Price Discrimination across PACs and the Consequences of Political Advertising Regulation

Sarah Moshary*

January 14, 2015

Abstract

The rapid growth of Political Action Committees – expenditures neared $500 million in the 2012 presidential election – is center-stage in the debate over money in American politics. The effect of PACs on elections depends on regulation and its interaction with imperfect competition. Congress requires stations to treat candidates to the same office equally, and to sell campaigns airtime at lowest unit rates (LURs) within sixty days of a general election. This paper examines pricing to PACs, which are not protected under the law, and the impact of political advertising regulation, in particular, lowest unit rate regulation. Using novel data on prices paid for individual ad spots from the 2012 presidential election, I find that stations price discriminate substantially across PACs for indistinguishable purchases. On average, PACs pay 40% markups above regulated rates. Republican PACs pay 14% higher prices on average, but there is substantial idiosyncratic variation in prices paid across ad spots. I develop and estimate a model of political demand for ad spots, exploiting misalignments of state borders and media markets to address potential price endogeneity. Findings indicate that pricing to PACs reflects buyer willingness-to-pay for viewer demographics. Taken together, these results indicate the current regulatory regime differentially subsidizes candidates depending on the characteristics of their base. Using a station price discrimination model, I then estimate a cost of regulation: strategic quantity withholding of airtime to keep regulated rates high. Using Bayesian MCMC methods, I estimate this effect is substantial – on the order of 7% of total advertising airtime – relative to a counterfactual without regulation.

*Department of Economics, MIT. Email: smoshary@mit.edu. I would like to thank Glenn Ellison, Nancy Rose, and Paulo Somaini for invaluable help and advice throughout this project. I would also like to thank Nikhil Agarwal, Isaiah Andrews, Josh Angrist, Mark Bell, Vivek Bhattacharya, David Colino, Kathleen Easterbrook, Sara Fisher Ellison, Ludovica Gazzè, Jerry Hausman, Mark Hou, Sally Hudson, Gaston Illanes, Anna Mikusheva, Arianna Moshary, Casey O'Brien, Ben Olken, Manisha Padi, Brendan Price, Otis Reid, Miikka Rokkanen, Adam Sacarny, Richard Schmalensee, Amy Ellen Schwartz, Duncan Simester, Bradley Shapiro, Ashish Shenoy, Daan Struyven, Melanie Wasserman, Nils Wernerfelt, and participants in the Industrial Organization seminar and field lunches at MIT. Please see the most recent version at: economics.mit.edu/grad/smoshary

Political Action Committee (PAC) spending in American elections skyrocketed to $1.3 billion in 2012.[1] This rise follows a series of Supreme Court decisions in 2010 abolishing limits on contributions to PACs, while leaving caps on donations to official campaigns untouched.[2] PAC growth is a boon for television stations, since TV advertising claims the lion's share of political dollars – it continues to be the primary vehicle of communication between candidates and the electorate in the United States (Fowler & Ridout 2013).[3] While the Federal Communications Commission (FCC) regulates television station sales of airtime to official campaigns, the law is silent on the treatment of PACs. Station owners potentially wield considerable power to shape PACs' access to the airwaves. Using novel data on television advertisement prices from the 2012 presidential election, this paper investigates how TV stations set prices for PACs. Since prices govern where and when PACs advertise, understanding station behavior is crucial in assessing how these groups will ultimately impact American elections. Station sales of airtime to PACs can also shed light on how existing regulation implicitly subsidizes candidates. I estimate a model of station pricing, and use this model to evaluate the existing regulatory regime.

Pricing to PACs depends first on the extent of TV station market power. Stations typically negotiate rates with commercial advertisers in an upfront market for ad spots, and they are suspected of selling airtime at different rates to different advertisers. As an example, large purchasers may receive substantial discounts (Blumenthal & Goodenough 2006). To ensure candidates equal access to viewers, the FCC mandates all campaigns pay the same prices for ad spots, set at lowest unit rates (LURs) (Karanicolas 2012).[4] In contrast, newspaper headlines decry stations' charging "super-gouge" rates to PACs. However, evidence on TV station pricing – to political or commercial advertisers – is thin.[5]

A first important finding is that on average, stations charge PACs 40% markups for air-

[1] I use PACs as an umbrella term for outside spending groups, including traditional PACs, super PACs, and 501(c) organizations. Spending estimates come from OpenSecrets.org [https://www.opensecrets.org/outsidespending/fes_summ.php?cycle=2012].

[2] For a full description of contribution laws, see: the Federal Election Commission. "Federal Election Campaign Laws." Washington, D.C.: 2008. Restrictions on donations to certain kinds of PACs remain. Crucially, there are no restrictions on donations to Super PACs, which accounted for 50% of spending in 2012.

[3] Wilner, Elizabeth. 2012. "Romney and Republicans Outspent Obama, but Couldn't Out-Advertise Him." Advertising Age. November 9.

[4] Lowest unit rate rules come into effect within 45 days of a primary election and 60 days of a general election. The Federal Communications Act of 1934, 47 U.S. Code §151. There is some ambiguity in implementation of the law. Law firms specialize in advertising law to help stations ensure they follow precedent.

[5] For example: Peters, Jeremy W., Nicholas Confessore and Sarah Cohen. 2012. "Obama is Even in TV Ad Race Despite PACs." The New York Times. Oct 28. Bykowicz, Julie 2012. "TV Stations Charge Super-Gouge Rates for Super PACs." Bloomberg News. Oct 6.

time above the campaign price. Higher prices ought to temper the effectiveness of PACs, since each PAC dollar is worth less than its campaign counterpart. Second, in a comparison of indistinguishable ads, Republican PACs pay higher prices than Democrat PACs (on the order of 14%), but there is substantial idiosyncratic variation in price differences across spots. While the literature (for example, Goettler 1999 or Bel & Domènech 2009) has documented correlations between average prices and program characteristics, such as audience size, this finding contributes to the limited empirical evidence that TV ad prices also differ substantially across buyers.[6]

Identifying the forces that drive ad pricing – and price differences across PACs – is essential to thinking about counterfactual regulatory regimes and evaluating current policy. One possibility is that station owners sell airtime more cheaply to the party they support privately. If owner bias were the primary driver of station pricing, then policy shielding PACs could be crucial in guaranteeing voice to diverse political ideologies. I test whether bias, measured by political donations, correlates with preferential pricing, but find little evidence linking donations to pricing. Alternatively, prices may reflect viewer taste for PAC advertisements, as Gentzkow & Shapiro (2010) find for print newspapers. Differences in negotiation costs could also drive pricing, as Goldberg (1996) finds for used-car dealer transactions with women and minorities. It may be that Republican PACs purchase airtime closer to run-dates, and this timing element accounts for price differences. I examine a fifth possibility: that stations price based on PAC taste for viewer demographics.

I develop a model of station price discrimination based on PAC willingness-to-pay for different demographics, and test whether station behavior is consistent with this model. The key ingredients for this test are Democrat and Republican PAC preferences for viewership. These preferences depend on the strategy PACs pursue; for example, a get-out-the-vote strategy involves PACs targeting their base, whereas a persuasion strategy necessitates targeting swing voters.[7] To estimate PAC preferences, I exploit the sensitivity of political demand to state borders. Some ad spots bundle viewers in contested and uncontested states; political advertising demand ought to be orthogonal to viewership in the latter. Since commercial advertisers value these extra viewers, audiences in uncontested states constitute a residual supply shock for political advertisers. Results provide evidence both of vote buying, where parties target swing viewers, and turnout buying, where parties target their bases.

Since estimation imposes no model of supply-side behavior, I construct a test of whether

[6] An older literature considered quantity discounts in television advertising, see Bagwell (2007) for a discussion.

[7] See Nichter (2008) for a complete categorization of strategies.

observed prices are consistent with a model of station price discrimination based on PAC willingness-to-pay. My main finding is that model-generated utility estimates strongly correlate with observed prices. This relationship is robust, even controlling for costs and unobserved quality, suggesting stations do price discriminate and charge higher prices for ads PACs value most. Estimates employ lowest unit rates to measure cost, since these approximate the opportunity cost of selling airtime to PACs.

Taken together, my findings suggest lowest unit rate regulations benefit campaigns that prize demographics unfavored by commercial buyers. For these campaigns, regulated rates are likely to be far below their willingness-to-pay. Parties able to channel donations through campaigns also benefit disproportionately. If redistribution across parties and campaigns is desirable – for example, from candidates with a small number of wealthy supporters to candidates with a large, but less affluent base – this may be interpreted as a regulatory success story. Further, it suggests that under the current regime, market forces drive inequalities in access to viewers for PACs, rather than media bias – a distinction potentially important in shaping future PAC regulations.

I then extend the model of station behavior to examine one cost of regulation: whether and how much lowest unit rate regulation distorts station behavior. Regulation ties the price campaigns pay for airtime to the price in the commercial market, incentivizing stations to raise prices for other advertisers. To extract rent from campaigns, stations may sell less airtime compared to a counterfactual without regulation. Duggan & Morton (2006) document similar distortions in prescription drug sales in response to Medicaid regulation. Since commercial sales data is proprietary (and therefore unobserved), I estimate quantity withholding using a Tobit-style structural model, which specifies both commercial and campaign demand. I estimate the model using Bayesian MCMC techniques to avoid difficulties in gradient-based optimization of discontinuous functions. The likelihood function then allows me to infer the quantity sold on the commercial market from the station's first order condition. The results suggest a 7% reduction in airtime.[8] These finding highlight the importance of considering market imperfections in creating government pricing policies.

The paper proceeds as follows. Section 1 describes the data sources and construction of key variables for my analysis. Section 2 provides reduced-form evidence on price discrimination across PACs. Sections 3.1 and 3.2 develop a model of PAC demand for ad spots. Sections 3.3-3.5 lay out my estimation strategy, which exploits state borders to recover demand parameters. Results on PAC taste for viewer characteristics are presented in section 3.6. Section 4 outlines a model of station price discrimination, and tests whether the model is consistent with observed prices. Section 5 builds and estimates a model of quantity

[8] If viewers dislike advertising, then this constitutes an upper bound on the welfare loss from regulation.

withholding by stations in response to LUR regulation. Section 6 concludes.

1 Data

In this section, I detail the three main data sources used in this study: an online FCC database on ad prices, Simmons survey data on viewership, and US Census data on market demographics. Then I describe statistics on viewership derived from the combined data sources.

1.1 Data Sources

The primary data for this paper is scraped from a newly mandated Federal Communications Commission online database. As of August 2nd, 2012, stations in the 50-largest Designated Market Areas (DMAs) are required to post detailed information about political ad sales online. This requirement only holds for the four largest stations in each DMA: CBS, NBC, ABC, and FOX affiliates.[9] The records include the station, client, media agency, show name, time, date, and purchase price for each transaction. Such detailed data is unique in the advertising arena (see Stratmann (2009) for a description of standard data sources). The extensive political science literature has employed fairly coarse data on prices in the past. As an example, researchers often impute ad exposure using campaign spending, potentially confounding quantity with quality.[10] CMAG's (Campaign Media Analysis Group) data on ad counts acquired via satellite technology is a popular alternative, but it contains no information about prices. Other work has employed TV station logs, but until the advent of the FCC online archive, large-scale data collection was prohibitively expensive. To my knowledge, this is the first paper to exploit the newly-available ad buy data on the archive.

While this new data is incredibly detailed, it is not without flaws. Stations upload data in a variety of formats. Some stations post only order forms or contracts (which do not include the specific date and time the ad is run, but only a date and time range), while others post actual invoices with as-run logs. The data quality varies by station; some stations have posted low-quality scans of official documents. These forms are parsed less accurately by optical character recognition software than are high-resolution documents. Therefore, this data is likely to be incomplete for stations that upload in this format (not that observed ads

[9] Federal Communications Commission. News Media Information. "FCC Modernizes Broadcast Television Public Inspection Files to Give the Public Online Access to Information Previously Available only at TV Stations." By Janice Wise. Washington, D.C: 2012.

[10] See Goldstein & Ridout (2004) for a detailed review of the literature.

are misreported, but that the program misses some ads altogether).

Advertising data is paired with viewership data from Simmons, which is based on their annual survey of 25,000 American households.[11] Since ad spots are not a homogenous good – in the data they range in price from $10 to $650,000 – data on viewership is instrumental in understanding pricing. Although ad spots are the unit of sale, advertiser demand is really for viewers. The Simmons data allows me to deconstruct each ad into a collection of viewers. For each show, it contains the number of viewers by race, gender, and age.

The final data set contains 128,051 ad-level observations placed between August 1st and November 6th, 2012. This represents a subsample of the ads actually run over the course of the entire election (approximately 15%)[12] for four reasons: (1) ads purchased prior to August 1st are not required to appear on the website, and so are not included here; (2) OCR software imperfectly parses photocopied invoices; (3) the FCC only required the 200 stations in the fifty largest DMAs to post on the website, excluding roughly 1,600 TV stations from my sample;[13] and (4) PACs explicitly focusing on non-presidential races are excluded from the analysis (approximately 16% of ads).[14] The final sample includes ads placed by over 60 political groups (42 pro-Republican and 20 pro-Democrat) at 37 TV stations in 19 DMAs. Table A1 shows the breakdown of ads by Political Action Committee.

The sample appears to be fairly representative based on comparisons to Fowler and Ridout's (2013) description of Kantar Media/CMAG's data. The CMAG sample includes all local broadcast, national cable, and national network ads for 2012, but contains no information about ad prices. As an example, the ratio of Romney to Obama campaign ads is the same across the samples (approximately 2:5). Fowler and Ridout report that the average price of an Obama campaign ad was strikingly lower than its Romney counterpart, a pattern mirrored in my data (table 1). This difference in average prices ($297), they attribute to different program choices – evidence, perhaps, that Obama successfully employed a more sophisticated ad-buying strategy. Alternatively, Romney ads might be more effective for different (and potentially more expensive) audiences, leading to differences in average prices. The Romney campaign also purchased higher viewership ads, which might account for the high prices. My sample includes a higher proportion of PAC to candidate advertisements than the CMAG data. Fowler and Ridout designate ads as "presidential" based on content, while my criteria includes any ad purchased by PACs that donated to a presidential campaign, had a clear political affiliation, and did not explicitly support a candidate in another

[11] Experian Marketing Services, Summer 2010 NHCS Adult Study 12-month. Simmons data is also used by Martin & Yurukoglu (2014) to assess the relationship between media slant and viewer ideology.

[12] Fowler & Ridout (2013) estimate 1,431,939 were run from January 1, 2012 to election day.

[13] Fung, Brian. 2014. "A Win for Transparency in Campaign Finance." The Washington Post. July 1.

[14] I discard observations at stations without dual PAC and campaign advertising.

race. Categorization of PACs is based on records from the Center for Responsive Politics.[15]

This new data on prices reveals important facts about the political ad market, and the scope for price discrimination. Figure 1 shows that prices (per viewer) increase in the run up to election day, consistent with stations' extracting rent from political advertisers. Figure 2 shows that advertising quantities also rise over time. Political groups are likely to value ads run later in the cycle for myriad reasons: impressions decay quickly; candidates may be trying to exhaust their budgets; and the identities of swing voters may become clearer as the election draws near. The average ad over the three-month period cost $1,260 and reached some 229,446 viewers.[16]

To get a sense of the importance of lowest unit rate regulation, I compute markups for PAC purchases above lowest unit rates during the 60 day period before the election. During this period, PACs (by law) pay weakly higher prices than campaigns.[17] On average, Republican PACs pay 35% (standard error of 2.9%) markups and Democrat PACs pay 46% (standard error of 4.6%) markups above lowest unit rates. These comparisons suggest LUR regulation provides a significant discount for campaigns. Candidates able to channel money through their official campaign therefore benefit most from regulation. Since current campaign finance laws restrict individual donations to campaigns, candidates with many, small donors can exploit regulation best.

1.2 Who Sees Political Ads?

Campaigns and PACs ultimately value winning elections. Ad spots are valuable because they reach viewers, viewers cast votes, and votes create winners. A contribution of this paper is to estimate ad exposures in the 2012 election. This exercise is similar in spirit to Ridout *et al.* (2012), who examine the distribution of campaign purchases across television shows for the 2000, 2004, and 2008 presidential elections. However, I combine survey data on viewership with demographic data in a novel fashion that maps ad spots to their underlying viewer composition.[18]

[15] I conducted searches on OpenSecrets.org, maintained by the Center for Responsive Politics. In two cases, I obtained political affiliations based on newspaper articles linking groups to partisan advertising when the organization was not categorized by OpenSecrets.org.

[16] I winsorize prices (1%) to mitigate the effect of outliers in the rest of the paper.

[17] Stations may try to circumvent regulation by redefining classes of time so that campaigns pay higher prices than PACs for ads that tend to air at the same time. However, creating a campaign-specific class of time is considered illegal. For some comparisons, campaigns therefore seem to be paying higher prices despite lowest unit rate rules (.2% or 22 out of 1,112 cases). I include these observations when calculating average markups. (See Wobble Carlyle Sandridge & Rice, LLP. 2014. "Political Broadcast Manual." Washington, D.C. By John F. Garziglia, Peter Gutmann, Jim Kahl and Gregg P. Skall.).

[18] Their work suggests Democrat and Republican candidates target different viewers, but does not consider how choice might be affected by differences in price across programs and differences in availability across programs.

I infer ad viewership by marrying three data sources: FCC data on show names, times, stations, and networks; Simmons data on the viewing habits of different demographic groups; and 2010 census data on the population demographics by DMA. I match each purchased ad spot from the FCC logs to viewership using show title or network and time (for example I assign average ABC 8am weekday viewership to all spots fitting that description without a discernible show title). Matching without a specific name is useful since invoices often describe purchases by these attributes rather than a "name." Also, this matching strategy allows me to analyze new shows (premiering after 2010) although they do not appear directly in the Simmons data.[19]

Let j denote the program and g denote a demographic group (e.g. white women under 65 years of age). π_{gj} is the probability a member of group g sees ad j, approximated by counts from the Simmons data. Let J_{cs} denote the set of ads broadcast in state s that support candidate c. Aggregating across this set produces total exposures for the demographic group in state s supporting candidate c.

$$A_{gsc} = \sum_{j \in J_{cs}} \pi_{gj}$$

Variation in ad viewership across states comes from demographic differences and differences in the composition of J_{sc} (ad purchases), rather than preference heterogeneity within the same group across states. Intuitively, in states with a higher proportion of individuals in group g, an ad that targets that group is more productive.

Estimated average exposures for each demographic group are displayed in table 2. Across all groups, viewers see approximately five times as many Republican PAC ads than their Democrat counterparts, which is consistent with Fowler and Ridout's findings. Based on ad-airings by the 12 largest PACs in the 2012 race, they calculate that Democrat spots accounted for 18% of political ads run. Interestingly, the skew in advertising is exacerbated at the exposure level; the difference in exposures across parties is higher than the ad counts would suggest. Republican PACs not only buy more ads, but they also buy higher viewership ads. Although ad counts put the Democrats ahead, these exposure estimates suggest Republican PACs and the Romney campaign reached more viewers that the Democrat PACs and the Obama campaign combined during the three months preceding the election.

Women see more political ads compared to men, and blacks see more spots compared to other racial groups. Both of these findings are in line with Ridout *et al.* (2012)'s tabulations for the 2008 election, and also with the broad TV watching habits of these demographic

[19]This assumes demographics are stable across years for each time slot. If networks replace shows strategically, this matching algorithm will under-predict the value of ad spots that air during new shows.

groups. As an example, the probability a person watches a show is 5.5% in my sample. For women, this probability is slightly higher (5.9%) and for men it is slightly lower (5%), so that on average, women are 20% more likely to watch a show than men. Based on viewership habits, then, it seems reasonable that women also see approximately 20% more political ads than men.

These aggregate statistics, while hinting at PAC demographic targeting, confound advertiser preferences over demographics and TV viewing differences across these demographics. To understand how much variation in exposures is due to advertiser choice requires reconstructing the menu of potential ad buys, rather than simply looking at purchased spots. Data on rejected ad spots will allow me to determine how purchase decisions relate to viewership composition. As an example, if rejected spots featured an even higher proportion of white women than the set of purchased spots, then it seems unlikely that they are a coveted demographic.

To construct the menu of potential spots, I partition each station-week into weekday/weekend spots, and then into 1-hour intervals (24×2 spots per station). However, Simmons only records viewership coarsely for early-morning shows, so I exclude programs airing between 12-5am, reducing the number of distinct products to 35 for each station, each week between August 1st and November 6th, 2012.[20] Spot viewership depends on local demographics and network programming. In total, there are 18,900 distinct products (36 stations $\times$ 15 weeks $\times$ 35 day parts).

Ad spots are often also described by a priority level, and an indicator for which particular days are permissible runtimes.[21] Priority level characterizes how easily a station can preempt an ad, should they oversell slots on a show. While stations air preempted ads on another show with similar characteristics, industry wisdom is that so-called "make-goods" are worse quality (Phillips & Young (2012)). Low priority purchases constitute a gamble on the level of residual supply. Purchasers can also specify the day of the week for ad spots. As an example, an ad spot could be described as "Wednesday's Today Show" or "Wednesday or Thursday's Today Show." Rather than defining these combinations as separate commodities, I will control for these features in demand estimation.[22]

2 Price Discrimination across PACs

Fear of inequitable media access across candidates is a key motivator for the regulation

[20] During primetime, intervals narrow to 30 minutes. During early early morning, intervals are wider. In the simplest model, stations have a 168 products each week, one for each hour of each day.

[21] If the station records only invoices with "as-run" logs, then it is often not possible to determine these characteristics of the purchase. I include a dummy in demand estimation as a flag for these missing values.

[22] For rejected shows, I assign characteristics in proportion to their presence in the purchased sample.

of political advertising (Karanicolas 2012). To shed light on whether these fears are well founded, I examine station behavior towards PACs, which is as yet unregulated. In particular, I test whether Republican and Democrat PACs pay the same prices for the same exact ad spots. To the contrary, I find that stations seem to price discriminate by political affiliation.

2.1 Do Republican and Democrat PACs Pay the Same Prices?

In this section, I compare prices paid by Democrat and Republican PACs for indistinguishable ad spots. It is unclear to what extent stations can tailor prices across different political buyers. Stations may lack the market power and information to price discriminate across political advertisers. Indeed, if the market for airtime were perfectly competitive, lowest unit rate regulations would be irrelevant, since all buyers would pay the same price for airtime. Because the presidential race is a national one, network affiliates compete both within and across DMAs for political dollars. High rates in one DMA would ostensibly induce substitution to other markets. More and more, stations also compete with other forms of media like Facebook and Twitter. Separate from competitive pressures, it is possible that stations lack the information to price discriminate. The first task of this paper, therefore, is to examine the extent and type of station price discrimination across PACs. Apart from providing insight into a counterfactual world with less regulation, PAC advertising, which nearly matched campaign expenditure in 2012, is itself an important piece of the competitive election puzzle.

I construct a price comparison for Democrats and Republicans using a restricted set of ad purchases. I consider cases where PACs supporting opposing candidates purchase airtime on the same program (identified by name), for the same date, on the same station, and at the same hour.[23] For this analysis, I treat the PACs supporting a particular candidate as a single entity, both for practical reasons (there are too few observations for one-on-one PAC comparisons) and also bearing in mind that like-minded PACs should value ad spots similarly, since they share an objective (elect their party's nominee). A price-discriminating station should therefore charge these PACs similar prices. On the other hand, if stations charge Democrat and Republican PACs similar prices for airtime, then it seems unlikely that stations are discriminating (unless these groups share the same willingness-to-pay for viewers – in which case, stations would not be able to engage in taste-based discrimination).

Table 3 shows the results of this same-show comparison. There are 717 shows where liberal and conservative PACs purchased exactly the same ad spots. In 212, they pay different prices for those ads. The average price difference is $196.88, approximately 26% of the total

[23] For this exercise, I consider only shows where the OCR software successfully scraped the full show name.

price. While Republicans pay more on average ($68.41), Democrats are almost equally likely to pay higher prices (among instances where Democrat and Republican PACs pay different prices, Democrat PACs pay more almost 50% of the time). That neither Democrats nor Republicans pay more across the board suggests price discrimination is more complicated than simple party favoritism (for example, stations always charging Republican PACs more).

Regulation provides a nice placebo test for this exercise: since federal law prohibits stations from charging the candidates different prices, the same comparison for the Obama and Romney campaigns should yield zero price discrepancies. Of the 103 shows where both campaigns purchase, candidates only pay different prices for 20. Further investigation reveals that half of these are errors in the data-gathering process (faults in the optical character recognition software). Reassuringly, the price differences between PACs are more than twice as large for candidates, suggesting that the PAC price gap is more than a coding error.

I also examine within-party price differences for the 37 Republican PACs and 17 Democrat PACs in my data. For each ad purchased by multiple PACs with the same political affiliation, I calculate the coefficient of variation for prices (the standard deviation divided by the mean). Table 4 shows the mean coefficient of variation for the full sample in column (1). Price dispersion is highest across parties. The coefficient of variation is 0.11 for the full sample of dual Republican and Democrat purchases. The standard deviation, on average, is over 10% of the price. In comparison, the coefficient of variation is an order of magnitude smaller for within-party comparisons.

There is a potential selection problem in the column (1) comparison, since the coefficient of variation is measured conditional on purchase. As an example, constructing the coefficient of variation for Republican PACs for a particular ad spot requires at least two Republican PACs purchase the same ad spot. The set of ad spots used to construct the coefficient of variation therefore differs across comparison group. I recompute the estimates using the intersection of the three samples (Republican-Republican), (Democrat-Democrat) and (Republican-Democrat). For this sample, price dispersion can be calculated both within and across groups. The estimates are presented in column (2). The qualitative results are unchanged. In fact, the coefficient of variation across parties grows. A test for whether dispersion across parties is larger than dispersion within the Republican PAC group rejects the null of equality at 5% (the t-statistic is 8.07).

2.2 Does Party Favoritism Explain Pricing?

Stations may charge Republican and Democrat PACs different prices for reasons separate from differences in PAC willingness-to-pay. As an example, station owners may offer cheaper rates to their favored party. To investigate this possibility, I examine whether station owner

and employees' political donations are linked to ad prices, and in particular, whether stations with a clear bias in donations have a similar bias in pricing. Data on donations comes from the Federal Elections Commission by way of the Sunlight Foundation.[24] For each owner, I construct the percentage of donations given to Republicans compared to Democrats. To measure bias in pricing, I construct a price dispersion index for each ad product sold to both groups (again using the restricted sample), where p_D is the Democrat PAC price and p_R is the Republican PAC price.

$$\psi = \frac{p_{\mathrm{R}} - p_D}{2\left(p_{\mathrm{R}} + p_D\right)}$$

I then average this measure across ads sold by the same media company (across stations and week). A virtue of this index is that it measures price differences relative to the average cost of the spot.[25]A value of 0 corresponds to no discrimination, while values of the ψ close to 1 (-1) indicate a strong pro-Democrat -Republican) bias in pricing.

Figure 3 shows that across owners, Democrats receive more favorable rates than Republicans. Across the five companies, $\bar{\psi}$ ranges from .02 to .07, which corresponds to Republican PACs paying 4% to 15% more than their Democrat counterparts. Weigel Broadcasting, which is connected only to donations to Democrat affiliates, charges Republicans the largest markup. On the other hand, the Journal Broadcast Group, with gives 91% of donations to Republican causes, still charges Republican PACs 7% more. Even within ownership company, there is substantial variation in the Republican-Democrat price gap across ads. The standard errors for the estimated mean dispersion indices are large and clearly not statistically significant. Nonetheless, there is a negative correlation between relative donations to Republicans and the Republican - Democrat price gap that warrants further investigation using data on more media companies. Taken together, however, these results suggest observed price differences are not simply an artifact of station bias. This is consistent with Gentzkow & Shapiro (2010), who find that newspaper bias explains only a small part of media slant. Were rates set by the "most favorable" seller from a Republican point of view, figure 3 indicates that Republican PACs would still benefit disproportionately from legislation prohibiting discrimination across political advertisers.

[24]The Sunlight Foundation maintains a database named "Influence Explorer," which catalogues donations by individuals and political groups affiliated with each station's parent company. Available: <data.influenceexplorer.com/contributions>.

[25]Others (for example, Daivs *et al.* (1996) and Chandra *et al.* (2013)) use this transformation in a similar spirit to prevent a few, large observations from skewing the measure of dispersion (or growth).

3 Political Demand for Ad Spots

Apart from media bias, price differences might reflect differences in willingness-to-pay across political ad buyers. Political parties may target different audiences depending on their strategy (Nichter (2008)).[26] As an example, a vote-buying strategy involves persuading indifferent voters to cast their ballot for your candidate. In contrast, a turnout buying strategy requires persuading folks who prefer your favored candidate to show up at the polls. If both Democrats and Republicans attempt vote-buying, then they ought to value similar demographics and the same ad spots. However, if at least one party focuses on turnout-buying, then Democrat and Republican preferences over demographics should be very different. Pricing based on willingness-to-pay could also account for the observed price disparities within groups if PACs adopt different strategies.

To investigate whether stations price based on PAC willingness-to-pay for ad characteristics, I develop a model of demand for ad spots rooted in PACs' allocating resources to maximize the probability of winning. The first building block of the model specifies how advertising affects voting. The second step embeds this vote production function into the PAC ad choice problem given a finite budget for advertising, and explicitly models the demand for a particular ad spot. In section 3.3 and 3.4, I present an instrumental variable estimation strategy for dealing with price endogeneity that exploits state borders. Section 3.5 discusses a selection correction for dealing with unobserved prices. I present results in 3.6, including parameters governing party-specific taste for demographics.

3.1 Effect of Advertising on Voting

Let V_{gsc} be the share of group g that votes for candidate c in state s. V_{gsc} depends on ad exposures favoring candidate c, A_{gsc}, and the efficacy of own advertising, $\check{\gamma}_{gc}$. It also depends on opponent's advertising, $A_{gsc'}$, and the efficacy of his advertising $\tilde{\gamma}_{gc}$ (for example, if his advertising convinces some viewers to switch allegiance or to stay home on election day). The share of group g that votes for c also depends on the raw taste for the candidate β_{gsc}, and a random variable ε_{sc}. ε_{sc} induces aggregate uncertainty in voting outcomes, and is important in rationalizing advertising in states that are ex-post uncontested. Political actors do not know which is the tipping-point state, the state whose electoral college vote decides the national election.[27] Assume that these elements define a linear vote production

[26] Nichter (2008) details these strategies in the context of candidates or parties targeting benefits to particular constituencies in return for voting behaviors. I adopt his terminology to describe ad targeting, which is similar in spirit.

[27] In other words, the least favorable state their candidate must win to carry the national election. I borrow Nate Silver's estimates of tipping point probabilities from his New York Times blog. (Silver, Nate. 2012. "FiveThirtyEight Forecast." <NewYorkTimes.com>. November 6.)

function.

$$V_{gsc} = \check{\gamma}_{gc} A_{gsc} - \tilde{\gamma}_{gc'} A_{gsc'} + \beta_{gsc} + \varepsilon_{sc} \tag{1}$$

Since electoral college votes are awarded in a winner-take-all fashion, political advertisers care about producing votes only insomuch as it affects the probability their candidate wins a state's majority.[28] Their bottom line is the probability that S_{sc}, the share of state s that votes for candidate c, is larger than his rival's share $S_{sc'}$. S_{sc} is a function of π_{gj}, the probability a member of group g sees ad j, and f_{gs}, the fraction of s's population in group g.

$$S_{sc} = \sum_{g \in G} f_{gs} V_{gsc} = \varepsilon_{sc} + \sum_{g \in G} f_{gs} \beta_{gsc} + \sum_{g \in G} f_{gs} \left(\check{\gamma}_{gc} \sum_{j \in J_{cs}} \pi_{gj} - \tilde{\gamma}_{gc'} A_{gsc'} \right)$$

Candidate c's vote share aggregates baseline preferences and advertising effects across demographic groups, in proportion to their presence in state s. The probability that candidate c wins the state s therefore depends on the distribution of ε_{sc} and $\varepsilon_{sc'}$, own and rival's ad choices, and state demographics:

$$\begin{aligned} &\mathbb{P}\{S_{sc} \geq S_{sc'}\} \\ &= \mathbb{P}\left\{ \varepsilon_{sc} - \varepsilon_{sc'} \geq \sum_{g \in G} f_{gs}(\beta_{gsc'} - \beta_{gsc}) + \sum_{g \in G} f_{gs} \left((\tilde{\gamma}_{gc'} + \check{\gamma}_{gc'}) \sum_{j \in J_{c's}} \pi_{gj} - (\tilde{\gamma}_{gc} + \check{\gamma}_{gc}) \sum_{j \in J_{cs}} \pi_{gj} \right) \right\}. \end{aligned}$$

If I estimated (1) directly, then I could potentially estimate $\check{\gamma}_{gc}$ and $\tilde{\gamma}_{gc}$ separately (although individual-level voting data would be needed to estimate β_{gsc}). $\check{\gamma}_{gc}$ is the effect of candidate c's advertising on the proportion of the total population in state s and group g that votes for him. $\tilde{\gamma}_{gc}$ is the effect of c's advertising on his rival's share. Winning the state depends only on relative shares, so that candidates and PACs ultimately care about the sum of these two effects. Let $\gamma_{gc} = \check{\gamma}_{gc} + \tilde{\gamma}_{gc}$. γ_{gc} is the impact of c's advertising on the difference in shares between the two candidates. This paper infers buyers' demographic preferences using a revealed preference approach, so that only γ_{gc}, the net effect, is identified. Note that while the vote production function is linear in advertising, the share of votes cast in c's favor (the vote share) is not. The impact of advertising on candidate c's vote share depends on

[28] In Nebraska and Maine, votes are split among districts. (FEC Office of Election Administration. "The Electoral College." By William C. Kimberling. 1992.)

the stock of own and rival advertising.[29]

For tractability, let $\varepsilon_{sc} - \varepsilon_{sc'}$ distribute uniformly $[-\kappa, \kappa]$, so that winning is described by a linear probability model

$$\mathbb{P}\{S_{sc} \geq S_{sc'}\} = \kappa + \sum_{g \in G} f_{gs}(\beta_{gsc} - \beta_{gsc'}) + \sum_{g \in G} f_{gs}\left(\gamma_{gc} \sum_{j \in J_{cs}} \pi_{gj} - \gamma_{gc'} \sum_{j \in J_{c's}} \pi_{gj}\right).$$

The probability c wins state s is then an affine function of a weighted difference in ad exposures (since $\gamma_{gc} \neq \gamma_{gc'}$) and the difference between the raw taste for candidates. This specification of advertising technology exhibits constant returns to scale, which precludes interactions between ad spots in vote production, but greatly simplifies demand estimation. Decreasing returns are embedded in the model since candidates can buy at most one ad spot on each program on a station in a city. [30] This assumption is best-suited to ad choice in states where the margin between candidates is thin, so that the effect of advertising is plausibly locally linear. These are exactly the states with data for empirical study. Running ad j in support of c in state s changes the probability c takes the state by

$$\Delta_{jsc} = \sum_{g \in G} f_{gs} \pi_{gj} \gamma_{gc}.$$

To compare ads run in different states, I weight Δ_{jsc} to reflect states' relative importance. Winning a state is only important inasmuch as it influences the likelihood of winning the national election, and some states loom much larger in this calculation. A state's importance depends on its likelihood of being the tipping-point state, the least favorable state a candidate must win to collect 270 electoral college votes. For the 2012 election, Nate Silver conveniently calculated a tipping point index (τ_s) that gives the probability each state play this roll. This index combines two forces that determine a state's importance in a presidential election: first, the likelihood the state flips between red and blue, and second, the probability the national outcome hinges on the the state outcome. The tipping-point index rationalizes, for example, the dearth of campaigning in states like California or Texas with substantial heft in the electoral college. They have a low tipping-point index because the state outcome is a

[29] Let $\mathcal{V}_{sc}$ be the vote share of candidate c in state s.

$$\mathcal{V}_{sc} = \frac{\sum_g f_{gs} V_{gsc}}{\sum_g f_{gs}\left(V_{gsc} + V_{gsc'}\right)} = \frac{\sum_g f_{gs}\left(A_{gsc}\tilde{\gamma}_{gc} - A_{gsc'}\tilde{\gamma}_{gc'}\right)}{\sum_g f_{gs}\left(\gamma_{gc}A_{gsc} - \gamma_{gc'}A_{gsc'} + \varepsilon_{sc} - \varepsilon_{sc'} + \beta_{gsc} - \beta_{gsc'}\right)}.$$

[30] Gordon & Hartmann (2013) utilize decreasing returns to scale of political advertising, but the returns may actually be convex – for cash-constrained campaigns, we may even see advertising on the convex part of the function.

forgone conclusion.[31] In sum, the effect of ad j in support of c in state s is v_{jsc}:

$$v_{jsc} = \tau_s \Delta_{jsc} = \sum_{g \in G} f_{gs} \pi_{gj} \gamma_{gc}. \tag{2}$$

3.2 Ad Selection

The political advertiser employs (2) in choosing ads to maximize the probability her candidate wins, subject to a budget constraint B. Let p_{jstc} be the price of ad j run in state s at week t in support of candidate c. J_{sc} is the set of chosen ads. The optimization problem is described by:

$$\max_{\{J_{sc}\}_{s=1}^{S}} \mathbb{P}\{c \text{ wins the election}\}$$
$$\text{st:} \sum_{\{j \in J_{sc}\}_{s=1}^{S}} p_{jtsc} \leq B$$

If advertisers can buy fractional ads, optimal purchasing follows a simple decision rule. If ${\tau_s}^{\Delta_{jsc}}/p_{jstc} \geq \alpha_c$, then she should buy, where

$$\alpha_c = \max_{j \notin \{J_{sc}\}_{s=1}^{S}} \left\{ \frac{\tau_s \Delta_{jsc}}{p_{jstc}} \right\}$$

is the highest utility per dollar among ads not purchased.[32][33] In other words, buy ads in descending order of utility per dollar until the budget is exhausted. Purchased ads then obey this decision rule. α_c is naturally interpreted as the marginal utility of a political dollar. Although fractional purchases are permitted, this specification generates unit demand except for the marginal ad at the cutoff.

The unknown parameters of this model are the effectiveness parameters, $\{\gamma_{gc}\}_{g=1}^{G}$, and the shadow value of funds, α_c. To estimate these parameters, I incorporate two unobservable components into ad value: ϵ_{jstc}, known only to buyers, and ξ_{jstc}, known to buyers and sellers. The econometrician observes neither. ϵ_{jstc} introduces uncertainty, on the part of the station, as to exactly which ads political buyers value most, creating a downward sloping demand

[31] In states of the world where Texas or California changes hands, their electoral college votes are gratuitous (extraneous to winning).

[32] Without fractional purchases, set-optimization is challenging because it involves linear programming with integer constraints.

[33] Instrumental to developing a tractable demand model is the assumption that PACs take tipping-point probabilities as given. As an example, a PAC assumes that even if it poured resources into California, it could not change the probability that California is the decisive state in the national election.

curve. ξ_{jstc} accommodates the typical concern in demand estimation that stations and advertisers have information about ad spots reflected in prices and quantities, but hidden from the econometrician. An ad product is identified by j, the program name, s, the state where it airs, and t, the week it airs. The price of the product is buyer-specific, so it also has a subscript c. To recast the model using simpler notation, let x_{jst} be the observable characteristics of an ad and (β_c, α_c) be the taste parameters of the party supporting candidate c. Then this model of purchasing behavior can be described by the latent utility of each ad $jstc$:

$$u^*_{jstc} = x_{jst}\beta_c - \alpha_c p_{jstc} + \xi_{jstc} + \epsilon_{jstc}.$$

Let y_{jstc} be an indicator for purchasing using the cutoff decision rule.

$$y_{jstc} = 1\{u^*_{jstc} \geq 0\}. \tag{3}$$

If $\epsilon_{jstc} \sim U[-\Gamma, \Gamma]$, then (3) becomes a linear probability model

$$\mathbb{P}\{y_{jstc} = 1\} \quad = \quad \frac{1}{2} + \frac{x_{jst}\beta_c - \alpha_c p_{jstc} + \xi_{jstc}}{2\Gamma}.$$

3.3 Instrument for Price

In this section, I propose an instrument for price to facilitate estimation of the PAC demand parameters from the preceding section. The goal is to estimate separate parameters for Democrat and Republican PACs. Recovery of these preferences permits investigation of how observed prices relate to PAC willingness-to-pay.

The difficulty in estimating demand parameters is two-fold: first, prices are only observed for purchased ads, and second, those prices are potentially correlated with the unobservable ($\mathbb{E}[\xi_{jstc}|p_{jstc}] \neq 0$). Endogeneity is a concern if stations price using information about ad quality that is unknown to the econometrician.

Putting aside the first difficulty of transactions data, estimation requires an instrumental variable. To find a suitable instrument, I exploit a unique feature of presidential political advertising: its sensitivity to state borders. DMAs often straddle state lines, so that viewers in different states are bundled together into a single ad spot. Ads with out-of-state viewers ought to be more valuable (relative to the same ad run without these extra viewers) to run-of-the-mill TV advertisers, thus raising the opportunity cost of selling to a PAC. Viewership levels in uncontested states do not affect the value of an ad to a PAC, so the number of "uncontested" viewers, as a shifter of the residual supply curve, is an appropriate instrument for political demand.

The misalignment of media markets and political boundaries has been used to assess other questions in political media, however not in an explicit instrumental variable approach. As an example, Snyder Jr & Strömberg (2010) use the geography of newspaper markets to assess whether media coverage disciplines politicians.[34] Ansolabehere *et al.* (2001) investigate whether congressional advertising on television declines in districts with more incidental (uncontested) viewers.[35]

The analogous ideal experiment is random assignment both of the distance of a DMA to a state border and the distribution of demographics across that border. Then ads near borders with valuable neighbor demographics would have higher opportunity costs for reasons unrelated to their political value. This instrument varies both within and across DMAs, since uncontested viewership depends on show demographics, state demographics and borders. In my sample, there are seven DMAs that broadcast to viewers in contested and uncontested states: Boston, Cincinnati, Denver, Jacksonville, Philadelphia, Pittsburgh and Washington, DC. Across these DMAs, ads reach a ratio of 1.2 uncontested viewers for each contested viewer. Figure 5 shows the geography of DMAs in the sample which broadcast to both contested and uncontested viewers.

As an example, in the 2012 election, the Boston DMA received substantial advertising because ads broadcast in Boston reach not only Massachusetts, but also New Hampshire viewers. The exclusion restriction is that Massachusetts viewership does not directly enter the PAC demand specification. The relevance condition requires Massachusetts viewership enter the demand of other advertisers, so that shows broadcast in Boston with higher Massachusetts viewership have a higher opportunity cost.

The exclusion restriction is violated if PACs care about influencing other elections, either because they directly support candidates to other offices or if there are positive spillovers between presidential and congressional advertising. In that case, viewers in states where the presidential election is a foregone conclusion might be valuable if the senate seat is up for grabs. I therefore include viewership in states with close senatorial races as an explicit demand characteristic. The exogenous variation in price comes from variation in viewership in states where neither the senatorial nor presidential race is contested.[36]

[34] They find that higher congruence between political and market boundaries leads to more local political stories, better informed constituents, and changes in House representatives' behavior.

[35] They find that congresspeople in districts with more incidental viewers do not spend more on advertising, suggesting a strong, robust relationship between the price of airtime and purchasing behavior. I take the next step, and exploit this relationship in an IV specification.

[36] This assumption might be violated if PACs purchase ads in an effort to fundraise in uncontested states.

3.4 Estimating Equations

The final demand specification is estimated separately for Democrat and Republican PACs. An ad product is a week-hour-station-weekend combination, where weekend is an indicator for Saturday or Sunday airtime. Demographic groups include the number of viewers who are female, black, white, and over 65 years old. For each group, I include $f_{gs}\pi_{gj}$, the fraction of the state in demographic group g watching program j. Ad prices and demographic composition are measured per contested viewer.[37] k_{jsct} includes controls: week dummies, and priority level[38] fixed effects, and the proportion of viewers living in states with contested senate races.[39] All demographic variables are multiplied by viewers' average tipping-point probability τ_s. The following system describes demand

$$p_{jstc} = \phi_{0c} + \phi_{1c}\tau_s + \phi_{2c}z_{js} + \sum_{g=1}^{G} \tau_s f_{gs}\pi_{gj}\phi_{gc} + k'_{jstc}\phi_{3c} + \eta_{jstc} \tag{4}$$

$$y_{jstc} = \gamma_{0c} + \gamma_{1c}\tau_s - \alpha_c p_{jstc} + \sum_{g=1}^{G} \tau_s f_{gs}\pi_{gj}\gamma_{gc} + k'_{jstc}\gamma_{2c} + \epsilon_{jstc} \tag{5}$$

In practice, I use the two sample IV estimator from Angrist & Krueger (1995) with bootstrapped standard errors. I include predicted prices, which are fits from (4), in lieu of price on the right-hand-side of (5). I estimate standard errors using the nonparametric bootstrap, since predicted prices are generated regressors. For robustness, I re-estimate the model with daypart[40] fixed effects, with an eye toward eliminating unobserved ad quality. Adding these fixed effects means estimation exploits only within hour/week-segment variation.

3.5 Heckman Selection Correction

My estimation strategy so far ignores the selection problem inherent in transactions data: price is only observed for purchased ads. Censoring does not affect the estimation of the reduced form, but it means the first stage is estimated using only this sample. Shows with high draws of the instrument have higher prices, and correspondingly lower purchase probabilities. If I observe a high value of the instrument, I therefore ought to infer a low draw of the unobservable in the price equation. In the selected sample, this induces negative

[37] Normalizing by the number of viewers weighs ads equally. Otherwise, high markups on ads with low viewership and low markups on ads with high viewership are observationally equivalent, despite there different economic interpretations.

[38] For this part of the analysis, I restrict to four priority levels: p1, p2, p3+ and missing.

[39] I use RealClearPolitics classification of "toss up" senate races in 2012 to measure whether a seat was contested. States include: Indiana, Massachusetts, Montana, Nevada, North Dakota, Virginia, and Wisconsin.

[40] e.g. 8 PM Weekend or 6 AM Weekday

bias in the estimation of the covariance between price and the cost shock.

I can recast this inference challenge as the canonical problem of estimating labor supply: attempting to estimate the impact of wages (prices) on labor force participation (purchasing), where wages (prices) are only observed for those who choose to work (purchase). In this spirit, this demand system can be rewritten as functions of an observed price p_{jstc} and a latent price p^*_{jstc} that is only observed if $y_{jstc} = 1$.

$$p^*_{jstc} = x_{jst}\varphi_{1c} + z_{js}\varphi_{2c} + \eta_{jstc} \tag{6}$$

where z_{js} is the instrument, and the observed price is truncated.

$$p_{jstc} = \begin{cases} p^*_{jstc} & \text{if } x_{jst}\beta_c - \alpha_c p^*_{jstc} + \epsilon_{jstc} \geq 0 \\ . & \text{if } x_{jst}\beta_c - \alpha_c p^*_{jstc} + \epsilon_{jstc} < 0 \end{cases}$$

Heckman (1979) devised a selection correction assuming ϵ, η distribute jointly normal with covariance ρ . In this model

$$\begin{aligned} y_{jstc} &= 1\{x_{jst}\beta_c - \alpha_c p^*_{jstc} + \epsilon_{jstc} \geq 0\} \\ &= 1\{x_{jst}\beta_c - \alpha_c(x_{jst}\gamma_{1c} + z_{js}\gamma_{2c} + \eta_{jstc}) + \epsilon_{jstc} \geq 0\} \\ &= 1\{x_{jst}\pi_{1c} + z_{js}\pi_{2c} + \omega_{jstc} \geq 0\} \end{aligned}$$

where $\omega = \epsilon - \alpha\eta \sim N(0, \alpha^2\sigma^2_\eta + \sigma^2_\epsilon - 2\alpha\rho\sigma_\epsilon\sigma_\eta)$, and $\sigma_\omega = 1$ is the free scale normalization. This specification allows for price endogeneity through unobserved product quality.[41] Estimation using Heckman's two-step estimator permits recovery of the structural parameters: $\rho, \sigma_\epsilon, \sigma_\eta, \gamma_{gc}, \beta_c, \alpha_c$. Note that without an exclusion restriction on z, we cannot separately identify β_c and α_c. It is important that z enter the selection equation only through its effect on prices, so that $\hat{\alpha}_c = \frac{-\hat{\pi}_2}{\hat{\gamma}_2}$, and is just identified.

The joint normality assumption is less than ideal. The bivariate normal distribution may only poorly approximate the true distribution of unobserved PAC taste and cost shocks. A more serious concern is that the Heckman model specifies a structural pricing equation potentially inconsistent with firm behavior. Price in (6) is a linear function of observed characteristics and an unobservable cost shock that distributes joint normal with the demand-

[41] Stata estimates σ^2_η and $\rho_{\eta\omega}$, and lets $\sigma^2_\omega = 1$ as the scale normalization. We then need to rescale the structural selection parameters using the standard deviation of the structural error term σ_ϵ. We can recover σ_ϵ using the following two equations: $\sigma^2_\omega = \alpha^2\sigma^2_\eta + \sigma^2_\epsilon - 2\alpha\rho_{\epsilon\eta}\sigma_\epsilon\sigma_\eta$ and $\rho_{\eta\omega} = \frac{cov(\eta, \epsilon - \alpha\eta)}{\sigma_\eta\sigma_\omega}$. Then we can estimate the variance of the structural selection equation as: $\hat{\sigma}^2_\epsilon = 2\hat{\alpha}(\hat{\sigma}_\eta\hat{\rho}_{\eta\omega} + \hat{\alpha}\hat{\sigma}^2_\eta) - \hat{\alpha}^2\hat{\sigma}^2_\eta$. Note that this allows for correlation between η and ϵ, e.g. if there were unobserved (to the econometrician) product quality.

side taste shock. However, since selection is a serious concern with transactions data, the Heckman adjustment provides a sense of the magnitude of selection bias in this setting.

3.6 Evidence on Willingness-To-Pay for Democrat and Republican PACs

In this section, I discuss results about PAC preferences over demographics, which are presented in table 5a (Republicans) and 5b (Democrats).

Results from my baseline IV estimation strategy, equation (5), are reported in column 3. First, findings indicate that both Democrat and Republican PACs prefer viewers over 65 years old to their younger counterparts. Seniors have historically broken for Republicans, but polls leadings up to election day 2012 showed a tight race between Obama and Romney for their votes. Perhaps equally important, senior citizens are more likely to go to the polls than other age groups, so advertising to seniors might have a bigger bang-for-your-buck in terms of vote production.[42] Calculating the average marginal effect of a change in demographics on the probability of purchase requires some manipulation of the coefficients in table 5, since the right hand side variables measure the product of demographics and tipping point probabilities:

$$\begin{array}{c}\text{average marginal effect of a}\\ \text{1 std dev increase in \% g}\end{array} = \frac{\hat{\gamma}_{gc}\sigma_{\pi_{jg}}}{N}\sum_{s=1}^{N}\tau_s f_{gs}$$

A 5 percentage point (one standard deviation) shift in senior viewership increases the probability of purchase by 1.9 points for Republicans and 3 point for Democrats. Both parties also value women above men. An 8 point (one standard deviation) increase in the percent women increases the likelihood of purchase by 7% for Republicans and 1.9% for Democrats. Like senior citizens, women were more likely to be swing voters in the 2012 election.[43] Taken together, these preferences are consistent with parties employing a "vote buying" strategy.

Second, and perhaps unsurprisingly, Republican PACs prefer white viewers, who are valued least compared to blacks and other non-whites by Democrat PACs. A 10 point increase in percent white increases the likelihood of a Republican purchase by 7.2%, but decreases the probability of a Democrat purchase by 9%. These racial preferences suggest parties also employ a turnout buying strategy, where parties target their own bases. This

[42] Gentile, Olivia. 2012. "Whether for Obama or Romney, Senior Citizens Exercise Political Muscle." The Boson Globe. October 4.

[43] Berg, Rebecca. 2012. "Few Voters are Truly Up for Grabs, Research Suggests." The New York Times. August 16.

is consistent with evidence from Ridout *et al.* (2012) on targeting in the 2010 midterms elections. If stations price based on willingness-to-pay, then the prices paid by Republican and Democrat PACs should reflect the differences in their bases' demographics.

Preferences for demographics are stable across IV specifications: column (4) reports estimates including a full set of daypart dummies and column (5) reports coefficients with a Heckman selection correction. It is reassuring that these demand estimates are similar in magnitude and sign to the baseline two sample least squares estimates. Since the qualitative results are not sensitive to the selection correction or the additional fixed effects, in the remaining analysis, I proceed with the IV baseline specification.

This model cannot tease apart different explanations for these preferences. PACs may prefer women and seniors either because their underlying taste for candidates is more responsive to advertising or because their turnout is more responsive to advertising – or both. The model combines both forces in mapping ad impressions to voting outcomes. However, these estimates reveal that preferences over demographics matter, both economically and statistically.

Identification of PAC preferences across all specifications relies on uncontested viewership moving prices for reasons unrelated to political demand. Column (2) contains the first stage results for the baseline model, which corresponds to estimating equation (4). I find a strong positive correlation between uncontested viewers and prices, both for shows purchased by Democrat and Republican PACs. The sign is consistent with a model where prices reflect commercial demand. The F-statistics are 31.79 and 37.9 respectively, suggesting finite sample bias of these two stage least squares is small (Sock *et al.* (2002)).

The price coefficient in the second stage (column 3) is large and negative for both groups. In the baseline IV specification, Democrat demand elasticity (at the average ad program characteristics) is -1.28, and Republican demand elasticity is -0.969. In the absence of an instrument, there is no variation in the purchase dummy conditional on price, so there an OLS regression of purchasing on price and characteristics is not possible. The closest OLS specification merely shows the relationship between purchase probability and demographic covariates (column 1). Unsurprisingly, including price as right-hand-side variable flips the sign on several of the viewer demographic coefficients, underscoring the importance of the IV strategy.

4 Price Discrimination Model

The big picture question is whether observed prices positively correlate with willingness-to-pay. Using demand estimates (equation 5), I can measure willingness-to-pay for each

ad spot and recover the simple correlation for the sample of purchased ads. This simple test can provide suggestive evidence about how taste differences inform pricing decisions, but two factors confound a causal interpretation: marginal cost and unobservable quality. To illustrate how these combine if stations price based on buyer-specific taste for product characteristics (ad demographics), I develop a structural model of station behavior in sections 4.1 and 4.2. Section 4.3 creates machinery to test that model, which requires model-free estimates of markups and model-generated optimal markups for comparison. Results are presented in section 4.4.

4.1 Monopoly Pricing with Lowest Unit Rate Regulations

The first step in the supply-side analysis is a simple model of stations as single-product monopolists facing LUR regulations. This model informs the construction of bounds for marginal cost. Modeling marginal cost is important for testing whether observed prices are consistent with taste-based price discrimination. If marginal cost is negatively correlated with willingness-to-pay, then failing to account for it in a regression of price on willingness-to-pay would camouflage price discrimination. On the other hand, if marginal cost is positively correlated with willingness-to-pay, excluding costs could lead to false positives for price discrimination.

The marginal cost of an ad spot is opportunity cost – the highest price another advertiser is willing to pay for those 30 seconds. Intuitively, LUR rates, the lowest price for the spots that were purchased by campaigns, should approximate marginal costs well. This model formalizes that intuition. The equilibrium conditions suggest LURs as an upper bound for marginal cost.

In determining how much to charge a PAC with demand $P_{PAC}(Q_{PAC})$ for airtime, a TV station considers two other sources of demand for those same seconds: campaign $\tilde{P}(\tilde{Q})$ and other, non-campaign demand $P(Q)$ that might include other PACs. Non-campaign demand is relevant because there are only T seconds of potential advertising time per show. Since airtime is not sold in a posted price market, I model the station as perfectly price discriminating against non-campaign advertisers.[44] Campaign demand is separate because stations are constrained to sell campaigns ads at the lowest price they command on the market. The LUR regulation therefore forces stations to employ linear pricing schemes in their dealings with campaigns. One consequence is that stations may not exhaust their capacity, since selling additional units comes with a loss on inframarginal units sold to

[44]Stations sell most airtime in an upfront market each May. While they print "rate cards," stations negotiate package buys with each buyer, chiefly through media agencies (Phillips & Young (2012)). Price disparities across PACs further motivates the perfect price discrimination assumption.

campaigns. In sum, the station faces the following constrained optimization problem

$$\max_{\tilde{Q},Q} \quad \pi = \left(\int_0^{Q_{PAC}} P_{PAC}(q)dq\right) + \left(\int_0^{Q} P(q)dq\right) + \tilde{Q}\tilde{P}(\tilde{Q})$$

$$\begin{aligned} \text{st: } & P(Q) \geq \tilde{P}(\tilde{Q}) && \text{(LUR 1)} \\ & P_{PAC}(Q_{PAC}) \geq \tilde{P}(\tilde{Q}) && \text{(LUR 2)} \\ & T \geq Q_{PAC} + Q + \tilde{Q} && \text{(Capacity Constaint)} \end{aligned}$$

Since the station can perfectly price discriminate against PACs and commercial advertisers, $P^*_{PAC} = P^*$ in equilibrium. Therefore, either both LUR constraints bind or neither binds. Let $\pi_{\neg PAC}$ be the profits from sales to campaigns and other advertisers:

$$\pi_{\neg PAC} = \left(\int_0^{Q} P(q)dq\right) + \tilde{Q} \cdot \min\{P(Q), \tilde{P}(\tilde{Q}), P_{PAC}(Q_{PAC})\}.$$

The opportunity cost is the change in $\pi_{\neg PAC}$ from an increase in Q_{PAC}

$$-\frac{\partial \pi_{PAC}}{\partial Q_{PAC}} = -(P(Q)\frac{\partial Q}{\partial Q_{PAC}} + \tilde{P}(\tilde{Q})\frac{\partial \tilde{Q}}{\partial Q_{PAC}} + \tilde{Q}\tilde{P}'\frac{\partial \tilde{Q}}{\partial Q_{PAC}}). \tag{7}$$

Condition (7) simplifies depending on which constraints bind. If and only if the station sells positive quantities to a campaign, then the LUR binds. However, given data on $Q_{PAC}, P_{PAC}, \tilde{Q}, \tilde{P}$, the econometrician does not know whether the capacity constraint binds. Given this information constraint, I bound marginal cost above by lowest unit rates. I show this bound holds under the three sets of conditions that potentially describe equilibrium:

1. **Both constraints bind.** The CC implies ${}^{\partial Q+\tilde{Q}}/_{\partial Q_{PAC}} = -1$, so that (7) simplifies:

$$-\frac{\partial \pi_{PAC}}{\partial Q_{PAC}} = P(Q) - \tilde{Q}\tilde{P}'\frac{\partial \tilde{Q}}{\partial Q_{PAC}}$$

   Determining the exact marginal cost requires assumptions on non-political ad demand (to estimate $\frac{\partial \tilde{Q}}{\partial Q_{PAC}}$). Without imposing such assumptions, I can bound the marginal cost in the following fashion:

$$P(Q) > -\frac{\partial \pi_{PAC}}{\partial Q_{PAC}} > P(Q) + \tilde{Q}\tilde{P}'(\tilde{Q})$$

   Lowest unit rates overestimate marginal cost, since selling more units leads to infra-marginal losses on units sold to campaigns. Based on estimates of campaign demand

(tables 5a and b), $\tilde{P}'(\tilde{Q})$ is small , so that the upper bound ought to be close to the true marginal cost.

2. **Only the lowest unit rate rule binds**. Selling additional units to the PAC forces stations to lower LURs, which means inframarginal losses on units sold to campaigns. Marginal cost is less than the lowest unit rate since $\partial Q+\tilde{Q}/\partial Q_{PAC} \leq -1$.

3. **Only the capacity constraint binds.** In this case, candidate demand is relatively low compared to other advertisers so that $\tilde{Q} = 0$. The equation for opportunity cost (7) becomes $-\frac{\partial \pi_{PAC}}{\partial Q_{PAC}} = P(Q)$, which is exactly the LUR. However, this rate is unobserved since campaigns do not purchase any ads. It is possible that $Q_{PAC} = 0$ if PAC demand for that particular ad is also very low.

4. **Neither constraint binds.** This case never occurs so long as advertising has non-negative returns (and disregarding the disutility of viewers). If the LUR rule does not bind, that means campaigns are not purchasing airtime. At the very least, non-political advertisers and PACs should have positive value for airtime, and since stations can perfectly price discriminate across units sold to these buyers, they should sell all of their airtime.

This model illustrates that lowest unit rates are a good proxy for marginal cost, albeit upper bounds. In the next section, I develop estimating equations based on the intuition from this model. In the final section, I incorporate LURs as marginal costs and explicitly test the stations' first order conditions.

4.2 Station's Optimal Pricing Condition

In this section, I adapt the continuous model to a discrete setting where the firm sells a single indivisible unit of each product. This model is the simplest that permits examination of price discrimination, the phenomenon of interest, but it may assign too much market power to stations. Since I have not imposed supply-side behavior in estimating demand, I can test the monopoly assumption jointly with the demand estimates. If the model poorly approximates true station behavior – because stations lack market power, demand estimates are incorrect, or pricing does not reflect PAC willingness-to-pay for demographics – then observed prices will be inconsistent with the monopolist's FOC for pricing ad product (jst) to a PAC supporting c:

$$
\begin{aligned}
p^*_{jstc} &= \underset{p_{jstc}}{\operatorname{argmax}} \quad (p_{jstc} - c_{jst})(1 - F_\epsilon(-(x_{jst}\beta_c + \xi_{jstc} - \alpha_c p_{jstc}))) \\
\implies p^*_{jstc} - c_{jst} &= \frac{1 - F_\epsilon(-(x_{jstc}\beta_c + \xi_{jstc} - \alpha_c p^*_{jstc})))}{\alpha_c f_\epsilon(-(x_{jstc}\beta_c + \xi_{jstc} - \alpha_c p^*_{jstc}))}. \qquad (8)
\end{aligned}
$$

This FOC ignores income effects by setting $\frac{\partial \alpha_c}{\partial p_{jstc}} = 0$. This assumption is standard in the IO literature for goods like ads that constitute but a small expenditure share of the budget (adding these effects restores complementarity between ad purchase decisions and greatly complicates both demand estimation and the pricing model). Essentially, I assume stations ignore cross-price elasticities. They assume that raising prices on a single ad has a negligible effect on demand for other ad buys. I also assume stations take tipping-point probabilities as given. This places the model somewhere on the spectrum between perfect competition and monopoly. These assumptions are most suspect when considering counterfactuals where the price of airtime may rise across the board, but a substantial discrepancy between observed and predicted prices from a model without income effects would suggest taste-based discrimination is unlikely to play an important role in this market.

This model incorporates three reasons for observed price differences between Democrat and Republican PACs: different marginal utilities of money (α_c), different values for the same demographics (γ_{gc}), and different values for other ad characteristics. It also points to another reason that Republican PACs pay higher prices on average: Republican PACs may purchase higher cost ads. Since the set of ad-products purchased by both parties is a selected sample, understanding the cost-side is key for drawing conclusions about the winners and losers under the current regulatory regime. To be clear, if differences in ad purchase decisions account for the lion's share of the difference in expenditures, then banning price discrimination across PACs ought to have but a small affect on the market. Conversely, if pricing is driven primarily by willingness-to-pay, such regulation would have real bite.

Imposing ϵ_{jstc} distributes uniformly simplifies the FOC (8), so that it is separable in the cost and preference-driven components of price:

$$
p^*_{jstc} = \frac{\Gamma}{2\alpha_c} + \frac{x_{jstc}\beta_c}{2\alpha_c} + \frac{c_{jst}}{2} + \frac{\xi_{jstc}}{2\alpha_c} \qquad (9)
$$

4.3 Testing Station Optimization

To examine whether prices reflect PAC willingness-to-pay, I develop a series of tests based on the TV station first order condition (8). As a first pass, I regress the observed price on

estimated utility per dollar separately for Democrats and Republican PACs. Willingness-to-Pay for each group is constructed using the demand parameters ($\hat{\beta}_c$ and $\hat{\alpha_c}$) estimated via (5)

$$\begin{aligned} \hat{u}_{jstc} &= \frac{x_{jst}\hat{\beta}_c}{\hat{\alpha}_c} \\ p_{jstc} &= \gamma_0 + \gamma_1 \hat{u}_{jstc} + \epsilon_{jstc}. \end{aligned} \tag{10}$$

This regression does not so much constitute a test of the particular monopoly model I propose as a test of whether prices reflect preferences. If yes, the estimate of γ_1 ought to be large, positive and statistically significant.

If marginal costs are small and there is limited variation in unobserved quality, then (10) also constitutes a test of the structural model (9). However, marginal cost is usually assumed to rise with quality. In this market, if commercial advertisers and PACs value similar characteristics, then marginal cost ought to be positively correlated with PAC willingness-to-pay. Here, I employ LURs as a measure of marginal cost and re-estimate the first order condition including this term. To test the model, I test the null $H_0 : \gamma_1 = \frac{1}{2}$, where γ_1 is the coefficient on the "taste" component of pricing

$$p_{jstc} = \gamma_0 + \gamma_1 \frac{x_{jst}\hat{\beta}_c}{\hat{\alpha}_c} + \gamma_2 c_{jstc} + \eta_{jstc}. \tag{11}$$

OLS estimation of (11) is still potentially biased due to selection on unobservables. If stations price according to the monopoly model, then the residual in (11) is a function of unobserved ad quality: $\eta_{jstc} = \frac{\xi_{jstc}}{2\alpha_c}$. Price is only observed conditional on purchase, so that $cov(\eta_{jstc}, \frac{x_{jst}\beta_c}{\alpha_c}) \leq 0$ in this sample (though not the population). Intuitively, if a PAC purchases an ad spot with poor observables, then that spot must have a high draw of the unobservable. This means OLS underestimates γ_c. The conditional expectation of p_{jstc} given c purchases an ad with characteristics x_{jst} is:

$$\mathbb{E}[p_{jstc}|y_{jstc} = 1, x\beta, \alpha] = \frac{\Gamma}{2\hat{\alpha}_c} + \frac{x_{jst}\hat{\beta}_c}{2\hat{\alpha}_c} + \frac{c_{jst}}{2} + \frac{\mathbb{E}[\xi_{jstc}|y_{jstc} = 1]}{2\hat{\alpha}_c}.$$

I can estimate the expectation of the omitted quality term if I specify a distribution for ξ_{jstc}. Let $\xi_{jstc} = \sigma_\xi \tilde{\xi} \sim N(0, \sigma_\xi^2)$. I model the CEF of ξ_{jstc} conditional on observables x_{jst}, estimated demand parameters α_c, β_c, costs c_{jst}, and purchase at the optimal price. (Conditioning on

the observed price is not possible, since observed price is the dependent variable).

$$\begin{aligned}
\mathbb{E}[\xi_{jstc}|y_{jstc}=1] &= \sigma_\xi \frac{\int_{-\infty}^{\infty} \tilde{\xi}(\Gamma + x_{jst}\hat{\beta}_c + \sigma_\xi\tilde{\xi} - \hat{\alpha}_c(\frac{\Gamma}{2\hat{\alpha}_c} + \frac{x_{jst}\hat{\beta}_c}{2\hat{\alpha}_c} + \frac{\sigma_\xi\tilde{\xi}}{2\hat{\alpha}_c})\phi(\tilde{\xi})d\tilde{\xi}}{\frac{1}{2}(x_{jst}\hat{\beta}_c + \Gamma)} \\
&= \frac{\sigma_\xi}{2}\frac{\int_{-\infty}^{\infty} \tilde{\xi}(\Gamma + x_{jst}\hat{\beta}_c\phi(\tilde{\xi})d\tilde{\xi}}{\frac{1}{2}(x_{jst}\hat{\beta}_c + \Gamma)} + \frac{\sigma_\xi^2}{2}\frac{\int_{-\infty}^{\infty} \tilde{\xi}^2\phi(\tilde{\xi})d\tilde{\xi}}{\frac{1}{2}(x_{jst}\hat{\beta}_c + \Gamma)} \\
&= \sigma_\xi \int_{-\infty}^{\infty} \tilde{\xi}\phi(\tilde{\xi})d\tilde{\xi} + \frac{\sigma_\xi^2}{x_{jst}\hat{\beta}_c + \Gamma} \\
&= \frac{\sigma_\xi^2}{x_{jst}\hat{\beta} + \Gamma}
\end{aligned}$$

Then I can test the FOC as:

$$p_{jstc} = \gamma_0 + \gamma_1\left(\frac{x_{jst}\hat{\beta}_c}{2\hat{\alpha}_c}\right) + \gamma_2\left(\frac{\hat{c}_{jst}}{2}\right) + \gamma_3\left(\frac{1}{2\hat{\alpha}_c(x_{jst}\hat{\beta}_c + \hat{\Gamma})}\right) + \omega_{jstc} \tag{12}$$

So far, the proposed tests of station behavior compare observed PAC-specific prices to measures of PAC valuation. They differ in the set of controls. A second variety of test compares Republican-Democrat PAC price differences to predicted price differences. This comparison requires no marginal cost or quality estimates above an assumption that these are independent of party affiliation.[45] The test specification is:

$$p_{jstR} - p_{jstD} = \gamma\left(\frac{x_{jst}\beta_R}{\alpha_R} - \frac{x_{jst}\beta_D}{\alpha_D}\right) + \psi_{jst} \tag{13}$$

The null hypothesis remains $H_0 : \gamma = \frac{1}{2}$.

4.4 Do Prices Reflect Willingness-to-Pay?

Table 6 reports the results from the first set of price discrimination tests. Columns (1) and (4) report the correlation between observed price and estimated utility (both measured in dollar terms) for Republican and Democrat PACs respectively. This specification corresponds to estimating equation (10). For both groups, the estimated coefficient is large, positive, and statistically significant at conventional levels. The coefficient is 0.67 for Democrats and 0.62 for Republicans, indicating price rises 1.2:1 with willingness-to-pay for both groups. While the difference in coefficients is statistically significant, it is economically negligible. Stations seem to extract rent from both political parties to a similar extent. Figures 6a and 6b show

[45] This would be a poor assumption, for example, if viewership (and ratings) are responsive to political advertiser identity.

this relationship graphically. I group observations into 20 bins by percentile of estimated utility, and plot each bin against its average price. The relationship appears strikingly linear.

Both the Democrat and Republican coefficients on willingness-to-pay are larger than predicted by the monopoly model. I can reject the null that the coefficient on is 0.5 for both groups at the 5% level. A positive correlation between utility and cost could cause an inflation of the coefficient estimate, and explain rejection of the model.

Columns (2) and (5) control for marginal cost using lowest unit rates, which corresponds to equation (11). The coefficients on willingness-to-pay are closer to the model's predictions. I cannot reject the null that the each coefficient is 0.5. The coefficients on cost, however, are smaller than theory indicates, which dovetails with lowest unit rates as upper bounds for marginal cost. As a robustness check, I estimate test specification (12) which includes a proxy for unobserved utility. Columns (3) and (6) present the results. Controlling for unobserved quality has almost no effect on the point estimates for the coefficient on willingness-to-pay, suggesting the variance in unobservable ad quality is small.

Table 7 reports results for the second set of tests, which compare observed price differences to estimated utility differences. Price disparities are a prime motivator for concern about discrimination, so a stringent test of the model is whether it can replicate this facet of the data. Column (1) reports the results of this test for the full set of ad-products where both Republicans and Democrats purchase. A $1 increase in Republican over Democrat utility per viewer corresponds to a $0.28 price hike for Republican versus Democrat PACs. Importantly, this test requires fewer assumptions on the cost side, since it lives only off of price differences. This small point estimate may be an artifact of the sample, since ad products are defined loosely as airtime at the same hour, station, and week. As an example, price differences may reflect cost differences between high and low priority purchases, rather than utility differences for the same level of priority. Column (2) restricts the sample to indistinguishable goods, where priority level and show name must be an exact match. Reassuringly, the coefficient estimate increases to a $0.61 price increase per dollar of utility.

Selection remains a concern because I can only perform this test conditional on a purchase. The ideal regression would have differences in offered prices as the dependent variable, rather than differences in purchase prices. The relationship between purchased price and utility may be attenuated if PACs are more likely to purchase shows where stations underprice. As a final test, I consider this relationship for the set of ad spots where stations actively price discriminate. In other words, I drop observations where Republicans and Democrats pay the exact same price (approximately half of the observations). The results indicate a $0.79 increase in price difference per $1 increase in utility difference (results reported in column (4)). For this restricted sample, I cannot reject the null hypothesis that the monopoly

model is true (that a 1:1 relationship between utility and price differences hold).

Taken together, these results indicate a robust relationship between buyer-specific taste for demographics and prices. Stations seem to be getting prices "right" by charging buyers more for more-desired demographics. Although other forces undoubtedly factor into the political ad market, including bundling and bargaining, my results suggest the monopoly model approximates station behavior fairly well.

5 Quantity Withholding in Response to Lowest Unit Rate Regulation

In this section, I extend the monopoly model to explore whether lowest unit rate regulation distorts stations' behavior. The goal of this model is to estimate the effect of lowest unit rate regulation on advertising quantities. Theory indicates that stations may sell below their capacity constraint in response to lowest unit rate rules, i.e. withhold quantity to drive up lowest unit rates.[46]

Duggan & Morton (2006) document similar distortions in response to regulation tying Medicaid prices to average price in the private market. They examine prices for drugs serving a large Medicaid population compared to drugs that target a non-Medicaid clientele. Since I do not observe outcomes in the commercial market directly, I adopt a different estimation strategy. Commercial quantities are necessary to measure how far equilibrium quantities fall short of capacity constraints. If I model station optimization and campaign and commercial demand, however, I can find the commercial quantities that best rationalize the data. I can then extrapolate the extent to which stations engage in quantity withholding. This sort of structural approach can be useful in a number of settings where government policies interact with imperfect competition in potentially unanticipated ways. To be clear, if stations had no market power, then there would be no strategic response to regulation. In markets like TV advertising, however, these distortions are potentially large.

Stations employ a number of tools to alter airtime: trim regular programming; replace network shows with other programs that have a higher ratio of commercials to content; cut ads promoting upcoming shows; and inserting local ads in place of network airtime.[47] I assume stations have 13.5 minutes of airtime available each hour, which they can sell

[46] Declining airtime during election years may seem counterintuitive. A model with disutility of advertising for voters could reconcile quantity withholding with more advertising time during election years, since stations might not exhaust their capacity absent political advertising. Election year would mean an increase in advertising, but a smaller increase than in a counterfactual absent regulation.

[47] Far, Paul. 2012. "Dilemma for D.C. Stations: So Many Political Ads, So Little Airtime." The Washington Post. October 22.

to the commercial market (including the network) or to political campaigns.[48] This hard constraint can be thought of as technological, i.e. the maximum airtime for advertising if a station deploys all of its tricks.[49]

In section 5.1, I develop a model of station optimal airtime allocation between the commercial and campaign markets. Section 5.2 maps this to an empirical demand specification, and section 5.3 delineates a Bayesian MCMC estimation procedure to back-out the demand parameters. Results are presented in section 5.4, including estimates of quantity withholding for the 2012 election cycle.

5.1 Station Quantity Decisions

In determining how to set LURs, a TV station considers both campaign $\tilde{P}(\tilde{Q})$ and commercial demand $P(Q)$ (that might include PACs). I assume that there are a maximum of T units of advertising per hour of television. Since advertising has negative externalities on viewers (see for example Bagwell (2007),Anderson & Gabzewikz (2006)), I set $T < M$, where M is the total amount of airtime. Again, I model the station as perfectly price discriminating against commercial advertisers. Campaign demand is separate because stations are constrained to sell campaigns ads at the lowest price they command on the market. The LUR regulation therefore forces stations to employ linear pricing schemes in their dealings with campaigns. One consequence is that stations may not exhaust their capacity, since selling additional units comes with a loss on infra-marginal units sold to campaigns. Absent regulation, a perfectly price discriminating station sells its entire capacity. In sum, the station faces the following constrained optimization problem:

$$\max_{\tilde{Q},Q} \pi = \int_0^Q P(q)dq + \tilde{Q}\tilde{P}(\tilde{Q})$$

$$\text{st: } T \geq Q + \tilde{Q} \qquad \text{(Capacity Constraint)}$$

$$P(Q) \geq \tilde{P}(\tilde{Q}). \qquad \text{(LUR)}$$

Three conditions potentially describe the optimal LUR, depending on whether there is an interior or boundary solution:

[48] This number is taken from Ad Week estimates of broadcast airtime: http://www.adweek.com/news/television/you-endure-more-commercials-when-watching-cable-networks-150575

[49] In theory, a station could dedicate all airtime to advertising by eschewing network programming, except that would harm viewership. This hard constraint on advertising time embeds this viewership response. If stations advertise more than 13.5 minutes per hour, then viewership plummets and airtime is useless for advertisers.

1. Only the capacity constraint binds: in this case, the station sells only to commercial advertisers, and the lowest unit rate is above the campaigns' willingness-to-pay for the first unit.

$$\begin{aligned} P^{LUR} &= P(T) \geq \tilde{P}(0) \\ \implies \pi^* &= \int_0^T P(q)dq \end{aligned}$$

If the campaign purchases zero units, I assume this condition describes the equilibrium.

2. Both the LUR and capacity constraints bind: in this case, the constraints perfectly determine the lowest unit rate.

$$\begin{aligned} P(Q^*) &= \tilde{P}(T - Q^*) \\ \implies \pi^* &= \int_0^{Q^*} P(q)dq + (T - Q^*)P(Q^*) \end{aligned} \tag{14}$$

3. Only the LUR binds:

$$\begin{aligned} \max_P \quad & P\tilde{Q}(P) + \textstyle\int_0^{Q(P)} P(q)dq \\ \text{FOC:} \quad & \tilde{Q}(P) + P\tilde{Q}'(P) + PQ'(P) \quad = 0. \end{aligned} \tag{15}$$

In this case, LUR regulation induces inefficiency, since too few ads are sold both to campaigns and non-political advertisers (ignoring the disutility of viewers, a first best allocation implies the capacity constraint binds).

5.2 Empirical Demand Specification

Consider a station selling T_i spots on a program with observed characteristics X_i. p_i is the price of the show, which I observe if at least one campaign purchases an ad spot. A market is a week-station-daypart combination.[50] This leaves approximately 2,480 products in my markets, since sales more than sixty days before the election are excluded.

To estimate the magnitude of efficiency loss (how often stations price according to (15) rather than (14)), I impose structure on the demand functions for each product. Campaigns demand quantity $M_i\tilde{s}_i$ of the M_i units of airtime on program i with observable characteristics

[50] Dayparts include early (5am-9am), daytime (9am-5pm), news (5pm-7pm), primetime (7pm-11pm) and late night (11pm-5am).

$\tilde{x}_i$, price p_i, and unobservable quality ξ_i. M_i is 120 times the length of the program in hours, of which T_i are potentially available for advertising. $\tilde{x}_i$ includes the characteristics from the baseline linear model (5): viewer pivotality, four demographic groups, priority level, and week aired. $\tilde{s}_i$ is modeled using the logit share function.

$$\tilde{s_i} = \frac{\exp\{\tilde{x}_i\tilde{\beta} - \tilde{\alpha}p_i + \xi_i\}}{1 + \exp\{\tilde{x}_i\tilde{\beta} - \tilde{\alpha}p_i + \xi_i\}}$$

Given a candidate parameter vector, an observed campaign share and price, the unobservable is:

$$\xi_i = \ln\left[\frac{\tilde{s}_i}{1 - \tilde{s}_i}\right] - \tilde{x}_i\tilde{\beta} + \tilde{\alpha}p_i.$$

Commercial advertisers value a larger set of covariates x_i, which include the number of viewers in uncontested states and demographics. Commercial advertisers also have a different unobservable utility component ω_i. Differences between the characteristics in x_i and $\tilde{x}_i$ help disentangle covariance between unobservables and common taste for observable characteristics. The share purchased by commercial buyers is

$$\mathrm{s}_i = \frac{\exp\{x_i\beta - \alpha p_i + \omega_i\}}{1 + \exp(x_i\beta - \alpha p_i + \omega_i\}}. \tag{16}$$

Assume that ξ_i, ω_i distribute bivariate normal with variances σ^2_ξ, σ^2_ω and covariance ρ. This share function can be micro-founded in a model where commercial advertisers value each unit j on the show at $v_{ij} = x_i\beta - \alpha p_i + \omega_i + \eta_{ij}$, where η_{ij} are distributed type I extreme value.[51] In this model, given a price p_i, commercial and campaign demand may exceed capacity (the shares need not sum to one).

5.2.1 Zero Shares

Empirically, there are many instances of zero shares – 1,330 zero shares in 2,480 markets – which hinders inversion to find mean utilities. Rather than faulting the expected logit share as a poor approximation to its empirical counterpart, I approach this as a missing data problem. Let $\tilde{q}_i$ be the observed quantity. I do not observe the true quantity when it

[51] Rather than the standard paradigm, where M consumers each decide whether or not to purchase a single unit, in this scenario, a single consumer (consider commercial advertisers as a single unit) decides whether to purchase or decline M times.

falls below the single-unit threshold:

$$\tilde{q}_i = \begin{cases} M_i \tilde{s}_i & \text{if } M_i \tilde{s}_i \geq 1 \\ 0 & \text{if } M_i \tilde{s}_i < 1 \end{cases}$$

A key distinction between this approach and alternative methods for handling zero shares (e.g. Gandhi *et al.* (2013)) is that the econometric difficulty does not stem from too few consumers relative to the number of products in a market. In that scenario, a zero realized share might mask a very high expected share. Rather, the difficulty here is that the data is 'binned' after it is generated. In contrast, a zero share, in my setting, rules out the possibility that the expected share was higher than $\frac{1}{M_i}$.

5.2.2 Correcting for Unobserved Commercial Quantity

Given campaign prices and demand parameters, the quantity sold to commercial advertisers takes two potential values – either the entire residual supply or a quantity defined by the firm's first order condition (15). I do not know which of these two conditions defines the optimum based on the observed data, but I know that one must hold. Both of these potential quantities implies a mean utility, and corresponding "supply" (commercial demand) shock.

1. If residual supply is exhausted (the station hits its capacity constraint), then I can back-out the commercial demand shock for the commercial market, ω_i^B, as a function of observed campaign share $\tilde{s}_i$ and parameters.

$$\begin{aligned} T_i &= M_i \tilde{s}_i + M_i s_i^B \\ s_i^B &= \frac{T_i}{M_i} - \tilde{s}_i \\ \implies \quad \delta_i^B &= \ln\left[\frac{s_i^B}{1 - s_i^B}\right] \\ \omega_i^B &= \delta_i^B - x_i\beta + \alpha p_i \end{aligned}$$

2. If the capacity constraint is not binding, then I recover the commercial demand shock, ω_i^I, from the first order condition that corresponds to (15)

$$\begin{aligned} M_i \tilde{s}_i - p_i M_i \tilde{s}_i \tilde{\alpha}(1 - \tilde{s}_i) &= \alpha M_i p_i s_i^I (1 - s_i^I) \\ \implies \delta_i^I &= \ln\left[\frac{s_i^B}{1 - s_i^B}\right] \\ \omega_i^I &= \delta_i^I - x_i\beta + \alpha p_i. \end{aligned}$$

I can then evaluate the likelihood of each draw ((ξ_i, ω_i^I) or (ξ_i, ω^B)) using the bivariate normal probability density function and a posited covariance matrix.

5.3 Bayesian Estimation Strategy

Identification of the preference parameters comes from three sources: an exclusion restriction in the campaign demand function, the stations' first order condition, and a joint normality assumption on the commercial and campaign unobservable. The exclusion restriction is identical to the PAC demand model I estimate in section 4, which permits recovery of the political advertiser preferences. Using the stations' first order condition lets me recover the unobserved commercial quantity, which is the object of interest. The parametric assumption on the distribution of campaign and commercial taste shocks is then necessary to correct for the selection bias in transactions data, analogous to the Heckman selection correction in section 4.5.

Taken together, these assumptions allow me to construct a likelihood function similar in spirit to a Tobit model. However, there are two aspects of the resulting likelihood function that preclude standard maximum likelihood estimation. First, the probability of a zero share must be simulated. In principle, this difficulty can be dealt with using maximum simulated likelihood techniques, but the finite-sample properties of MSL estimators have serious drawbacks. For consistency, the number of simulations must grow faster than the number of observations (Train (2009)). Since simulations are computationally expensive (they involve root-finding), in practice, the number of draws per observation is constrained. A more serious concern is that the likelihood function is not differentiable in the parameter space. For some parameter vectors, the inversion of the station's first order condition, which amounts to solving the quadratic equation, produces complex numbers for commercial shares, inducing a discrete jump from positive probability to zero probability in the likelihood function.[52] Gradient-based optimization is therefore quite tricky (the gradient may not exist). Instead, I implement a Bayesian Markov Chain Monte Carlo estimation procedure using a Metropolis-Hastings algorithm with a random walk.[53] Each step of the Markov chain requires a Monte Carlo integration of the probability of censoring in my data. See the appendix for a detailed description of the likelihood function and integration procedure.

The statistic of interest is the amount of airtime stations withhold from the market to bolster lowest unit rates. The total amount of airtime available for sale at stations in the

[52] For some values of the parameters, there are no draws of the unobservables that rationalize the data. Rather than assign zero probability to those parameter values, I penalize the likelihood function by 10^{-x} where $x = 10 + 10 \cdot \text{fraction unrationalizable}$. In practice, this amounts to approximately 3% of observations.

[53] I use a flat prior and a normal proposal density. I adjust the variance of the normal to regulate the acceptance probability to be between 0.25 and 0.4.

sample is $A = \sum_{i=1}^{N} T_i$. I calculate the fraction of airtime unsold, L, for each MCMC step

$$L = \frac{1}{A}\sum_{i=1}^{N}\left(T_i - M_i\tilde{s}_i - M_i s_i^I\right)\frac{\mathbb{P}\left\{\omega_i^I, \xi_i\right\}}{\mathbb{P}\left\{\omega_i^I, \xi_i\right\} + \mathbb{P}\left\{\omega_i^B, \xi_i\right\}}$$

and estimate $\hat{L}$ as the posterior mean of the distribution.

5.4 Evidence on Quantity Withholding

Table 8 presents parameter estimates and credible interval from a random-walk metropolis chain of 200,000 draws with a burn-in of 50,000 draws. The goal of the model is to estimate the distortionary effects of LUR regulation on the total amount of airtime sold. The posterior mean of the quantity withholding is 7.52% of total available advertising time. The credible interval (analogous to a 95% confidence interval) extends from 7.39% to 7.62%, so the estimate is fairly precise. These findings suggest the distortionary effects are of first-order importance in evaluating LUR regulation.

The estimated covariance between the commercial and campaign taste shocks is positive, consistent with an unobserved quality dimension valued by both groups of advertisers. As an example, both commercial and campaign advertisers might prize primetime shows, which may attract viewers who are otherwise hard to reach (Phillips & Young (2012)). The estimated variance parameters are large. These parameters rationalize the variation in campaign shares across ad products in the data. Campaigns have no observed purchases for most ad products, but for a small subset, they purchase a large share of the inventory. This model explains this pattern through high and low draws of the unobservable.

The campaign price coefficient is smaller in magnitude (-11.23) than its commercial counterpart (-183.26). A shallower demand curve is consistent with campaigns' having limited alternative advertising opportunities relative to commercial advertisers. Campaigns prioritize tipping-point DMAs in a small time window – the months preceding the election – relative to commercial advertisers who are not beholden to the peculiarities of the electoral college. Although campaigns are less price sensitive, the estimated campaign intercept is much smaller than its commercial counterpart; commercial advertising still swamps political airtime.

Apart from the price coefficient, parameter estimates suggest campaigns value viewers in states more likely to play the tipping-point roll and also older viewers. These preferences are consistent with the results on PACs presented in section 3. Campaigns value black viewers above whites, which is also consistent with PAC preferences. Since the Obama campaign

accounted for the lion's share of campaign buying in 2012, an aggregate campaign preference for blacks is in line with Obama employing a get-out-the-vote strategy.

6 Conclusion

Since Lyndon B. Johnson's infamous "Daisy" commercial aired in 1964, industry wisdom holds that paid TV advertising is necessary to a successful political campaign and, since 1934, the FCC requires television stations to sell airtime to all official campaigns at the same price – in fact, at lowest unit rates (West 2010). Regulation advocates fear that, without restrictions, campaigns might face different prices, leading to large – and unfair – discrepancies in media presence. This paper examines station treatment of Political Action Committees, not subject to such restrictions, to shed light on whether, and to what extent, such fears are well-founded.

To be clear, PACs loom large on the political advertising scene – spending neared $500 million in the 2012 presidential race – because campaign finance regulations require large donations go through PACs.[54] Importantly, stations have a free hand in their dealings with PACs, and their pricing decisions have direct consequences for inequalities in political speech. Further, the prices PACs pay can guide our expectations about prices official campaigns would pay absent regulation.

Novel data on ad-level prices reveals two stylized facts. First, PACs pay substantial markups above regulated rates. Since PACs face higher prices, a candidate should prefer donations come through his official campaign. When campaign finance regulation diverts funds to PACs, the candidate gets a lower bang-for-his-buck. A candidate's ad purchasing power, therefore, depends on the distribution of donation dollars across his supporters. Second, stations charge Democrat and Republican PACs different prices for indistinguishable ads. Price differences have several potential causes: station owner bias, viewer preferences over parties, differences in purchase timing, and PAC willingness-to-pay for ad characteristics, to name a few. I find little evidence that media bias, measured using data on political donations, drives pricing. Rather, findings indicate that prices (and price differences) reflect each party's preferences for viewer demographics.

To recover PAC willingness-to-pay for different viewers, I develop a model of demand for advertising spots and estimate preference parameters separately for Democrats and Republicans. To mitigate concerns about price endogeneity, I exploit the sensitivity of political demand to state borders. Viewership in uncontested states constitutes a residual supply

[54] Ferrell, Stephanie, Matea Gold, Maloy Moorem, Anthony Pesce, and Daniel Schonhaut. 2012. "Outside Spending Shapes 2012 Election." LA Times. Nov 20. <graphics.latimes.com/2012-election-outside-spending>.

shift for political advertisers. This permits identification of PAC demand curves under the assumption that PACs only value audiences in states that are potentially pivotal in the presidential election. Results suggest parties place a premium on the demographics of their base, which is consistent with a get-out-the-vote strategy.

Using these demand estimates, I develop and test a model of monopoly TV station behavior. TV stations are widely thought to price discriminate in sales of airtime to commercial advertisers, but this behavior has not been systematically studied in the literature. My findings confirm these suspicions; observed prices are consistent with a monopoly pricing model, indicating regulation actively prevents stations from price discriminating across candidates. Further, this result suggests lowest unit rate regulation differentially subsidizes candidates in a second fashion. Regulation benefits candidates who prize viewer demographics that are relatively undervalued by the commercial market. For these candidates, regulated rates are likely to fall short of their true value for ad spots.

I extend the monopoly pricing model to estimate a distortionary effect of lowest unit rate regulations. In marrying the campaign price to the lowest price paid by any other advertiser, regulation incentivizes stations to withhold quantity to keep lowest unit rates high. To estimate this effect, I model commercial demand for advertising, and back out unobserved quantities based on station optimization. If advertising has but small negative externalities on viewers, then this decline in commercial airtime constitutes a large loss in efficiency, on the order of 7% of total time. This suggests that extending lowest unit rates to PACs, for example, would have first-order effects on commercial advertising markets. Both this efficiency loss and the distributional consequences of the current regulatory regime warrant consideration in an ultimate welfare calculus for government intervention in political media markets.

References

Anderson, Simon P., & Gabzewikz, Jean J. 2006. The Media and Advertising: A Tale of Two-Sided Markets. *Chap. 18, pages 567–613 of: Handbook of the Economics of Art and Culture*. North-Holland.

Angrist, Joshua D, & Krueger, Alan B. 1995. Split-Sample Instrumental Variables Estimates of the Return to Schooling. *Journal of Business & Economic Statistics*, **13**(2), 225–235.

Ansolabehere, Stephen, Gerber, Alan S, & Snyder, James M. 2001. *Does TV Advertising Explain the Rise of Campaign Spending?*

Bagwell, Kyle. 2007. Handbook of Industrial Organization. **3**(06).

Bel, Germà, & Domènech, Laia. 2009. What Influences Advertising Price in Television Channels?: An Empirical Analysis on the Spanish Market. *Journal of Media Economics*, **22**(3), 164–183.

Blumenthal, Howard, & Goodenough, Oliver R. 2006. *The Business of Television*. Billboard Books.

Chandra, Amitabh, Finkelstein, Amy, Sacarny, Adam, & Syverson, Chad. 2013. *Healthcare Exceptionalism? Productivity and Allocation in the U.S. Healthcare Sector*.

Daivs, Steven, Haltiwanger, John C, & Schuh, Scott. 1996. *Job Creation and Job Destruction*. Cambridge: MIT Press.

Duggan, Mark, & Morton, Fiona M Scott. 2006. The Distortionary Effects of Government Procurement: Evidence from Medicaid Perscription Drug Purchasing. *The Quarterly Journal of Economics*, **CXXI**(1), 1–30.

Fowler, Erika Franklin, & Ridout, Travis N. 2013. Negative, Angry, and Ubiquitous: Political Advertising in 2012. *The Forum*, **10**(4), 51–61.

Gandhi, Amit, Lu, Zhentong, & Shi, Xiaoxia. 2013. *Estimating demand for differentiated products with error in market shares*.

Gentzkow, Matthew, & Shapiro, Jesse. 2010. What Drives Media Slant? Evidence From U.S. Daily Newspapers. *Econometrica*, **78**(1), 35–71.

Goettler, Ronald. 1999. Advertising Rates, Audience Composition, and Competition in the Network Television Industry.

Goldberg, Pinelopi Koujianou. 1996. Dealer Price Discrimination in New Car Purchases: Evidence from the Consumer Expenditure Survey. *Journal of Political Economy*, **104**(3), 622–654.

Goldstein, Kenneth, & Ridout, Travis N. 2004. Measuring the Effects of Televised Political Advertising in the United States. *Annual Review of Political Science*, **7**(1), 205–226.

Gordon, Brett R., & Hartmann, Wesley R. 2013. Advertising Effects in Presidential Elections. *Marketing Science*, **32**(1), 19–35.

Heckman, James J. 1979. Sample Selection Bias as a Specification Error. *Econometrica*, **47**(1), 153–161.

Karanicolas, Michael. 2012. *Regulation of Paid Political Advertising: A Survey*. Tech. rept. March. Centre for Law and Democracy.

Martin, Gregory J, & Yurukoglu, Ali. 2014. *Bias in Cable News: Real Effects and Polarization*.

Nichter, Simeon. 2008. Vote Buying or Turnout Buying? Machine Politics and the Secret Ballot. *American Political Science Review*, **102**(01), 19–31.

Phillips, Robert, & Young, Graham. 2012. Television Advertisement Pricing in the United States. *Pages 181–190 of:* Ozer, Ozalp, & Robert, Phillips (eds), *The Oxford Handbook of Pricing Management*. Oxford: Oxford University Press.

Ridout, Travis N., Franz, Michael, Goldstein, Kenneth M., & Feltus, William J. 2012. Separation by Television Program: Understanding the Targeting of Political Advertising in Presidential Elections. *Political Communication*, **29**(1), 1–23.

Snyder Jr, James M, & Strömberg, David. 2010. Press Coverage and Political Accountability David Stro. *Journal of Political Economy*, **118**(2), 355–408.

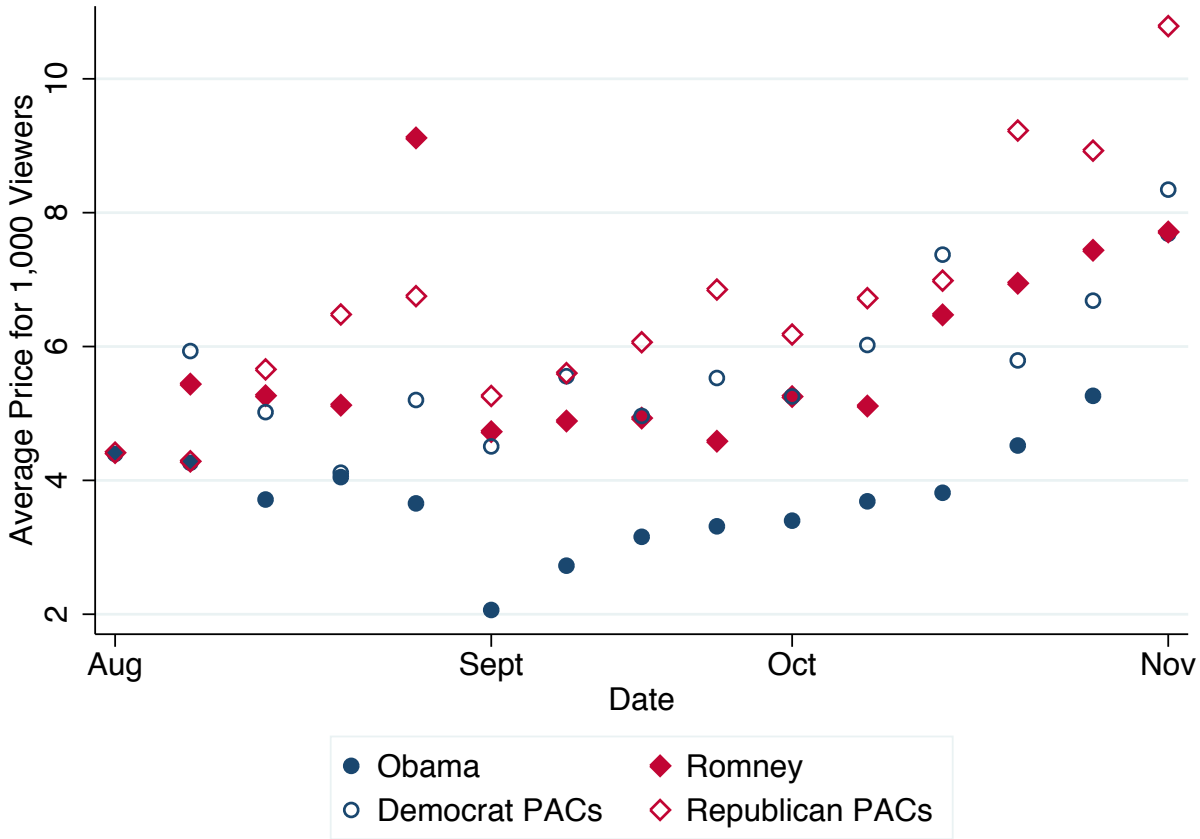
Sock, James H., Wright, Jonathan H., & Yogo, Motohiro. 2002. A Survey of Weak Instruments and Weak Identification in Generalized Method of Moments. *Journal of Business & Economic Statistics*, **20**, 518–529.

Stratmann, Thomas. 2009. How Prices Matter in Politics: the Returns to Campaign Advertising. *Public Choice*, **140**(3/4), 357–3777.

Train, Kenneth E. 2009. *Discrete Choice Methods with Simulation*. Second edn. Cambridge: Cambridge University Press.

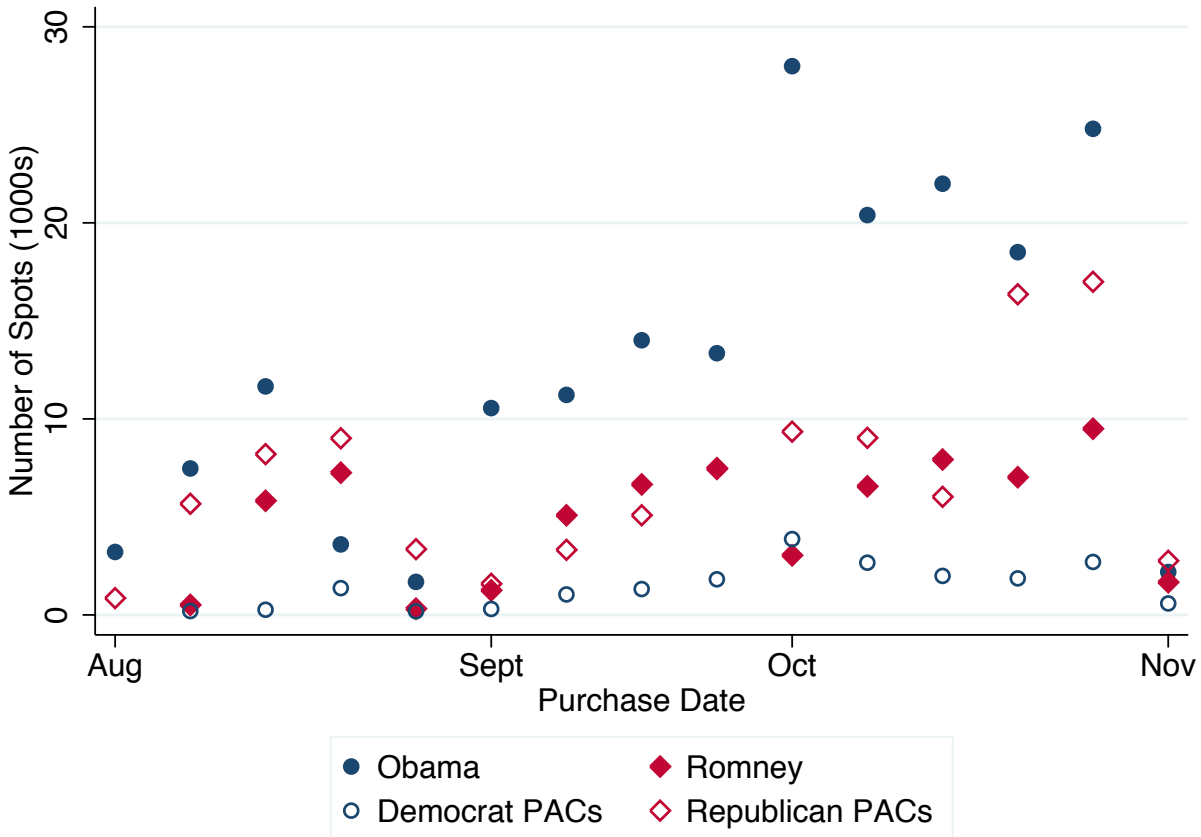
West, Darrell M. 2010. *Air Wars: Television Advertising in Election Campaigns 1952-2008*. Fifth edn. Washignton, DC: CQ Press.

Figure 1: Prices Leading Up to Election Day 2012



Notes: Figure 1 shows that prices (per viewer) increase in the run-up to election day. Since advertising effects are suspected to decay rapidly, ads placed close to November 6, 2012 are likely to be more valuable. High prices near election day is consistent with stations' extracting rent from political ad buyers.

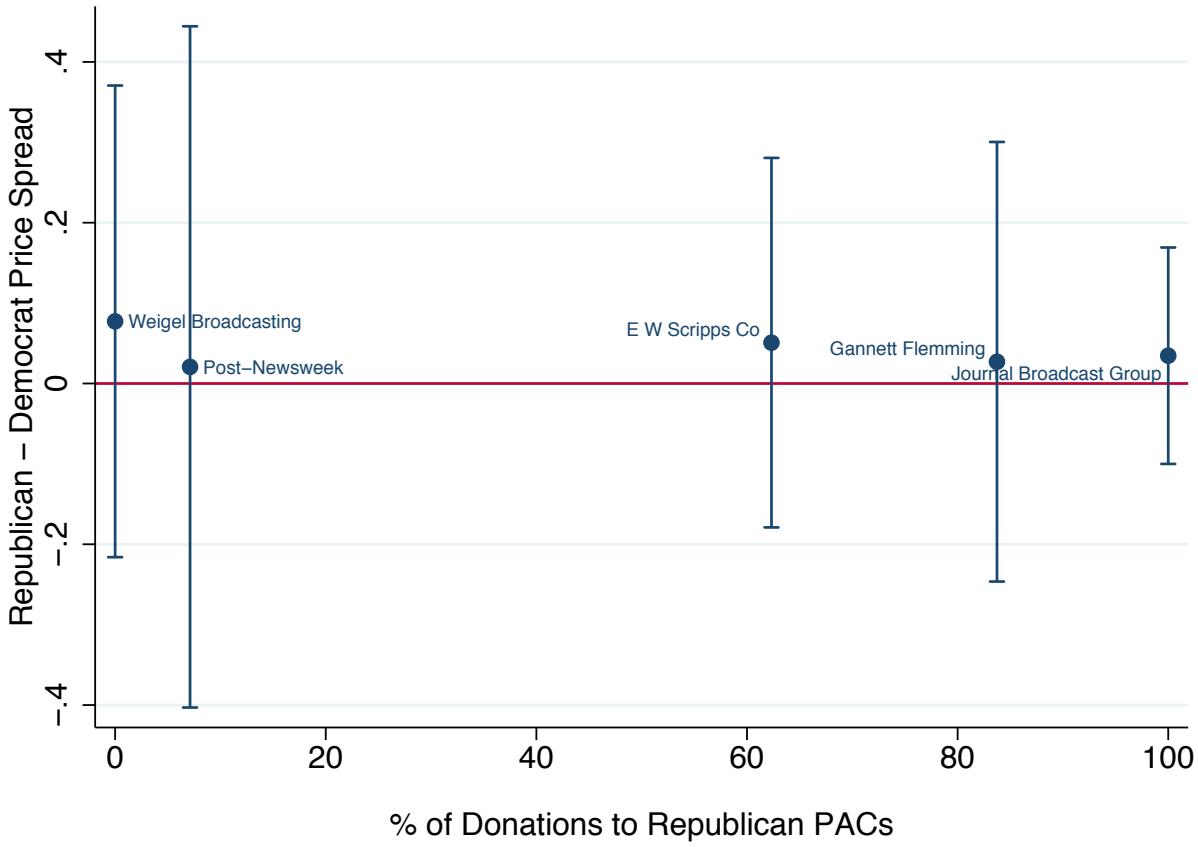
Figure 2: Ad Quantities Leading Up to Election Day 2012



Notes: Figure 2 shows that political ad volumes increase in the run-up to election day, despite price increases.

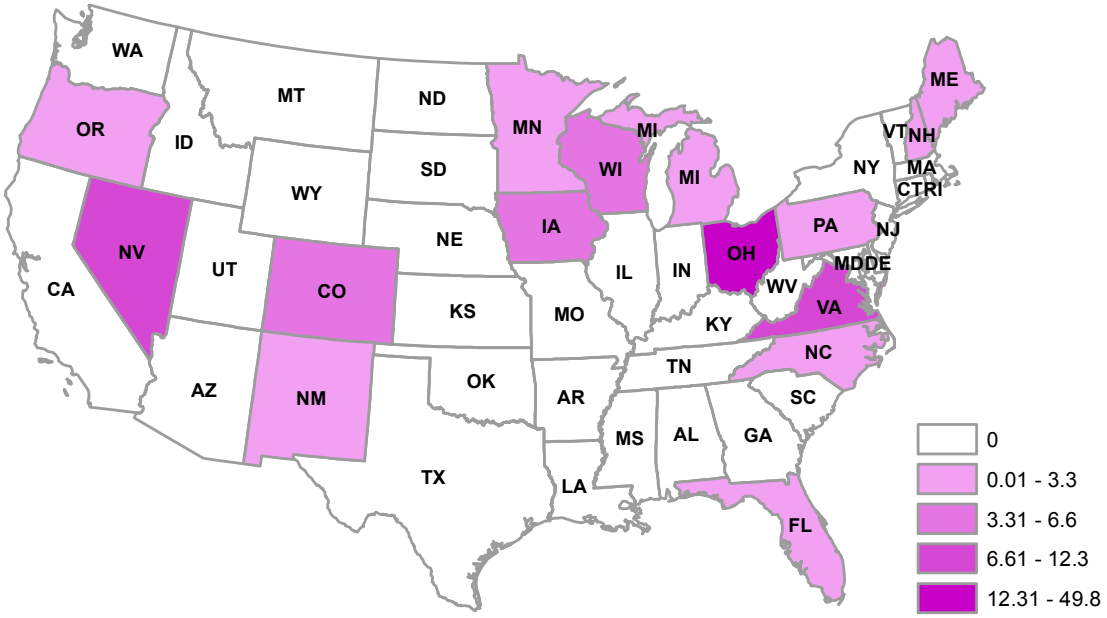
Figure 3: Station Political Donations & PAC Price Disparities

95% Confidence Intervals for Mean Price Disparity



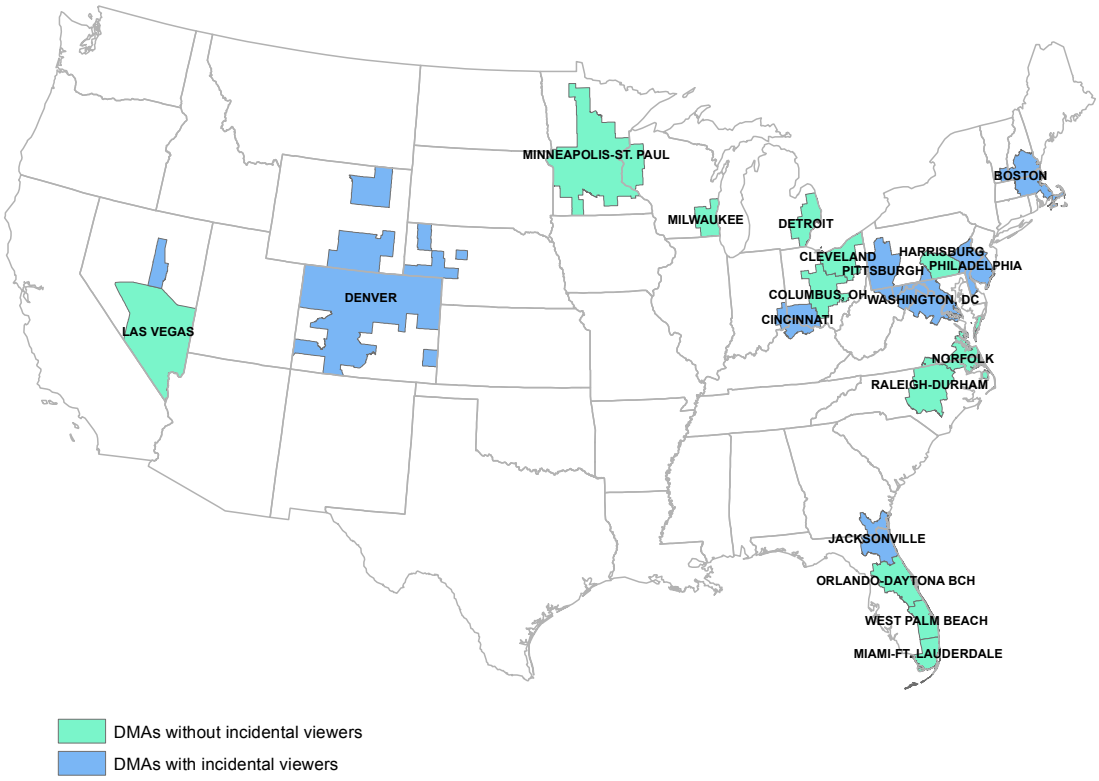
Notes: Figure 3 shows confidence intervals for mean values of the Republican - Democrat price spread by media conglomerate. Price spread is measured as $\frac{P_{Rep}-P_{Dem}}{\frac{1}{2}(P_{Rep}+P_{Dem})}$. While there appears to be a negative correlation of donations to Republicans and offering Republicans lower prices (relative to Democrats), the effect seems small. Station bias in pricing is hard to discern in such a small sample, and warrants further investigation.

Figure 4: Pivotal Probabilities Across States



Notes: Figure 4 displays pivotal (tipping point) probabilities by state in the 2012 Presidential Race. Probabilities are borrowed from Nate Silver's *New York Times* blog.

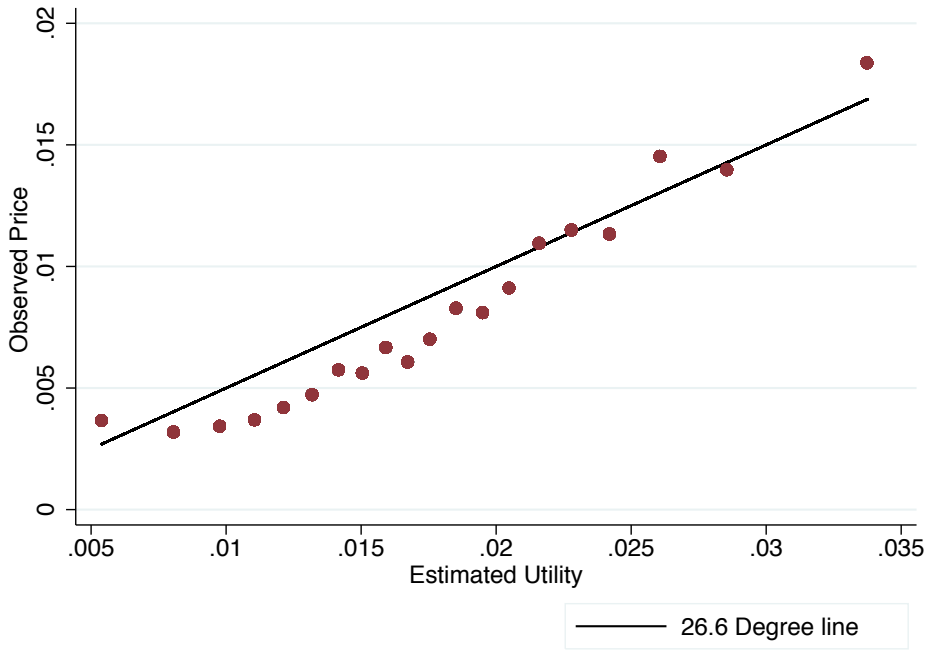
Figure 5: The Geography of Uncontested viewers



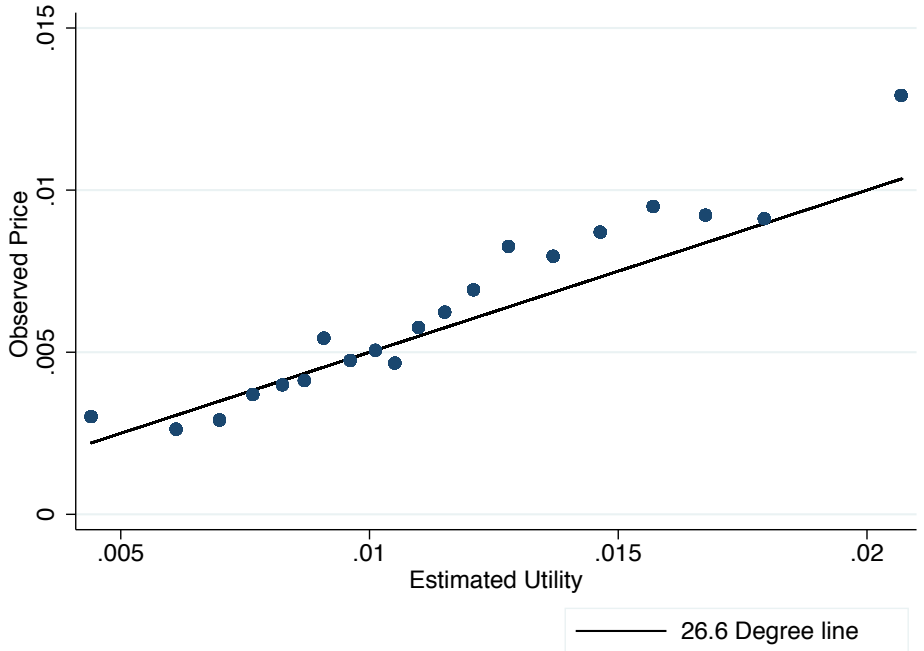
Notes: Figure 5 shows the geography of Designated Market Areas that broadcast to both contested and uncontested (incidental) viewers. Incidental viewers are those viewers who reside in states where the 2012 race as a foregone conclusion.

Figure 6: Prices vs Estimated Utility

(a) Republican PACs



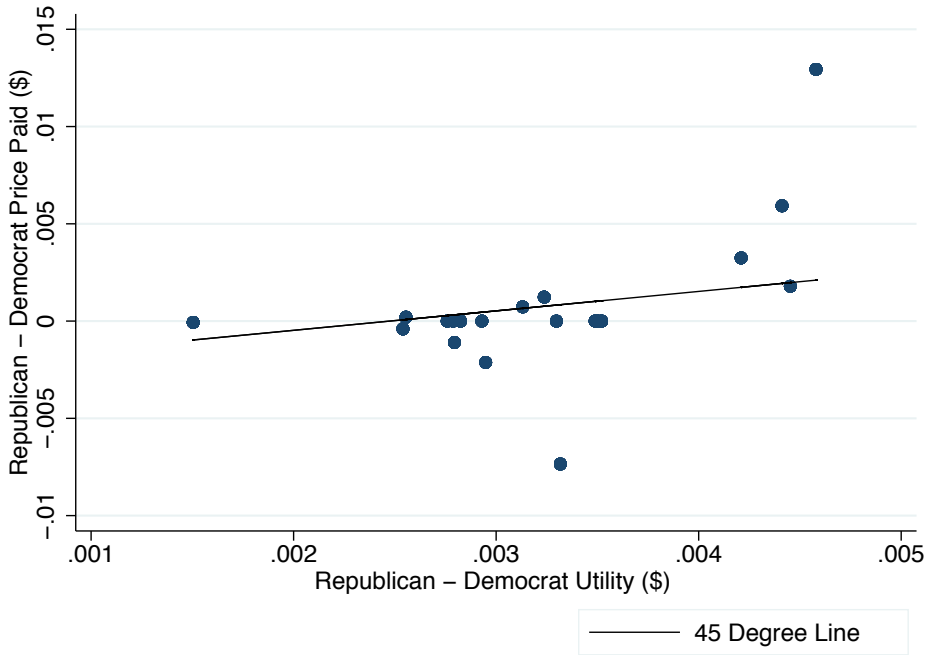
(b) Democrat PACs



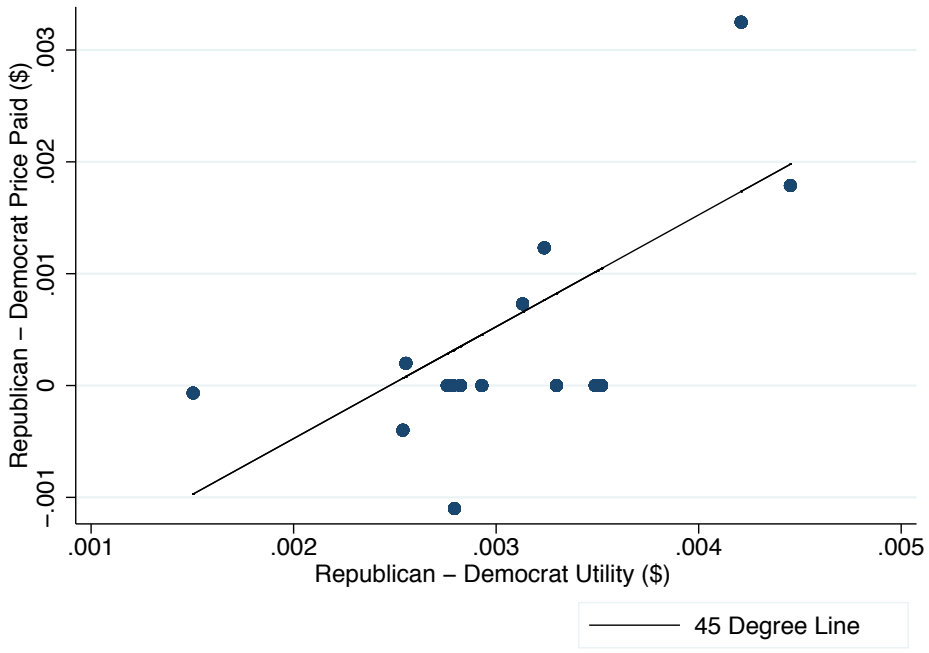
Notes: Figure 6 shows the relationship between observed prices and estimated utilities. The strong, positive correlation suggests stations price, at least in part, on willingness-to-pay. Prices and utilities are measured per viewer in a contested state. Obsverations are grouped into 20 bins according to estimated utilities, each containing five percent of the data.

Figure 7: Price Differences vs Estimated Utility Differences

(a) Full Sample



(b) Interdecile Range



Notes: Figures 7 (a) and (b) show the relationship between observed price differences and estimated utility differences. Subfigure (a) shows the full sample, while (b) restricts to the interdecile range for legibility. The strong, positive correlation suggests differences in willingess-to-pay between Republicans and Democrats help explain the differences in observed prices. The comparison is conducted only for spots where Republican and Democrat PAC purchases are indistinguishable. Prices and utilities are measured per viewer in a contested state. Obsverations are binned into groups of five percentiles.

Table 1
Summary Statistics for Ad Spots Purchased by Political Group

	Democrat PACs	Republican PACs	Obama Campaign	Romney Campaign
Price ($)	1,019.02	1,311.47	835.14	1,135.22
	(1,341.788)	(2,081.08)	(1,741.09)	(1,918.05)
Total Viewership (10,000)	21.28	22.96	22.06	23.23
	(14.05)	(15.59)	(15.94)	(16.14)
Pivotal Viewership (10,000)	19.34	20.74	20.08	20.69
	(13.12)	(14.63)	(15)	(15.05)
Average Pivotality	13.4	19.1	20.13	19.3
	(18.1)	(21.2)	(21.1)	(21.5)
% Women	55.15	54.94	54.32	55.34
	(8.43)	(8.61)	(9.39)	(8.71)
% White	78.4	79	76.58	78.34
	(11.05)	(10.52)	(12.22)	(9.47)
% Black	16.78	16.63	18.99	17.33
	(10.42)	(10.25)	(11.96)	(9.20)
% Over 65	15.92	15.63	14.72	15.8
	(5.09)	(5.39)	(5.48)	(5.45)
Observations	9,326	45,278	53,442	23,520

Notes: Table 1 presents means and standard deviations in parentheses for ads purchased starting August 1, 2012 - November 6, 2012 that were successfully scraped from the FCC website. The average price of a Republican purchase is higher than its Democrat counterpart, but this naive comparison potentially confounds two effects. Stations may charge PACs of different affiliations different prices, but the two groups may also purchase different types of ad spots. As an example, Republican PACs buy higher viewership ad spots, which are costlier.

Table 2
Estimated Ad Exposures in Tipping Point States by Demographic Group and Political Party

	Democrat PACs				Republican PACs			
	Ages 18-64		Ages 65+		Ages 18-64		Ages 65+	
	Men	Women	Men	Women	Men	Women	Men	Women
White	11.73	13.62	8.55	10.26	59.32	68.32	42.89	51.21
Black	11.80	14.18	9.63	11.95	67.05	82.23	54.43	68.51
Other	8.30	9.60	6.17	7.40	39.54	44.84	28.40	33.73

	Obama Campaign				Romney Campaign			
	Ages 18-64		Ages 65+		Ages 18-64		Ages 65+	
	Men	Women	Men	Women	Men	Women	Men	Women
White	68.58	79.43	46.65	56.59	30.97	36.20	22.08	26.71
Black	84.75	102.41	63.61	80.21	36.33	45.02	29.29	37.31
Other	30.91	51.63	68.58	36.98	14.74	23.46	30.97	17.81

Notes: Table 2 calculates expected ad exposures for each of twelve demographic groups based on the purchases in my data. Exposures are calculated based on the programs where ads air and the proclivity of members of each group to watch those programs. The difference in exposure between the Obama and Romney campaign highlights the importance of outside spending in the 2012 election.

Table 3
Price Differences across Political Parties for Indistinguishable Ad Purchases

	PACs (Republicans - Democrats)	Candidates (Romney - Obama)
% Zero Price Difference	41.28	80.34
% Higher Republican Price	30.26	8.28
% Higher Democrat Price	28.45	11.38
Measure of Price Dispersion:		
Absolute Value of Price Difference	196.88	96.21
	(12.63)	(12.98)
Absolute Value of % Price Difference	26	14
	(3.00)	(2.00)
Raw Price Difference	68.41	-33.45
	(14.39)	(14.02)
% Raw Price Difference	14	4
	(3.00)	(3.00)
# Observations	717	290

Notes: Table 3 describes price differences between Republican and Democrat PACs for indisringuishable ad purchases (ad purchases with the same show name, priority level, aired during the same week, at the same station). When there are multiple puchases by different PACs within the same party, I compare the order statistics of the Republican and Democrat prices (for example, the highest Republican and Democrat purchase prices and the lowest purchase prices). The signs and magnitudes of the comparisons are similar if instead I compare average prices.

Table 4
Price Dispersion across vs. within Parties

	Coefficient of Variation			
	(1) Full Sample		(2) Balanced Sample	
Across Republican & Democrat PACs	0.11	622	0.14	224
	(0.15)		(0.18)	
Within Republican PACs	0.03	3400	0.05	224
	(0.10)		(0.10)	
Within Democrat PACs	0.00	664	0.00	224
	(0.01)		(0.00)	

Notes: Table 4 presents the mean coefficient of variation across purchases of ads with indistinguishable characteristics. I estimate the mean both within and across parties. The mean estimate across parties is an order of magnitude larger than the coefficient within party (for either Republicans or Democrats). The coefficient of variation is the standard deviation divided by the mean price for each ad product. Standard deviations are reported in parentheses. The number of observations is reported in the column to the right of coefficients.

Table 5a
Republican PAC Demand for Ad Products Using State Border Design

	OLS	First Stage	IV	IV with FE	Heckman Selection coefficients	Heckman Selection marginal effects
	(1)	(2)	(3)	(4)	(5)	
Price			-47.3***	-32.37***	-82.26***	-36.09
			(10.87)	(5.14)	(13.67)	
Tipping-point probability × Fraction in Demographic Group						
Women	45.16***	0.41***	65.38***	34.99***	80.38***	35.27
	(1.44)	(0.06)	(5.73)	(3.37)	(12.05)	
Aged 65+	31.08***	-0.06	26.05***	26.69***	68.25***	29.95
	(3.99)	(0.10)	(6.80)	(5.45)	(18.06)	
Black	-35.44***	0.74***	14.40	-1.66	-17.92	-7.87
	(7.35)	(0.19)	(17.38)	(9.88)	(33.30)	
White	-25.25***	1.19***	50.08*	13.97	18.40	8.08
	(6.74)	(0.19)	(22.16)	(10.38)	(32.11)	
Instrument						
Viewers in uncontested states		0.179***				
		(0.032)				
Show Fixed Effects				Y	Y	
Observations	18221	5204	18221	18221	18221	
First-stage F-statistic		31.79				
Estimated Elasticity			-0.969	-0.663	-5.87	

Notes: Table 5a presents Republican PAC demand estimates for ad characteristics, and particularly, ad viewer demographics. All variables are measured per viewer in a contested state. Standard errors in (3) & (4) estimated using the N-out-of-N nonparametric bootstrap (1,000 repetitions). IV estimates with fixed effects use only within-program variation across DMAs (first stage results are re-estimated to include FE). Week and priority fixed effects are included in all specifications. Tipping point probability = state pivotality/state population in 1,000,000s. A main tipping point probability variable is also included as a control, as is the proportion of viewers in a state where the senate race is contested. Elasticities are estimated at average ad characteristics.

Table 5b
Democrat PAC Demand for Ad Products Using State Border Design

	OLS	First Stage	IV	IV with FE	Heckman Selection coefficients	Heckman Selection marginal effects
	(1)	(2)	(3)	(4)	(5)	
Price			-41.52***	-36.98***	-189.01**	-25.115
			(8.883)	(6.366)	(71.768)	
Tipping-point probability × Fraction in Demographic Group						
Women	13.212***	0.078	17.50***	9.97***	54.49*	7.240
	(0.990)	(0.061)	(2.959)	(3.043)	(25.506)	
Aged 65+	32.521***	0.291**	42.23***	47.75***	226.31**	30.071
	(2.957)	(0.095)	(5.891)	(6.773)	(86.505)	
Black	-68.332***	-0.224	33.89***	-42.53***	-290.58**	-38.611
	(6.458)	(0.167)	(9.521)	(8.370)	(107.613)	
White	-56.817***	0.098	-62.98***	-64.74***	-178.21*	-23.680
	(6.092)	(0.152)	(9.561)	(7.867)	(69.618)	
Instrument						
Viewers in uncontested states		0.203***				
		(0.033)				
Show Fixed Effects				Y	Y	
Observations	18221	1976	18221	18221	18221	
First-stage F-statistic		37.9				
Estimated Elasticity			-1.208	-1.076	-5.815	

Notes: Table 5b presents Democrat demand estimates for ad characteristics, and particularly, ad viewer demographics. All variables are measured per viewer in a contested state. Standard errors in (3) & (4) estimated using the N-out-of-N nonparametric bootstrap (1,000 repetitions). IV estimates with fixed effects use only within-program variation across DMAs (first stage results are re-estimated to include FE). Week and priority fixed effects are included in all specifications. Tipping point probability = state pivotality/state population in 1,000,000s. A main tipping point probability variable is also included as a control, as is the proportion of viewers in a state where the senate race is contested. Elasticities are estimated at average ad characteristics.

Table 6
Price Paid vs Estimated Utility
Tests of Station Optimization

	Price Paid by Republican PACs			Price Paid by Democrat PACs		
	(1)	(2)	(3)	(4)	(5)	(6)
Estimated Utility	0.623***	0.534***	0.598***	0.676***	0.478***	0.472***
	(0.020)	(0.036)	(0.049)	(0.031)	(0.059)	(0.059)
Cost		0.210***	0.208***		0.374***	0.374***
		(0.050)	(0.050)		(0.082)	(0.082)
Constant	0.000	-0.002***	-0.004***	-0.001***	-0.001***	-0.001
	(0.000)	(0.000)	(0.001)	(0.000)	(0.001)	(0.001)
Expected Unobserved Utility			Y			Y
Observations	5204	2049	2049	1976	943	943

Notes: Table 6 shows the relationship between estimated PAC willingness-to-pay for ad spots and purchase prices. Under the null hypothesis that stations are single-product monopolists, the coefficient on estimated utility is 0.5. All variable are measured per contested viewer. Cost is the lowest unit rate paid by campaigns for an indistinguishable product during the 60-day window before the general election; there are fewer observations in regressions including cost since LUR data is available only if a campaign purchases. Heteroskedasticity-robust standard errors in parentheses. Coefficients are statistically significant at the * .05, ** .01 and, *** .001 level.

Table 7
Price Differences vs Estimated Utility Differences
Tests of Station Optimization

	Republican - Democrat Price Paid			
Sample:	Full	Indistinguishable Add-Ons	Non-zero Price Difference	Indistinguishable Add-Ons & Non-zero Price Difference
	(1)	(2)	(3)	(4)
Utility Difference ($)	0.277**	0.610*	0.392**	0.793*
	(0.090)	(0.267)	(0.126)	(0.331)
Observations	1501	196	996	103

Notes: Table 7 describes the relationship between observed price differences and model-generated utility differences. If price differences reflect differences in WTP for the same ad spot, then coefficient estimates should be positive and statistically significant. Under the monopoly pricing model described in Section 4, the coefficient on utility differences should be 1. Robust standard errors are reported in parentheses. All variables are measured per viewer in a contested state.

Table 8

Bayesian Parameter Estimates for Quantity Withholding, Commercial and Campaign Demand

	Initial Guess	Posterior Mean	Credible Interval	
% Inventory Withheld		7.52	7.39	7.61
Commercial Parameters				
Constant	-2.22	-2.36	-2.38	-2.27
Fraction Black	0.36	1.01	1.00	1.08
Fraction White	0.36	0.05	0.04	0.07
Fraction Old	0.45	0.63	0.56	0.66
Fraction Female	-0.34	0.08	0.06	0.09
Ratio of Non-contested to Contested Viewers	-0.01	0.08	0.07	0.08
Week	0.02	-0.04	-0.05	-0.02
Price Per Viewer	-24.67	-183.26	-183.52	-183.18
Campaign Parameters				
Constant	-5.12	-4.43	-4.57	-3.80
Pivotality	0.08	0.23	0.21	0.25
Fraction Black	-0.84	-0.20	-0.21	-0.17
Fraction White	-0.12	-1.11	-1.12	-1.08
Fraction Old	2.59	3.32	3.24	3.41
Fraction Female	-0.15	0.00	-0.03	0.01
Week	0.02	-0.05	-0.06	-0.05
Price Per Viewer	-31.53	-11.23	-11.23	-11.21
Error Covariance				
Commercial Variance (σ^2_ω)	4.04	9.00	7.58	10.13
Campaign Variance (σ^2_ξ)	10.60	6.00	4.61	6.48
Covariance ($\sigma_{\xi\omega}$)	0.52	0.34	0.01	0.56

Notes: Table 8 shows estimates for the parameters of the demand model outlined in section 5. The goal of the model is to estimate quantity withholding, the amount of inventory stations do not sell in order to keep LURs high. Estimates are based on 250,000 draws using a random-walk Metropolis-Hastings sampling algorithm, and a burn-in period of 50,000 draws. 3.3% of observed prices are not rationalizable at the posterior mean of the parameters. The acceptance rate is regulated to 0.37. The credible interval is asymptotically equivalent to the 95% CI.

Table A1
Political Action Committee Classification

Republican PACS	Number of Ads	Democrat PACS	Number of Ads
60 Plus Association	625	AFL-CIO	27
Special Operations OPSEC Education Fund	198	AFSCME	1,167
American Action Network	5,832	Alliance for a Better MN	202
American Chemimstry Council	74	Committee for Justice & Fairness	249
American Energy Alliance	21	DNC	206
American Future Fund	1,164	Florida Democratic Party	65
American Unity PAC	9	Independence USA PAC	167
Americans for Job Security	1,769	League of Conservation Voters	486
Americans for Prosperity	3,200	MN United for All Families	721
Americans for Tax Reform	37	MoveOn.org	38
Campaign for American Values	54	Moving Ohio Forward	169
Center for Individual Freedom	101	National Education Association	427
Checks and Balances for Economic Growth	26	Patriot Majority PAC	574
Club for Growth Action Committee	279	Planned Parenthood	415
American Crossroads/Crossroads GPS	16,296	Priorities USA	3,306
Emergency Committee for Israel	16	SEIU	904
Ending Spending PAC	118	Women Vote!	203
Freedom Fund	74	Total	9326
Freedom PAC	68		
Government Integrity Fund	170		
Judicial Crisis Network	54		
Live Free or Die PAC	290		
National Association of Manufacturers	197		
National Federation of Independent Business	262		
National Republican Trust	65		
National Rifle Association	142		
Now or Never PAC	119		
Republican Jewish Coalition	1,017		
Republican Party of Florida	150		
Restore Our Future	5,529		
RNC	5,806		
Securing Our Safety	46		
SuperPAC for America	31		
US Chamber of Commerce	1,020		
Women Speak Out PAC	22		
Young Guns Action Fund	397		
Total	45278		

Notes: PACs are classified as Republican or Democrat based on the classification (conservative or liberal) at OpenSecrets.org, a website maintained by the Center for Responsive Politics.

Table A2
Selection of Ads from the Online FCC Database

Criteria	Number Dropped	Percent of Raw Sample
Missing show name	1,048	0.46
Aired before 08/01/2012	7,020	3.09
Longer or shorter than 30 seconds	9,406	4.14
Non-presidential PAC	37,031	16.29
PAC purchased < 20 spots	398	0.18
No clear party affiliation	15,201	6.69
Station with single-party advertising	14,716	6.47
Station without presidential advertising	7,835	3.45
Total eliminated	92,655	

Notes: Table A2 describes how I refine the raw data for demand estimation in section 3. Shows that have no identifiable name cannot be matched to viewership data, so they are excluded from the demand analysis. Shows airing before August 1, 2012 are excluded because stations are not required to post invoices predating August, 2 2012; those that choose to may be a selected sample. I do not consider sales of airtime that are longer or shorter than the standard 30 second spot (e.g. some of these are zeros, indicating time was not sold after all). The analysis also excludes purchases by very small PACs or PACs with no clear party affiliation. Stations with single-party advertising or without campaign advertising are excluded as these suggest purchasing for other races. 134,671 observations remain in the sample.

Table A3
Advertising Counts by Station

Station	Designated Market Area	Observations	Percent of Sample
KARE	Minneapolis	2,020	1.58
KCNC-TV	Denver	5,428	4.24
KDVR	Denver	9,581	7.48
KMGH-TV	Denver	6,218	4.86
KMSP-TV	Minneapolis	592	0.46
KTNV-TV	Las Vegas	5,483	4.28
KYW-TV	Philadelphia	698	0.55
WBZ-TV	Boston	1,579	1.23
WCCO-TV	Minneapolis	913	0.71
WCPO-TV	Cincinnati	5,176	4.04
WCVB-TV	Boston	918	0.72
WDJT-TV	Milwaukee	3,759	2.94
WEWS-TV	Cleveland	15,635	12.21
WFLX	West Palm Bch	1,979	1.55
WFOR-TV	Miami	2,605	2.03
WFXT	Boston	1,295	1.01
WHTM-TV	Harrisburg	948	0.74
WISN-TV	Milwaukee	4,060	3.17
WJXX	Jacksonville	4,320	3.37
WKMG-TV	Orlando	7,653	5.98
WKYC	Cleveland	7,077	5.53
WLWT	Cincinnati	3,499	2.73
WPLG	Miami	5,150	4.02
WPMT	Harrisburg	477	0.37
WPVI-TV	Philadelphia	44	0.03
WRAL-TV	Raleigh	1,647	1.29
WSYX	Columbus, OH	4,364	3.41
WTAE-TV	Pittsburgh	1,155	0.90
WTLV	Jacksonville	1,701	1.33
WTTE	Columbus, OH	2,411	1.88
WTVJ	Miami	1,866	1.46
WUSA	Washington, DC	5,038	3.93
WVBT	Norfolk	7,749	6.05
WVEC	Norfolk	948	0.74
WWJ-TV	Detroit	340	0.27
WXIX-TV	Cincinnati	3,725	2.91
Total		128,051	

Notes: The FCC 2012 archive includes only affiliates of the four major networks in top-50 DMAs. Data is scraped from using OCR software, so that some stations are omitted because the software could not parse their upload formats. Despite these limitations, to my knowledge, this is the most comprehensive set of advertising price data from the presidential election.

Technical Appendix: Likelihood Function for Bayesian Estimation

In this appendix, I develop the likelihood function that I use to estimate quantity withholding in Section 5. The difficulty in estimation is that I only observe data from the campaign side of the market, and quantity withholding depends on the total quantity of airtime sold (to both commercial and campaign advertisers). The strategy is to infer commercial sales from the firm's decisions, assuming the firm set prices optimally. I split the likelihood function into two pieces that depend on whether the campaign price and quantity is observed.

Campaign Price and Quantity is observed

Given invoice data on price and quantity, I can back out the mean utility of show i for campaigns ($\tilde{\delta}_i$), and the implied demand shock: $\xi_i = \tilde{\delta}_i - \tilde{x}_i\tilde{\beta} + \tilde{\alpha}p_i$. I know the quantity sold on the commercial market came either from an interior solution to the firm's optimization problem or from a boundary solution. If it came from an interior solution, then there is an efficiency loss from quantity withholding. Let $\mathbb{P}\{p_i, \tilde{s}_i\}$ be the probability of observing price p_i and campaign demand $\tilde{s}_i$ (with corresponding campaign mean utility $\tilde{\delta}_i$):

$$\mathbb{P}\{p_i, \tilde{s}_i\} \quad = \quad \mathbb{P}\{p_i, \tilde{s}_i, \text{ boundary solution}\} + \mathbb{P}\{p_i, \tilde{s}_i, \text{ interior solution}\}.$$

Observed outcomes p_i and $\tilde{s}_i$ are the product of a boundary solution if $\delta_i = \delta_i^B$. They are the product of an interior solution if $\delta_i = \delta_i^I$.

Capacity constrained optimum (observed)

If the optimum is at the boundary, then the commercial quantity is immediately known: it is the residual amount of airtime. Mean commercial utility is then perfectly observed: $\delta_i^B = \ln\left[\frac{T_i/M_i - \tilde{s}_i}{1 - T_i/M_i + \tilde{s}_i}\right]$. The supply shock is simply the residual difference between this mean utility and the observed components of utility: $\omega_i = \delta_i^B - x_i\beta + \alpha p_i$. Once these shocks are calculated, it's imperative to check that they are consistent with a boundary solution – i.e. that the observed price is indeed optimal given the implied shocks. If not, then I assign zero likelihood to the capacity constrained optimum.

The final step in the likelihood is to calculate the modulus of the Jacobin corresponding

to a change-in-variables from $(\tilde{\delta}, p)$ to (ξ, ω).

$$\mathbb{P}\{\tilde{\delta}, p, \delta = \delta^B\} \;=\; \mathbb{P}\{\xi = \tilde{\delta} - \tilde{x}\tilde{\beta} + \tilde{\alpha}p, \omega = \delta^B - x\beta + \alpha p\} 1\{p = p^*(\omega, \xi)\} \begin{vmatrix} \partial\xi/\partial\tilde{\delta} & \partial\xi/\partial p \\ \partial\omega/\partial\tilde{\delta} & \partial\omega/\partial p \end{vmatrix}$$

The elements of the Jacobian for the change-of-variables between observed mean utilities $\tilde{\delta}$ and prices p are:

$$\begin{aligned} \frac{\partial\xi}{\partial\tilde{\delta}} &= 1 \qquad & \frac{\partial\xi}{\partial p} &= \tilde{\alpha} \\ \frac{\partial\omega}{\partial\tilde{\delta}} &= -1 \qquad & \frac{\partial\omega}{\partial p} &= \alpha \\ \implies det &= & \alpha + \tilde{\alpha}. \end{aligned}$$

Re-writing the probability:

$$\mathbb{P}\{\tilde{\delta}_i, p_i, \delta_i = \delta^B\} = \mathbb{P}\{\xi = \tilde{\delta} - \tilde{x}\tilde{\beta} + \tilde{\alpha}p, \omega = \delta^B - x\beta + \alpha p\} 1\{p = p^*(\omega, \xi)\} |\alpha + \tilde{\alpha}|.$$

Interior optimum (observed)

If the observed campaign share and price arose from an interior optimum, then I can use the station's first order condition to back out the unobserved commercial share:

$$s^I \;=\; \frac{1}{2} \pm \sqrt{\frac{1}{4} - \frac{\tilde{s} - \tilde{\alpha}p\tilde{s}(1 - \tilde{s})}{\alpha p}}.$$

There are up to two roots (commercial shares) consistent with the observed data (given parameter values $\alpha, \tilde{\alpha}$). If $s^{(i)}$ $(i \in \{1, 2\})$ is a root of the quadratic equation, then $s^{(i)} \in R$ constitutes a viable equilibrium if $s^{(i)} \in [0, \frac{T}{M} - \tilde{s}]$. Let $\delta^{(i)}$ be the mean utility corresponding to $s^{(i)}$. Then we can use the implied supply shocks $(\omega^{(i)})$ to create a likelihood:

$$\begin{aligned} \mathbb{P}\{\tilde{\delta}, p, \delta = \delta^I\} \;=\; & \mathbb{P}\{\xi = \tilde{\delta} - \tilde{x}\tilde{\beta} + \tilde{\alpha}p, \omega = \delta^{(1)} - x\beta + \alpha p\} \cdot 1\{s^{(1)} \in [0, \frac{T}{M} - \tilde{s}]\} \cdot \begin{vmatrix} \partial\xi/\partial\tilde{\delta} & \partial\xi/\partial p \\ \partial\omega/\partial\tilde{\delta} & \partial\omega/\partial p \end{vmatrix} \\ + \; & \mathbb{P}\{\xi = \tilde{\delta} - \tilde{x}\tilde{\beta} + \tilde{\alpha}p, \omega = \delta^{(2)} - x\beta + \alpha p\} \cdot 1\{s^{(2)} \in [0, \frac{T}{M} - \tilde{s}]\} \cdot \begin{vmatrix} \partial\xi/\partial\tilde{\delta} & \partial\xi/\partial p \\ \partial\omega/\partial\tilde{\delta} & \partial\omega/\partial p \end{vmatrix}. \end{aligned}$$

The Jacobian is not the same as in the boundary solution case, since the relationship between ω_i and p_i is now given by the FOC.

$$
\begin{aligned}
\frac{\partial \omega}{\partial \tilde{\delta}} &= \frac{\partial \delta}{\partial s}\frac{\partial s}{\partial \tilde{s}}\frac{\partial \tilde{s}}{\partial \tilde{\delta}} \\
&= \frac{1}{s(1-s)}\frac{\partial s}{\partial \tilde{s}}\frac{\partial \tilde{s}}{\partial \tilde{\delta}} \\
&= \pm\frac{1}{2s(1-s)}\left(\frac{1}{4}-\frac{\tilde{s}-\tilde{\alpha}p\tilde{s}(1-\tilde{s})}{\alpha p}\right)^{-\frac{1}{2}}\frac{(\tilde{\alpha}p(1-2\tilde{s})-1)}{\alpha p}\frac{\partial \tilde{s}}{\partial \tilde{\delta}} \\
&= \pm\frac{\tilde{s}(1-\tilde{s})}{2s(1-s)}\left(\frac{1}{4}-\frac{\tilde{s}-\tilde{\alpha}p\tilde{s}(1-\tilde{s})}{\alpha p}\right)^{-\frac{1}{2}}\frac{(\tilde{\alpha}p(1-2\tilde{s})-1)}{\alpha p} \\
&= \begin{cases} \frac{\tilde{s}(1-\tilde{s})(\tilde{\alpha}p(1-2\tilde{s})-1)}{\alpha ps(1-s)(2s-1)} & \text{if pos root} \\ \frac{\tilde{s}(1-\tilde{s})(\tilde{\alpha}p(1-2\tilde{s})-1)}{\alpha ps(1-s)(1-2s)} & \text{if neg root} \end{cases} \\
\frac{\partial \omega}{\partial p} &= \alpha + \frac{\partial \delta}{\partial s}\frac{\partial s}{\partial p} \\
&= \alpha + \frac{1-s}{s}(\frac{1}{1-s}+\frac{s}{(1-s)^2})\frac{\partial s}{\partial p} \\
&= \alpha + \frac{1}{s(1-s)}\left(\pm\frac{\frac{\tilde{s}}{\alpha p^2}}{2\sqrt{\frac{1}{4}-\frac{\tilde{s}-\tilde{\alpha}p\tilde{s}(1-\tilde{s})}{\alpha p}}}\right) \\
&= \begin{cases} \alpha + \frac{\tilde{s}}{\alpha p^2 s(1-s)(2s-1)} & \text{if pos root} \\ \alpha + \frac{\tilde{s}}{\alpha p^2 s(1-s)(1-2s)} & \text{if neg root} \end{cases}
\end{aligned}
$$

Likelihood if campaigns make no observed purchases

The integral of interest is the probability the campaign share is less than $1/M$ given product characteristics $x, \tilde{x}$. For tractability, split this piece of the likelihood into two components, depending on whether draws of (ξ, ω) imply an interior or boundary solution.

Interior optimum (unobserved)

Integrating over (ξ, ω) space, the likelihood of an unobserved interior optimum corresponds to:

$$
\mathbb{P}\left\{\tilde{s} \leq \frac{1}{M}, \delta = \delta^I\right\} = \int_{\xi,\omega:\tilde{s}(\xi,\omega)\leq\frac{1}{M},\delta=\delta^I} f(\xi,\omega)d\xi\omega
$$

Unfortunately, the domain is not closed-form, and sampling from the full distribution of (ω, ξ) might require a large number of simulations to produce draws within the bounds. Instead, consider integration over mean utility $(\tilde{\delta}, \delta)$ space. Let p^I be the price given by the FOC (interior solution) and p^B be the price given by the capacity constraint (boundary solution). The requirement $\tilde{s} \leq \frac{1}{M}$ at an unconstrained optimum amounts to $\tilde{\delta}(p^I) \leq \ln \frac{1}{M-1}$. Using the change-of-variables:

$$\mathbb{P}\left\{\tilde{s} \leq \frac{1}{M}, \delta = \delta^I\right\} = \int_{-\infty}^{\ln \frac{1}{M-1}} \int_{-\infty}^{\infty} f\left(\xi(\tilde{\delta}, \delta), \omega(\tilde{\delta}, \delta)\right) \begin{vmatrix} \partial\xi/\partial\tilde{\delta} & \partial\xi/\partial p \\ \partial\omega/\partial\tilde{\delta} & \partial\omega/\partial p \end{vmatrix} d\delta d\tilde{\delta}$$

Given a draw $(\tilde{\delta}_s, \delta_s)$, an interior optimal price is defined by the logit share equation and the FOC:

$$\begin{aligned} \tilde{s}_s(\tilde{\delta}_s) &= \frac{\exp(\tilde{\delta}_s)}{1 + \exp(\tilde{\delta}_s)} \\ p_s^I &= \frac{\tilde{s}_s}{\tilde{\alpha}\tilde{s}_s(1 - \tilde{s}_s) + \alpha s_s(1 - s_s)} \end{aligned}$$

This implies values of the unobservable:

$$\begin{aligned} \xi_s &= \tilde{\delta}_s - \tilde{x}_i\tilde{\beta} + \tilde{\alpha}p_s^I \\ \omega_s &= \delta_s - x_i\beta + \alpha p_s^I \end{aligned}$$

I draw $\tilde{\delta}_s$ from a normal distribution with mean $\ln \frac{1}{M-1}$, variance $\hat{\sigma}_\xi^2$, truncated above at $\ln \frac{1}{M-1}$. I draw $\delta_s \sim N\big(x_i\beta, 10\hat{\sigma}_\omega^2\big)$. $\hat{\sigma}_\omega^2$ and $\hat{\sigma}_\xi^2$ are the estimated variances of the shocks based on an initial IV regression. If I draw $j = 1, ..., S$ simulations, then I can estimate this probability as:

$$\hat{\mathbb{P}}\left\{\tilde{s} \leq \frac{1}{M}, \delta = \delta^I\right\} = \frac{1}{S}\sum_{s=1}^{S} \frac{F\big(\xi_s, \omega_s\big) \begin{vmatrix} \partial\xi/\partial\tilde{\delta} & \partial\xi/\partial p \\ \partial\omega/\partial\tilde{\delta} & \partial\omega/\partial p \end{vmatrix}}{\frac{\phi\left(\frac{\delta_s - x\beta}{\sqrt{10}\hat{\sigma}_\omega}\right)}{\Phi\left(\frac{\delta_s + \ln(T-1)}{\sqrt{10}\hat{\sigma}_\omega}\right)} \frac{\phi\left(\frac{\tilde{\delta}_s + \ln(M-1)}{\sqrt{10}\hat{\sigma}_\xi}\right)}{10\hat{\sigma}_\xi\hat{\sigma}_\omega}}$$

Constrained optimum (unobserved)

It is also difficult to sample (ξ, ω) where there is mass in constrained optima, and the observed share is below $\frac{1}{M}$. Instead, I sample from $\tilde{\delta} \geq \ln \frac{1}{M-1}$ – I try to find mean utilities where at the interior optimal price (p^I), the campaign share exceeds the observed bound and the capacity constraint is also violated. In those cases, it is possible that the boundary condition will push the optimal campaign share below the observation threshold. For a candidate draw of $(\tilde{\delta}_s, \delta_s)$, I find the implied FOC price as:

$$p_s^I = \frac{\tilde{s}_s(\tilde{\delta}_s)}{\tilde{\alpha}\tilde{s}_s(\tilde{\delta}_s)(1-\tilde{s}_s(\tilde{\delta}_s)) + \alpha s_s(\delta_s)(1-s_s(\delta_s))}$$

Then I back-out the implied taste shocks using this interior price:

$$\begin{aligned} \xi_s &= \tilde{\delta}_s - \tilde{x}\tilde{\beta} + \tilde{\alpha}p^I \\ \omega_s &= \delta_s - x\beta + \alpha p^I \end{aligned}$$

By construction, at p^I, the capacity constraint is violated. So I use these shocks to find the boundary price (which must be the unobserved, equilibrium price). The boundary price, p^B, solves the following nonlinear equation:

$$\frac{T}{M} = \frac{\exp\left(\tilde{x}\tilde{\beta} - \tilde{\alpha}p^B + \xi_s\right)}{1+\exp\left(\tilde{x}\tilde{\beta} - \tilde{\alpha}p^B + \xi_s\right)} + \frac{\exp\left(x\beta - \alpha p^B + \omega_s\right)}{1+\exp\left(x\beta - \alpha p^B + \omega_s\right)}$$

I can approximate the probability that the campaign quantity fell below 1 unit and equilibrium price was from a boundary solution as:

$$\begin{aligned} \mathbb{P}\left\{\tilde{s} \leq \frac{1}{M}, \delta = \delta^B\right\} &= \int_{\ln\frac{1}{M_i-1}}^{\infty}\int_{-\infty}^{\infty} f(\xi,\omega)\begin{vmatrix} \partial\xi/\partial\tilde{\delta} & \partial\xi/\partial p \\ \partial\omega/\partial\tilde{\delta} & \partial\omega/\partial p \end{vmatrix} d\delta d\tilde{\delta} \\ \hat{\mathbb{P}}\left\{\tilde{s} \leq \frac{1}{M}, \delta = \delta^B\right\} &= \frac{1}{S}\sum_{s=1}^{S} \frac{F\left(\xi_s,\omega_s\right)\begin{vmatrix} \partial\xi/\partial\tilde{\delta} & \partial\xi/\partial p \\ \partial\omega/\partial\tilde{\delta} & \partial\omega/\partial p \end{vmatrix}}{\frac{\phi\left(\frac{\delta_s - x\beta}{\sqrt{10}\hat{\sigma}_\omega}\right)}{1-\Phi\left(\frac{\delta_s+\ln(M-1)}{\sqrt{10}\hat{\sigma}_\omega}\right)}\frac{\phi\left(\frac{\tilde{\delta}_s+\ln(T-1)}{\sqrt{10}\hat{\sigma}_\xi}\right)}{10\hat{\sigma}_\xi\hat{\sigma}_\omega}} \cdot 1\left\{\tilde{s}(p_s^B) \leq \frac{1}{M}\right\} \end{aligned}$$