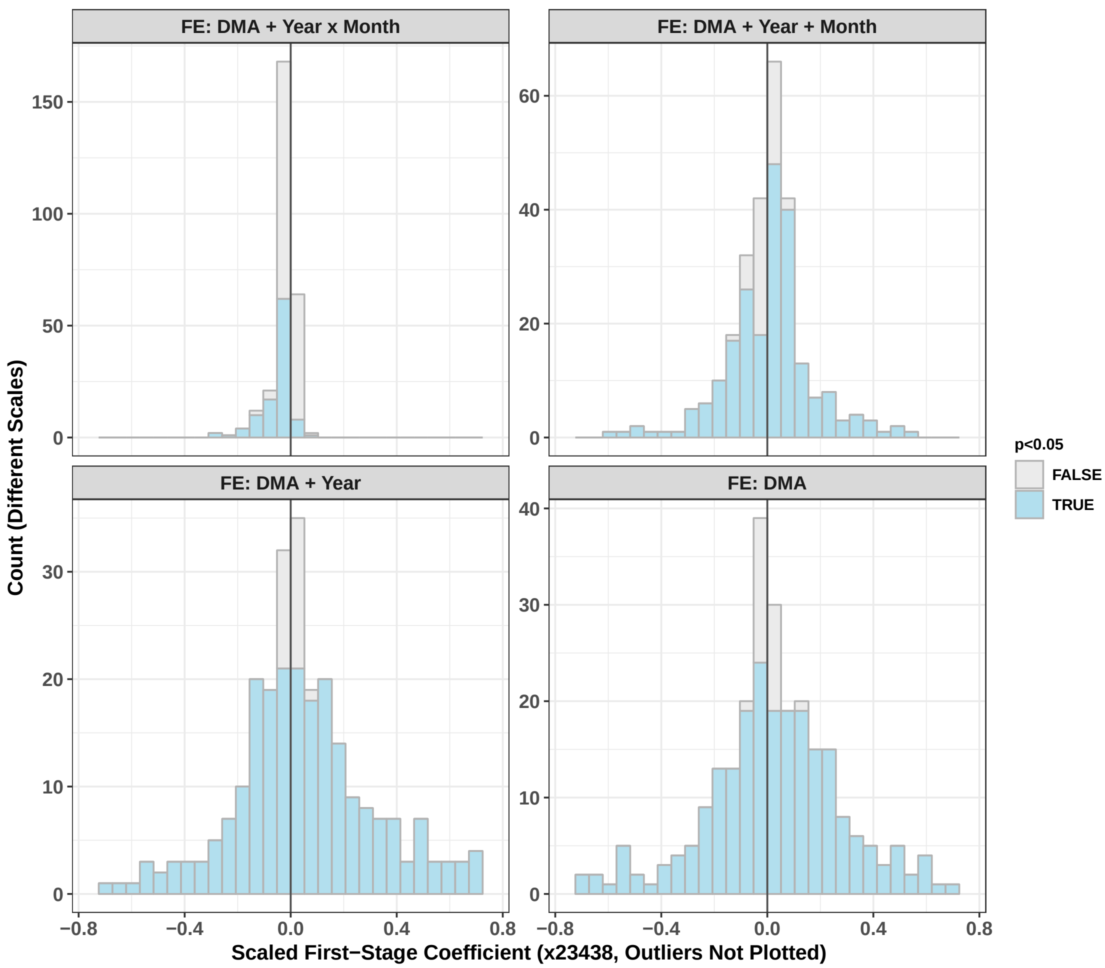
In this appendix, we provide suggestive evidence that DMA and year-month fixed effects are necessary for satisfying the monotonicity condition (as discussed in Section 2.3). Specifically, we estimate Equation (2) with three alternative sets of fixed effects: DMA + Year + Month, DMA + Year, and DMA only.

The number of categories with positive and negative first-stage coefficients under these specifications are reported in Table 12, and the distributions of coefficients are reported in Figure 26. Only 9 out of 274 categories (3.3%) have positive significant coefficients under the DMA and year-month fixed effects, which is roughly what would be predicted just due to chance. In contrast, more than 130 categories, or almost half of all categories, have positive and significant first-stage coefficients under each of the three alternative specifications. This shows that without the year-month fixed effects, the theory does not work as prescribed. This is consistent with failures of the monotonicity condition described in Section 2.3.

Table 12: Fraction of Positive / Negative First-Stage Coefficients Under Alternative Fixed Effects

Fixed Effects	Negative		Positive	
	$p < 0.05$	$p \geq 0.05$	$p < 0.05$	$p \geq 0.05$
DMA + Year × Month	96	112	9	57
DMA + Year + Month	92	31	131	20
DMA + Year	109	11	139	15
DMA	110	16	136	12

Figure 26: Distribution of First-Stage Coefficients Under Alternative Fixed Effects



G Implementation with FCC Data

In this appendix, we provide a roadmap for implementing the political advertising instrument with publicly available data from the Federal Communications Commission. The chief advantage of the FCC dataset is that it is available in real-time to all audiences; our hope is that it will broaden the applicability of this IV approach. Since August 2012, the FCC mandates that television stations in the fifty largest DMAs post data about their sales of airtime to political groups, including campaigns and Political Action Committees, to an online database, available at `https://publicfiles.fcc.gov`. The posts must be in real-time, and as of July 2014, the mandate was expanded to include all stations.[46] In September 2018, we collected data on the total number of files uploaded to the database by stations in each market in each month since August 2012. This dataset presents two challenges: first, files comprise invoices, order forms, contracts, and occasionally other forms such as a federal candidate certificates.[47] Each file therefore represents an unknown number of GRPs. The second challenge is that files must only be retained for two years, so that our data from August 2012-August 2016 may contain only a selected sample of stations which have chosen to retain their files for longer than required. This limitation is somewhat temporary, however, as we plan to update our file counts in real-time going forward. To explore whether file counts capture political advertising volumes, Figure 27 compares counts to GRPs measured by Ad Intel. We find a robust, positive relationship between the two measures. Note also that neither challenge in the FCC data threatens our identification strategy so long as file counts correlate with commercial airtime and also satisfy the exclusion restriction from section 2. The use of file counts merely requires that compliance must not be driven by a third factor that affects product markets directly.

[46] FCC, About Public Inspection Files, Accessed December 3, 2018.

[47] These must be ratified in order for an official campaign to receive lowest unit rates.

Figure 27: FCC File Counts vs Political GRPs

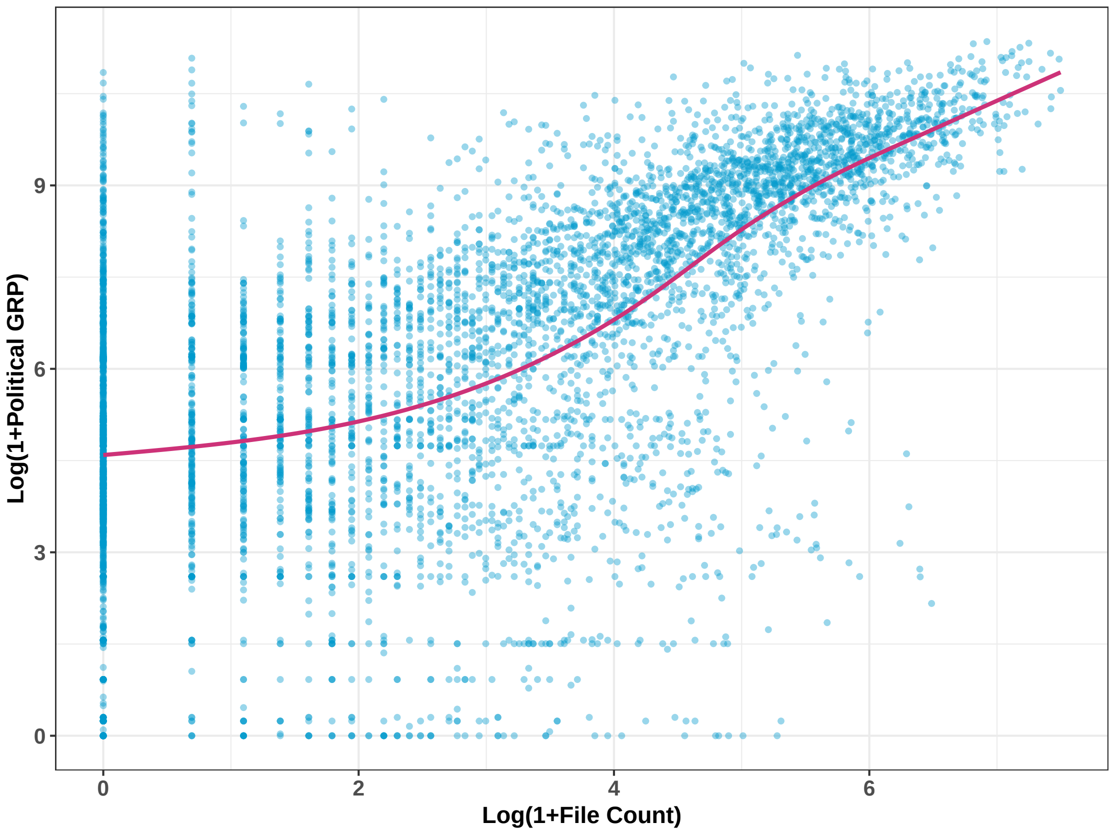


Table 13: FCC File Counts and Spots

	Mean	SD	Min	Percentile			Obs
				10	50	90	
Spots in an FCC File	57.71	74.66	1	7	37	124	3,869

Notes: Based on invoices, contracts, and order forms uploaded to the FCC online database during the 2012 Presidential Election. See Moshary [Forthcoming] for a detailed description of the data.

We compare the "free" FCC political file count instrument with the "paid" political GRP instrument under the log-linear specification. Table 14 tabulates the numbers of categories by first-stage strength for the two instruments: the FCC instrument achieves $F > 25$ in 15 categories and $10 < F < 25$ in 17 categories, compared to 28 and 25 for the political GRP instrument. Meanwhile, the number of categories with partial $R^2 > 0.01$ is 14, compared to 21 for the benchmark. In general, we find that the FCC instrument is weaker than the political GRP instrument at all levels.

The 15 categories in which the FCC instrument achieves $F > 25$ are listed in Table 15 along with their partial F and partial R^2 for both instruments. The comparison shows that the noise of the FCC instrument has a large impact on both the first-stage strength and explanatory power. However, the list also shows that the FCC instrument is mostly consistent with the political GRP instrument in terms of the categories where they are strong. In fact, most of the strong categories for the FCC instrument are the strongest ones for political GRP: all 15 categories in the list have $F > 50$ and $R^2 > 0.01$ in the benchmark specification. Therefore, the free FCC instrument may be an attractive option for researchers interested in one of these categories.

Table 14: Classification of Categories by First Stage Strength: FCC v.s. Political GRP Instrument

Instrument		FCC			Total (Political GRP)
		$F > 25$	$10 < F < 25$	$F < 10$	
Political GRP	$F > 25$	15	7	6	28
	$10 < F < 25$	0	7	18	25
	$F < 10$	0	3	218	221
Total (FCC)		15	17	242	

Table 15: List of 15 Categories With $F > 25$: FCC Instrument

Category	Partial F		Partial R^2		Disp. Effect	
	FCC	Pol. GRP	FCC	Pol. GRP	FCC	Pol. GRP
Household Furnishings & Appliance Stores	82	330	0.054	0.058	-0.07	-0.14
Miscellaneous Organization Advertising	74.3	154	0.054	0.05	0.08	0.13
Miscellaneous Professional Services	73.3	141.2	0.042	0.028	-0.04	-0.05
Autos & Light Truck - Dealerships	72.5	280.2	0.032	0.035	-0.1	-0.2
Light Trucks & Vans-Factory:new & Cpo	49	106.1	0.02	0.024	-0.03	-0.05
Hospitals,physicians & Misc.physical Culture	48.4	192.9	0.024	0.03	-0.06	-0.12
Optical Goods and Services	45.9	183.2	0.021	0.018	-0.08	-0.14
Restaurants, Hotel Dining & Nightclubs	35.4	134.6	0.022	0.026	-0.02	-0.03
Misc Entertainment & Combination Copy	34	146.7	0.008	0.02	-0.1	-0.27
Dance, Theater, Concerts, Opera	33.2	90.2	0.01	0.014	-0.09	-0.15
Passenger Cars-Factory:new & Cpo	31	143.4	0.017	0.026	-0.02	-0.04
Schools & Camps, Seminars	28.8	72.6	0.014	0.014	-0.03	-0.06
Construction, Engineering & Archit Srvcs	28	120.9	0.013	0.017	-0.07	-0.15
Miscellaneous Retail	27.5	112.7	0.018	0.023	-0.02	-0.04
Passenger Cars-Dealer Assn:new & Cpo	25.2	110	0.01	0.018	-0.1	-0.21