Essays in Price Discrimination and Regulation

by

Sarah N. S. Moshary

SUBMITTED TO THE DEPARTMENT OF ECONOMICS IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY
AT THE
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

SEPTEMBER 2015

© 2015 Sarah N. S. Moshary. All rights reserved.

The author hereby grants MIT permission to reproduce and to distribute publicly paper and electronic copies of this thesis document in whole or in part in any medium now known or hereafter created.

Signature of Author: [Signature redacted]

Department of Economics
August 15, 2015

Certified By: [Signature redacted]
Glenn Ellison
Gregory K. Palm Professor of Economics
Thesis Supervisor

Certified By: [Signature redacted]
Nancy Rose
Charles P. Kindleberger Professor of Economics
Thesis Supervisor

Certified By: [Signature redacted]
Paulo Somaini
Assistant Professor of Economics
Thesis Supervisor

Accepted By: [Signature redacted]
Ricardo Caballero
Ford International Professor of Economics
Chairman, Departmental Committee on Graduate Studies
ABSTRACT

Chapter 1 studies price discrimination in advertising sales to Political Action Committees (PACs) in the 2012 Presidential Election. These groups have grown rapidly – expenditures neared $500 million in the 2012 presidential election – and their effect on elections depends on regulation and its interaction with imperfect competition. While the government tightly proscribes station behavior vis-a-vis official campaigns, it does not protect Political Actions Committees (PACs). Television stations potentially wield considerable power to shape access to the electorate. Using novel data on prices paid for individual ad spots from the 2012 presidential election, I find PACs pay a 40% markup above campaign rates, and that there are differences in prices paid by Republican and Democratic groups for indistinguishable purchases. I then 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.

Chapter 2 investigates spillover effects of regulation protecting campaign advertising purchases, a most favored nation clause. This regulation guarantees campaigns the lowest rate received by any advertiser, incentivizing stations to sell less airtime to commercial advertisers to buoy campaign prices. Using spot-level data on presidential campaign advertising purchases from 2012, I find that campaign ad prices drop following the institution of rate regulation (sixty days preceding election day). I then develop a model of station price discrimination, and estimate the effect of regulation on campaign and commercial prices relative to a counterfactual without regulation.

Chapter 3, co-authored with Gaston Illanes, studies the effects of potential entry on market outcomes in the context of Washington state’s 2012 privatization of liquor sales. Theory indicates that entry, and even the threat of entry, plays a key role in disciplining market outcomes. We exploit the post-reform licensure requirement that stores have 10,000 square feet of retail space to estimate the impact of an additional store on price competition. We compare prices and product variety in markets with stores just above versus just below the square footage cutoff.
Thesis Supervisor: Glenn Ellison
Title: Gregory K. Palm Professor of Economics

Thesis Supervisor: Nancy Rose
Title: Charles P. Kindleberger Professor of Economics

Thesis Supervisor: Paulo Somaini
Title: Assistant Professor of Economics
Acknowledgements

I would like to express tremendous gratitude to my advisors, Glenn Ellison, Nancy Rose and Paulo Somaini. Glenn’s insight, attention to detail, and guidance through the job market process were invaluable. Nancy’s perspective prevented me from losing the forest for the trees, and her positive attitude buoyed me through the tough patches of dissertation-writing. I also appreciate Paulo’s help with the nitty-gritty, as I learned to write and estimate a structural model for the first time. Their generosity with time and patience has been truly remarkable.

I would like to thank all of the faculty at MIT who provided feedback and help at seminars, lunches, and meetings over the past years. In particular, I thank Nikhil Agarwal, Josh Angrist, Victor Chernozhukov, Sara Ellison, Jerry Hausman, Chris Knittel, Anna Mikusheva, and Dick Schmalensee. My thanks also to Heidi Williams and Joe Doyle, who helped me get smart about the research process and continue to offer advice on steering a course through the economics profession. I am very grateful to Oliver Hart and Hamid Mehran, whose teaching and advising brought me to MIT in the first place.

Thanks to all of my classmates who enriched my graduate school experience and made the long hours pass so quickly. In particular, I thank Gaston Illanes, Ashish Shenoy, Sally Hudson, Bradley Shapiro, Bradley Larsen, Xiao Yu Wang, Nils Wernerfelt, Isaiah Andrews, and Manisha Padi. I feel truly lucky to have travelled the grad school road with such smart and supportive companions. Also, a special thanks to all of the wonderful folks at Maseeh Hall, who introduced me to the world of MIT beyond the Economics Department.

Words cannot express the debt of gratitude I owe to my parents, Amy Schwartz and Fred Moshary, and my sisters, Arianna and Leila. They are my biggest fans, tireless cheerleaders, and late-night help line. Thank you.
## Contents

1 Price Discrimination across Political Action Committees: Evidence from the 2012 Presidential Election

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1 Data</td>
<td>9</td>
</tr>
<tr>
<td>1.1.1 Data Sources</td>
<td>9</td>
</tr>
<tr>
<td>1.1.2 Who Sees Political Ads?</td>
<td>12</td>
</tr>
<tr>
<td>1.2 Price Discrimination across PACs</td>
<td>14</td>
</tr>
<tr>
<td>1.2.1 Do Republican and Democrat PACs Pay the Same Prices?</td>
<td>14</td>
</tr>
<tr>
<td>1.2.2 Does Party Favoritism Explain Pricing?</td>
<td>16</td>
</tr>
<tr>
<td>1.3 Political Demand for Ad Spots</td>
<td>17</td>
</tr>
<tr>
<td>1.3.1 Effect of Advertising on Voting</td>
<td>17</td>
</tr>
<tr>
<td>1.3.2 Ad Selection</td>
<td>20</td>
</tr>
<tr>
<td>1.3.3 Instrument for Price</td>
<td>21</td>
</tr>
<tr>
<td>1.3.4 Estimating Equations</td>
<td>22</td>
</tr>
<tr>
<td>1.3.5 Heckman Selection Correction</td>
<td>23</td>
</tr>
<tr>
<td>1.3.6 Evidence on Willingness-To-Pay for Democrat and Republican PACs</td>
<td>24</td>
</tr>
<tr>
<td>1.4 Price Discrimination Model</td>
<td>26</td>
</tr>
<tr>
<td>1.4.1 Monopoly Pricing with Lowest Unit Rate Regulations</td>
<td>27</td>
</tr>
<tr>
<td>1.4.2 Station’s Optimal Pricing Condition</td>
<td>29</td>
</tr>
<tr>
<td>1.4.3 Testing Station Optimization</td>
<td>30</td>
</tr>
<tr>
<td>1.4.4 Do Prices Reflect Willingness-to-Pay?</td>
<td>32</td>
</tr>
<tr>
<td>1.5 Conclusion</td>
<td>33</td>
</tr>
</tbody>
</table>

2 Advertising Market Distortions from a Most Favored Nation Clause for Political Campaigns

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1 Introduction</td>
<td>54</td>
</tr>
<tr>
<td>2.2 Advertising in the 2012 Presidential Election</td>
<td>57</td>
</tr>
<tr>
<td>2.3 FCC, CPS, and Simmons Household Data</td>
<td>58</td>
</tr>
<tr>
<td>2.4 Effects of LUR on Campaign Prices</td>
<td>60</td>
</tr>
</tbody>
</table>
Chapter 1

Price Discrimination across Political Action Committees: Evidence from the 2012 Presidential Election

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 eliminating limits on contributions to PACs.² 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. As an example, they might offer cheaper prices PACs that support their favorite candidate. Alternatively, stations might charge PACs commensurately with their willingness-to-pay for airtime. Using novel data on television advertisement prices from the 2012 presidential election, this paper investigates how TV stations set prices for PACs. I estimate PAC demand for airtime, and test whether and how much prices reflect willingness-to-pay.

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).³ In contrast, newspaper

¹I 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.
³Lowest unit rate rules come into effect within 45 days of a primary election and 60 days of a general election. The Federal Election Campaign Act of 1971. There is some ambiguity in implementation of the law. Some law firms specialize in advertising law to help stations ensure they follow precedent.
headlines decry stations’ charging “super-gouge” rates to PACs. However, evidence on TV station pricing – to political or commercial advertisers – is thin.4

A first important finding is that on average, stations charge PACs 40% markups for airtime above the campaign price. Higher prices ought to temper the effect 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.5

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.6 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


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

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 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, rather than media bias, drive inequalities in access to viewers for PACs; this distinction is potentially important in shaping future PAC regulations.

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 concludes.

1.1 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.

1.1.1 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. 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. 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 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. 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%) for four reasons: (1) OCR software imperfectly parses photocopied invoices; (2) ads purchased prior to August 1st are not required to appear on the website, and so are not included here; (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

---


8 See Goldstein & Ridout (2004) for a detailed review of the literature.

9 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.

10 Fowler & Ridout (2013) estimate 1,431,939 were run from January 1, 2012 to election day.
my sample;¹¹ and (4) PACs explicitly focusing on non-presidential races are excluded from the analysis (approximately 16% of ads).¹² 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/C MAG’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). 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 race. Categorization of PACs is based on records from the Center for Responsive Politics.¹³

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; many donations arrive late in the election cycle;¹⁴ 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.¹⁵

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.¹⁶ 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

¹²discard observations at stations without dual PAC and campaign advertising.
¹³I 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.
¹⁴In the 2012 presidential race, October was the most lucrative month for both parties, followed by September, and then August (Ashkenas et al. (2012)).
¹⁵winsorize prices (1%) to mitigate the effect of outliers in the rest of the paper.
¹⁶stations 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.)
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.

1.1.2 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. In this section, I estimate ad exposures in the 2012 presidential race by combining survey data on viewership with market demographic data and data on ad purchases.

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.\(^{17}\)

Let \(j\) denote the program and \(g\) denote a demographic group (e.g. white women under 65 years of age). \(\pi_{gj}\) is the probability a member of group \(g\) sees ad \(j\), approximated by counts from the Simmons data. Let \(J_c\) 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 \(c\).

\[
A_{gsc} = \sum_{j \in J_c} \pi_{gj}
\]

Variation in ad viewership across states comes from demographic differences and differences in the composition of \(J_c\) (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

\(^{17}\)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.
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 groups. As an example, women are 20% more likely to watch a show than men (5.9% compared to 5%). 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 x 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. Spot viewership depends on local demographics and network programming. In total, there are 18,900 distinct products (36 stations x 15 weeks x 35 day parts).

Ad spots are often also described by a priority level, and an indicator for which particular days are permissible runtimes. 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

---

18During 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.

19If 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.
control for these features in demand estimation.  

1.2 Price Discrimination across PACs

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

1.2.1 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.  

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

---

20 For rejected shows, I assign characteristics in proportion to their presence in the purchased sample.
21 For this exercise, I consider only shows where the OCR software successfully scraped the full show name.
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 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
1.2.2 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. 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 \( p_D \) and \( p_R \) are the Democratic and Republican PAC prices, respectively.

\[
\psi = \frac{p_R - p_D}{p_R + p_D}
\]

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. A value of 0 corresponds to no discrimination, while values of the \( \psi \) close to 1 (-1) indicate a strong pro-Democratic -Republican) bias in pricing.

Figure 3 shows that across owners, Democrats receive more favorable rates than Republicans. Across the five companies, \( \psi \) 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, the Republican - Democratic 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 finding 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

---

\(^{22}\)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>.

\(^{23}\)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).
point of view, figure 3 indicates that Republican PACs would still benefit disproportionately from legislation prohibiting discrimination across political advertisers.

1.3 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)). 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.

1.3.1 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, $\gamma_{gc}$. It also depends on opponent’s advertising, $A_{gsc}$, and the efficacy of his advertising $\gamma_{gc}$ (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 $\beta_{gsc}$, and a random variable $\varepsilon_{sc}$. $\varepsilon_{sc}$ induces aggregate uncertainty in voting outcomes, and is important in rationalizing advertising in states that are ex-post uncontested. Political

\footnote{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.}
actors do not know which is the tipping-point state, the state whose electoral college vote decides the national election.\textsuperscript{25} Assume that these elements define a linear vote production function.

\[ V_{gsc} = \gamma_{gc} A_{gsc} - \gamma_{gc'} A_{gsc'} + \beta_{gsc} + \varepsilon_{sc} \] (1.1)

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.\textsuperscript{26} 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 \( \pi_{gj} \), the probability a member of group \( g \) sees ad \( j \), and \( f_{gs} \), the fraction of \( s \)’s population in group \( g \).

\[ S_{sc} = \sum_{g \in G} f_{gs} V_{gsc} = \varepsilon_{sc} + \sum_{g \in G} f_{gs} \beta_{gsc} + \sum_{g \in G} f_{gs} \left( \gamma_{gc} \sum_{j \in J_{es}} \pi_{gj} - \gamma_{gc'} A_{gsc'} \right) \]

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 \( \varepsilon_{sc} \) and \( \varepsilon_{sc'} \), own and rival’s ad choices, and state demographics:

\[
\mathbb{P}\{S_{sc} \geq S_{sc'}\} = \mathbb{P}\left\{ \varepsilon_{sc} - \varepsilon_{sc'} \geq \sum_{g \in G} f_{gs} (\beta_{gsc'} - \beta_{gsc}) + \sum_{g \in G} f_{gs} \left( (\gamma_{gc'} + \gamma_{gc'}) \sum_{j \in J_{es}} \pi_{gj} - (\gamma_{gc} + \gamma_{gc}) \sum_{j \in J_{es}} \pi_{gj} \right) \right\}.
\]

If I estimated (1.1) directly, then I could potentially estimate \( \gamma_{gc} \) and \( \gamma_{gc'} \) separately (although individual-level voting data would be needed to estimate \( \beta_{gsc} \)). \( \gamma_{gc} \) 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. \( \gamma_{gc} \) 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 \( \gamma_{gc} = \gamma_{gc} + \gamma_{gc'} \). \( \gamma_{gc} \) 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 \( \gamma_{gc} \), 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

\textsuperscript{25}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.)

\textsuperscript{26}In Nebraska and Maine, votes are split among districts. (FEC Office of Election Administration. “The Electoral College.” By William C. Kimberling. 1992.)
(the vote share) is not. The impact of advertising on candidate c’s vote share depends on the stock of own and rival advertising.\textsuperscript{27}

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

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

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.\textsuperscript{28} 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 \(\Delta_{jsc}\) 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 \((T)\) 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 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

\textsuperscript{27}Let \(V_{sc}\) be the vote share of candidate \(c\) in state \(s\).

\[
V_{sc} = \frac{\sum_g f_{gs} V_{gsc}}{\sum_g f_{gs}(V_{gsc} + V_{gsc'})} = \frac{\sum_g f_{gs}(A_{gsc}\gamma_{gc} - A_{gsc'}\gamma_{gc'})}{\sum_g f_{gs}(\gamma_{gc} A_{gsc} - \gamma_{gc'} A_{gsc'} + \varepsilon_{sc} - \varepsilon_{sc'} + \beta_{gsc} - \beta_{gsc'})}.
\]

\textsuperscript{28}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. In sum, the effect of ad \( j \) in support of \( c \) in state \( s \) is \( v_{jsc} \):

\[
v_{jsc} = \tau_s \Delta_{jsc} = \sum_{g \in G} f_{gs} \pi_{gj} \gamma_{gc}.
\] (1.2)

1.3.2 Ad Selection

The political advertiser employs (1.2) 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:

\[
\max_{\{J_{sc}\}} \mathbb{P}\{c \text{ wins the election}\}
\text{st:} \sum_{\{j \in J_{sc}\}} p_{jstc} \leq B
\]

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

\[
\alpha_c = \max_{j \notin \{J_{sc}\}} \left\{ \frac{\tau_s \Delta_{jsc}}{p_{jstc}} \right\}
\]

is the highest utility per dollar among ads not purchased. In other words, buy ads in descending order of utility per dollar until the budget is exhausted. Purchased ads then obey this decision rule. \( \alpha_c \) 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, \( \alpha_c \). To estimate these parameters, I incorporate two unobservable components into ad value: \( \epsilon_{jstc} \), known only to buyers, and \( \xi_{jstc} \), known to buyers and sellers. The econometrician observes neither. \( \epsilon_{jstc} \) introduces uncertainty, on the part of the station, as to exactly which ads political buyers value most, creating a downward sloping demand curve. \( \xi_{jstc} \) accommodates the typical concern in demand estimation that stations and advertisers have information about ad spots reflected in prices and quantities, but hidden.

---

29 In states of the world where Texas or California changes hands, their electoral college votes are gratuitous (extraneous to winning).
30 Without fractional purchases, set-optimization is challenging because it involves linear programming with integer constraints.
31 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.
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 $(\beta_c, \alpha_c)$ 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.

$$y_{jstc} = 1\{u_{jstc}^* \geq 0\}. \quad (1.3)$$

If $\epsilon_{jstc} \sim U[-\Gamma, \Gamma]$, then (1.3) becomes a linear probability model

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

1.3.3 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 $(\mathbb{E}[\epsilon_{jstc} | p_{jstc}] \neq 0)$. 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. Ansolabehere et al. (2001) investigate whether congressional advertising on television declines in districts with more incidental (uncontested) viewers.

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.

1.3.4 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

---

32 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.

33 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.

34 This assumption might be violated if PACs purchase ads in an effort to fundraise in uncontested states.
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. $k_{jstc}$ includes controls: week dummies, and priority level fixed effects, and the proportion of viewers living in states with contested senate races. All demographic variables are multiplied by viewers’ average tipping-point probability $\tau_s$. The following system describes demand

$$p_{jstc} = \phi_{0c} + \phi_{1c}\tau_s + \phi_{2c}z_{js} + \sum_{g=1}^{G} \tau_s f_{gs} \pi_{gj} \phi_{gc} + k_{jstc}' \phi_{3c} + \eta_{jstc}$$ (1.4)$$

$$y_{jstc} = \gamma_{0c} + \gamma_{1c}\tau_s - \alpha_c p_{jstc} + \sum_{g=1}^{G} \tau_s f_{gs} \pi_{gj} \gamma_{gc} + k_{jstc}' \gamma_{2c} + \epsilon_{jstc}$$ (1.5)

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 (1.4), in lieu of price on the right-hand-side of (1.5). I estimate standard errors using the nonparametric bootstrap, since predicted prices are generated regressors. For robustness, I re-estimate the model with daypart fixed effects, with an eye toward eliminating unobserved ad quality. Adding these fixed effects means estimation exploits only within hour/week-segment variation.

1.3.5 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 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

---

35 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.

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

37 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.

38 e.g. 8 PM Weekend or 6 AM Weekday
$p^*_{jstc}$ that is only observed if $y_{jstc} = 1$.

$$p^*_{jstc} = x_{jst} \beta_c + z_{js} \gamma_2 + \eta_{jstc} \tag{1.6}$$

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

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

Heckman (1979) devised a selection correction assuming $\epsilon, \eta$ distribute jointly normal with covariance $\rho$. In this model

$$y_{jstc} = 1\{x_{jst} \beta_c - \alpha_c p^*_{jstc} + \epsilon_{jstc} \geq 0\}$$

$$= 1\{x_{jst} \beta_c - \alpha_c (x_{jst} \gamma_1 + z_{js} \gamma_2 + \eta_{jstc}) + \epsilon_{jstc} \geq 0\}$$

$$= 1\{x_{jst} \pi_1 + z_{js} \pi_2 + \omega_{jstc} \geq 0\}$$

where $\omega = \epsilon - \alpha \eta \sim N(0, \alpha^2 \sigma_\epsilon^2 + \sigma_\eta^2 - 2 \alpha \rho \sigma_\epsilon \sigma_\eta)$, and $\sigma_\omega = 1$ is the free scale normalization. This specification allows for price endogeneity through unobserved product quality. 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 $\beta_c$ and $\alpha_c$. It is important that $z$ enter the selection equation only through its effect on prices, so that $\alpha_c = -\frac{\sigma_\eta}{\sigma_\omega}$, 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 (1.6) is a linear function of observed characteristics and an unobservable cost shock that distributes joint normal with the demand-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.

### 1.3.6 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 (1.5), 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 leading 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.\textsuperscript{40} 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:

\[
\text{average marginal effect of a } 1 \text{ std dev increase in } \% g = \frac{\hat{\gamma}_{pc} \sigma_{\pi g}}{N} \sum_{s=1}^{N} r_{sf_{gs}}
\]

A 5 percentage point (one standard deviation) shift in senior viewership increases the probability of purchase by 1.9 points for Republicans and 3 points 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.\textsuperscript{41} 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 is consistent with evidence from Ridout \textit{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.


25
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 (1.4). 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.

1.4 Price Discrimination Model

The big picture question are whether and to what extent willingness-to-pay matters for pricing, and whether the profit-maximizing station model explains pricing. Using demand estimates (equation 1.5), 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.
1.4.1 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 $\hat{P}(\hat{Q})$ 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.\(^{42}\) 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 campaigns. In sum, the station faces the following constrained optimization problem

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

st: $P(Q) \geq \hat{P}(\hat{Q})$ \hspace{1cm} (LUR 1)

$P_{PAC}(Q_{PAC}) \geq \hat{P}(\hat{Q})$ \hspace{1cm} (LUR 2)

$T \geq Q_{PAC} + Q + \hat{Q}$ \hspace{1cm} (Capacity Constraint)

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

\(^{42}\)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.
bonds. Let \( \pi_{\text{PAC}} \) be the profits from sales to campaigns and other advertisers:

\[
\pi_{\text{PAC}} = \left( \int_0^Q P(q) dq \right) + \hat{Q} \cdot \min \{P(Q), \hat{P}(\hat{Q}), P_{\text{PAC}}(Q_{\text{PAC}})\}.
\]

The opportunity cost is the change in \( \pi_{\text{PAC}} \) from an increase in \( Q_{\text{PAC}} \)

\[
-\frac{\partial \pi_{\text{PAC}}}{\partial Q_{\text{PAC}}} = -(P(Q) \frac{\partial Q}{\partial Q_{\text{PAC}}} + \hat{P}(\hat{Q}) \frac{\partial \hat{Q}}{\partial Q_{\text{PAC}}} + \hat{Q} \hat{P}'(\hat{Q}) \frac{\partial \hat{Q}}{\partial Q_{\text{PAC}}}).
\]  \hspace{1cm} (1.7)

Condition (1.7) 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 \( Q_{\text{PAC}}, P_{\text{PAC}}, \hat{Q}, \hat{P} \), 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 \( \frac{\partial Q + \hat{Q}}{\partial Q_{\text{PAC}}} = -1 \), so that (1.7) simplifies:

\[
-\frac{\partial \pi_{\text{PAC}}}{\partial Q_{\text{PAC}}} = P(Q) - \hat{Q} \hat{P}' \frac{\partial \hat{Q}}{\partial Q_{\text{PAC}}}.
\]

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

\[
P(Q) > -\frac{\partial \pi_{\text{PAC}}}{\partial Q_{\text{PAC}}} > P(Q) + \hat{Q} \hat{P}'(\hat{Q})
\]

Lowest unit rates overestimate marginal cost, since selling more units leads to inframarginal losses on units sold to campaigns. Based on estimates of campaign demand (tables 5a and b), \( \hat{P}'(\hat{Q}) \) 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 \( \frac{\partial Q + \hat{Q}}{\partial Q_{\text{PAC}}} \leq -1 \).

3. **Only the capacity constraint binds.** In this case, candidate demand is relatively low compared to other advertisers so that \( \hat{Q} = 0 \). The equation for opportunity cost (1.7) becomes \( -\frac{\partial \pi_{\text{PAC}}}{\partial Q_{\text{PAC}}} = P(Q) \), which is exactly the LUR. However, this rate is unobserved since campaigns do not purchase any ads. It is possible that \( Q_{\text{PAC}} = 0 \) 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.

### 1.4.2 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 \(c\):

\[
p_{jstc}^* = \arg\max_{p_{jstc}} (p_{jstc} - c_{jst})(1 - F_c(-(x_{jstc}\beta_c + \xi_{jstc} - \alpha_c p_{jstc})))
\]

\[
\implies p_{jstc}^* - c_{jst} = \frac{1 - F_c(-(x_{jstc}\beta_c + \xi_{jstc} - \alpha_c p_{jstc}^*))}{\alpha_c F_c(-(x_{jstc}\beta_c + \xi_{jstc} - \alpha_c p_{jstc}^*))}.
\]

(1.8)

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 ($\alpha_c$), different values for the same demographics ($\gamma_{gc}$), 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 $\epsilon_{jstc}$ distributes uniformly simplifies the FOC (1.8), so that it is separable in the cost and preference-driven components of price:

$$p_{jstc}^* = \frac{\Gamma}{2\alpha_c} + \frac{x_{jstc}\beta_c}{2\alpha_c} + \frac{c_{jst}}{2} + \frac{\xi_{jstc}}{2\alpha_c}$$  \hspace{1cm} (1.9)

### 1.4.3 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 (1.8). 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 (1.5)

$$\hat{u}_{jstc} = \frac{x_{jst}\hat{\beta}_c}{\hat{\alpha}_c}$$

$$p_{jstc} = \gamma_0 + \gamma_1 \hat{u}_{jstc} + \epsilon_{jstc}.$$  \hspace{1cm} (1.10)

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 $\gamma_1$ ought to be large, positive and statistically significant.

If marginal costs are small and there is limited variation in unobserved quality, then (1.10) also constitutes a test of the structural model (1.9). 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 $\gamma_1$ is
the coefficient on the "taste" component of pricing

\[ p_{jstc} = \gamma_0 + \gamma_1 \frac{x_{jst\beta_c}}{\tilde{\alpha}_c} + \gamma_2 c_{jstc} + \eta_{jstc}. \]  

(1.11)

OLS estimation of (1.11) is still potentially biased due to selection on unobservables. If stations price according to the monopoly model, then the residual in (1.11) is a function of unobserved ad quality: \( \eta_{jstc} \sim \xi_{jstc} \). Price is only observed conditional on purchase, so that \( \text{cov}(\eta_{jstc}, \frac{x_{jst\beta_c}}{\alpha_c}) \leq 0 \) 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 \( \gamma_c \). The conditional expectation of \( p_{jstc} \) given \( c \) purchases an ad with characteristics \( x_{jst} \) is:

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

I can estimate the expectation of the omitted quality term if I specify a distribution for \( \xi_{jstc} \).

Let \( \xi_{jstc} = \sigma_c \xi \sim N(0, \sigma^2_c) \). I model the CEF of \( \xi_{jstc} \) conditional on observables \( x_{jst} \), estimated demand parameters \( \alpha_c, \beta_c \), costs \( c_{jst} \), and purchase at the optimal price. (Conditioning on the observed price is not possible, since observed price is the dependent variable).

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

\[ = \frac{\sigma_c}{2} \int_{-\infty}^{\infty} \tilde{\xi}(\Gamma + x_{jst}\hat{\beta}_c + \tilde{\xi})d\tilde{\xi} + \frac{\sigma^2_c}{2} \int_{-\infty}^{\infty} \tilde{\xi}^2 \phi(\tilde{\xi})d\tilde{\xi} \]

\[ = \frac{\sigma^2_c}{x_{jst}\hat{\beta}_c + \Gamma} \]

Then I can test the FOC as:

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

(1.12)

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.\footnote{This would be a poor assumption, for example, if viewership (and ratings) are responsive to political advertiser identity.} The test specification is:

$$p_{jstR} - p_{jstD} = \gamma \left( \frac{x_{jst\beta_R}}{\alpha_R} - \frac{x_{jst\beta_D}}{\alpha_D} \right) + \psi_{jst}$$  \hspace{1cm} (1.13)

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

**1.4.4 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 (1.10). 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 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 (1.11). 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 (1.12) 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 $0.61 \text{ price increase per dollar of utility.}$

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. Selection could drive a correlation, in the purchased sample, between price and willingness-to-pay because buyers only purchase expensive spots when their willingness-to-pay is high. Under a pure selection story, however, stations sell both high and low willingness-to-pay slots to PACs at low prices; if stations successfully price discriminate, they never sell high willingness-to-pay slots at low prices. Figure 8 shows the distribution of transacted prices against willingness-to-pay; there are very few low transacted prices for the high WTP shows, consistent with a price discrimination story.

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 \text{ 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.

1.5 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 1971, Congress 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. 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 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.

---


34
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 1.2: 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 1.3: 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 1.4: 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 1.5: 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 1.6: 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. Observations are grouped into 20 bins according to estimated utilities, each containing five percent of the data.
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 willingness-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. Observations are binned into groups of five percentiles.
Figure 1.8: Prices vs Estimated Utility

(a) Republican PACs

(b) Democrat PACs

Notes: Figure 8 shows the relationship between observed prices and estimated utilities, without binning the data. Prices and utilities are measured per viewer in a contested state. The strong, positive correlation suggests stations price, at least in part, based on willingness-to-pay. Importantly, the bottom-end of the transacted price distribution shifts up for higher WTP slots.
### Table 1
Summary Statistics for Ad Spots Purchased by Political Group

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

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

<table>
<thead>
<tr>
<th></th>
<th>Democrat PACs</th>
<th></th>
<th>Republican PACs</th>
<th></th>
<th>Romney PACs</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Ages 18-64</td>
<td>Ages 65+</td>
<td>Ages 18-64</td>
<td>Ages 65+</td>
<td>Ages 18-64</td>
</tr>
<tr>
<td></td>
<td>Men</td>
<td>Women</td>
<td>Men</td>
<td>Women</td>
<td>Men</td>
</tr>
<tr>
<td>White</td>
<td>11.73</td>
<td>13.62</td>
<td>8.55</td>
<td>10.26</td>
<td>59.32</td>
</tr>
<tr>
<td>Black</td>
<td>11.80</td>
<td>14.18</td>
<td>9.63</td>
<td>11.95</td>
<td>67.05</td>
</tr>
<tr>
<td>Other</td>
<td>8.30</td>
<td>9.60</td>
<td>6.17</td>
<td>7.40</td>
<td>39.54</td>
</tr>
</tbody>
</table>

Obama Campaign

<table>
<thead>
<tr>
<th></th>
<th>Ages 18-64</th>
<th>Ages 65+</th>
<th>Ages 18-64</th>
<th>Ages 65+</th>
<th>Ages 18-64</th>
<th>Ages 65+</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Men</td>
<td>Women</td>
<td>Men</td>
<td>Women</td>
<td>Men</td>
<td>Women</td>
</tr>
<tr>
<td>White</td>
<td>68.58</td>
<td>79.43</td>
<td>46.65</td>
<td>56.59</td>
<td>30.97</td>
<td>36.20</td>
</tr>
<tr>
<td>Black</td>
<td>84.75</td>
<td>102.41</td>
<td>63.61</td>
<td>80.21</td>
<td>36.33</td>
<td>45.02</td>
</tr>
<tr>
<td>Other</td>
<td>30.91</td>
<td>51.63</td>
<td>68.58</td>
<td>36.98</td>
<td>14.74</td>
<td>23.46</td>
</tr>
</tbody>
</table>

Romney Campaign

<table>
<thead>
<tr>
<th></th>
<th>Ages 18-64</th>
<th>Ages 65+</th>
<th>Ages 18-64</th>
<th>Ages 65+</th>
<th>Ages 18-64</th>
<th>Ages 65+</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Men</td>
<td>Women</td>
<td>Men</td>
<td>Women</td>
<td>Men</td>
<td>Women</td>
</tr>
<tr>
<td>White</td>
<td>30.97</td>
<td>36.20</td>
<td>22.08</td>
<td>26.71</td>
<td>30.97</td>
<td>17.81</td>
</tr>
<tr>
<td>Black</td>
<td>36.33</td>
<td>45.02</td>
<td>29.29</td>
<td>37.31</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>14.74</td>
<td>23.46</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th></th>
<th>PACs (Republicans - Democrats)</th>
<th>Candidates (Romney - Obama)</th>
</tr>
</thead>
<tbody>
<tr>
<td>% Zero Price Difference</td>
<td>41.28</td>
<td>80.34</td>
</tr>
<tr>
<td>% Higher Republican Price</td>
<td>30.26</td>
<td>8.28</td>
</tr>
<tr>
<td>% Higher Democrat Price</td>
<td>28.45</td>
<td>11.38</td>
</tr>
</tbody>
</table>

Measure of Price Dispersion:

<table>
<thead>
<tr>
<th></th>
<th>PACs</th>
<th>Candidates</th>
</tr>
</thead>
<tbody>
<tr>
<td>Absolute Value of Price Difference</td>
<td>196.88 (12.63)</td>
<td>96.21 (12.98)</td>
</tr>
<tr>
<td>Absolute Value of % Price Difference</td>
<td>26 (3.00)</td>
<td>14 (2.00)</td>
</tr>
<tr>
<td>Raw Price Difference</td>
<td>68.41 (14.39)</td>
<td>-33.45 (14.02)</td>
</tr>
<tr>
<td>% Raw Price Difference</td>
<td>14 (3.00)</td>
<td>4 (3.00)</td>
</tr>
</tbody>
</table>

# Observations 717 290

Notes: Table 3 describes price differences between Republican and Democrat PACs for indisguishable ad purchases (ad purchases with the same show name, priority level, aired during the same week, at the same station). When there are multiple purchases 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

<table>
<thead>
<tr>
<th>Coefficient of Variation</th>
<th>(1) Full Sample</th>
<th>(2) Balanced Sample</th>
</tr>
</thead>
<tbody>
<tr>
<td>Across Republican &amp; Democrat PACs</td>
<td>0.11 (0.15)</td>
<td>0.14 (0.18)</td>
</tr>
<tr>
<td></td>
<td>622</td>
<td>224</td>
</tr>
<tr>
<td>Within Republican PACs</td>
<td>0.03 (0.10)</td>
<td>0.05 (0.10)</td>
</tr>
<tr>
<td></td>
<td>3400</td>
<td>224</td>
</tr>
<tr>
<td>Within Democrat PACs</td>
<td>0.00 (0.01)</td>
<td>0.00 (0.00)</td>
</tr>
<tr>
<td></td>
<td>664</td>
<td>224</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th></th>
<th>OLS (1)</th>
<th>First Stage (2)</th>
<th>IV (3)</th>
<th>IV with FE (4)</th>
<th>Heckman Selection coefficients (5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Price</td>
<td>-47.3*** (10.87)</td>
<td><strong>-32.37</strong>* (5.14)</td>
<td><strong>-82.26</strong>* (13.67)</td>
<td><strong>-36.09</strong>* (13.67)</td>
<td></td>
</tr>
<tr>
<td><strong>Tipping-point probability x Fraction in Demographic Group</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Women</td>
<td>45.16*** (1.44)</td>
<td>0.41*** (0.06)</td>
<td>65.38*** (5.73)</td>
<td>34.99*** (3.37)</td>
<td>80.38*** (12.05)</td>
</tr>
<tr>
<td>Aged 65+</td>
<td>31.08*** (3.99)</td>
<td>-0.06 (0.10)</td>
<td>26.05*** (6.80)</td>
<td>26.69*** (5.45)</td>
<td>68.25*** (18.06)</td>
</tr>
<tr>
<td>Black</td>
<td>-35.44*** (7.35)</td>
<td>0.74*** (0.19)</td>
<td>14.40 (17.38)</td>
<td>-1.66 (9.88)</td>
<td>-17.92 (33.30)</td>
</tr>
<tr>
<td>White</td>
<td>-25.25*** (6.74)</td>
<td>1.19*** (0.19)</td>
<td>50.08* (22.16)</td>
<td>13.97 (10.38)</td>
<td>18.40 (32.11)</td>
</tr>
</tbody>
</table>

**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>
<thead>
<tr>
<th></th>
<th>OLS (1)</th>
<th>First Stage (2)</th>
<th>IV (3)</th>
<th>IV with FE (4)</th>
<th>Heckman Selection coefficients (5)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Price</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>-41.52***</td>
<td>-36.98***</td>
<td>-189.01**</td>
<td>-25.115</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(8.883)</td>
<td>(6.366)</td>
<td>(71.768)</td>
<td></td>
</tr>
<tr>
<td><strong>Tipping-point probability × Fraction in Demographic Group</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Women</td>
<td>13.212***</td>
<td>0.078</td>
<td>17.50***</td>
<td>9.97***</td>
<td>54.49* 7.240</td>
</tr>
<tr>
<td></td>
<td>(0.990)</td>
<td>(0.061)</td>
<td>(2.959)</td>
<td>(3.043)</td>
<td>(25.506)</td>
</tr>
<tr>
<td>Aged 65+</td>
<td>32.521***</td>
<td>0.291**</td>
<td>42.23***</td>
<td>47.75***</td>
<td>226.31** 30.071</td>
</tr>
<tr>
<td></td>
<td>(2.957)</td>
<td>(0.095)</td>
<td>(5.891)</td>
<td>(6.773)</td>
<td>(86.505)</td>
</tr>
<tr>
<td>Black</td>
<td>-68.332***</td>
<td>-0.224</td>
<td>33.89***</td>
<td>-42.53***</td>
<td>-290.58** -38.611</td>
</tr>
<tr>
<td></td>
<td>(6.458)</td>
<td>(0.167)</td>
<td>(9.521)</td>
<td>(8.370)</td>
<td>(107.613)</td>
</tr>
<tr>
<td>White</td>
<td>-56.817***</td>
<td>0.098</td>
<td>-62.98***</td>
<td>-64.74***</td>
<td>-178.21* -23.680</td>
</tr>
<tr>
<td></td>
<td>(6.092)</td>
<td>(0.152)</td>
<td>(9.561)</td>
<td>(7.867)</td>
<td>(69.618)</td>
</tr>
<tr>
<td><strong>Instrument</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Viewers in uncontested states</td>
<td>0.203***</td>
<td></td>
<td>0.000 **</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.033)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Show Fixed Effects</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>18221</td>
<td>1976</td>
<td>18221</td>
<td>18221</td>
<td>Y</td>
</tr>
<tr>
<td>First-stage F-statistic</td>
<td>37.9</td>
<td></td>
<td></td>
<td></td>
<td>Y</td>
</tr>
<tr>
<td>Estimated Elasticity</td>
<td>-1.208</td>
<td>-1.076</td>
<td></td>
<td></td>
<td>-5.815</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th></th>
<th>Price Paid by Republican PACs</th>
<th></th>
<th>Price Paid by Democrat PACs</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
</tr>
<tr>
<td>Estimated Utility</td>
<td>0.623*** (0.020)</td>
<td>0.534*** (0.036)</td>
<td>0.598*** (0.049)</td>
<td>0.676*** (0.031)</td>
</tr>
<tr>
<td>Cost</td>
<td>0.210*** (0.050)</td>
<td>0.208*** (0.050)</td>
<td></td>
<td>0.374*** (0.082)</td>
</tr>
<tr>
<td>Constant</td>
<td>0.000 (0.000)</td>
<td>-0.002*** (0.000)</td>
<td>-0.004*** (0.001)</td>
<td>-0.001*** (0.001)</td>
</tr>
<tr>
<td>Expected Unobserved Utility</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>5204</td>
<td>2049</td>
<td>2049</td>
<td>1976</td>
</tr>
</tbody>
</table>

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>
<thead>
<tr>
<th>Sample:</th>
<th>Republican - Democrat Price Paid</th>
<th>Add-Ons</th>
<th>Price Difference</th>
<th>Price Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Utility Difference ($)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.277**</td>
<td>0.610*</td>
<td>0.392**</td>
<td>0.793*</td>
<td></td>
</tr>
<tr>
<td>(0.090)</td>
<td>(0.267)</td>
<td>(0.126)</td>
<td>(0.331)</td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>1501</td>
<td>196</td>
<td>996</td>
<td>103</td>
</tr>
</tbody>
</table>

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

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

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

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

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

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.
Chapter 2

Advertising Market Distortions from a Most Favored Nation Clause for Political Campaigns

2.1 Introduction

American election season is a boon for TV stations and a nightmare for TV advertisers. The influx of dollars in politically contested states drives up advertising rates, crowding out commercial advertisers (Sinkinson & Starc 2015). This crowd-out effect is potentially large; in swing states, political sales amount to as much as 70% of inventory for local stations.¹ Political advertising regulation potentially exacerbates this phenomenon. Regulation provides candidates to elected office a most favored nation in airtime sales. This regulation guarantees candidates the lowest price received by any other advertiser for airtime. But it also incentivizes stations to increase these lowest rates in order to extract rent from political campaigns. This sort of strategic response both undermines regulatory intent (to provide airtime to candidates cheaply), and also imposes a negative externality on local commercial advertisers.² I develop and estimate a model of television station behavior to explore whether and how much regulation increases the lowest unit rate for commercial advertisers and distorts advertising quantities. This paper provides evidence that regulation decreased campaign ad prices approximately 40%, but increased lowest rates two-fold in the 2012 presidential election. Further, findings indicate that stations sell less total advertising time in order to buoy campaign prices; stations sell approximately 10% less airtime in order to maintain high lowest unit rates.

The first contribution of this paper is to demonstrate that rate regulation affects cam-

²E.g. Kirchen (2012)
paign prices for airtime. Rate reduction was an express goal of regulation. Regulators hoped that lower campaign costs would encourage challengers (Lipsitz 2011). If, however, airtime were sold in a uniform price market, then regulators would need a different mechanism to accomplish this aim. Rate regulation is necessary and effective only if TV stations price discriminate across buyers, so that the unregulated station charges candidates and commercial buyers different prices. There is good reason to suspect stations of tailoring prices to each client. As an example, TV stations employ quantity discounts for large buyers of airtime (Bagwell 2007). There are even law firms that specialize in advising TV stations on compliance with LUR regulation, and in particular, on avoiding selling airtime too cheaply to candidates. Wiley Rein, a DC law firm, puts out a media guide warning stations against giving favored clients complementary spots during election season, lest they be forced to give politicians free time.\(^3\) Another guide counsels stations against selling unused inventory cheaply (a fire sale):

> “Before selling any spots at deeply discounted rates, a broadcaster should calculate the cost of rebating political advertisers down to the lowest unit charge established by the fire sale. In some instances, particularly when there is a heavy volume of political business on the air, it may be less expensive to retain the unsold spots.”\(^4\)

I exploit the timing of LUR regulation to identify its effect on prices. Rate rules take effect sixty days before the general election. Before the sixty-day mark, stations can charge campaigns prices higher than the commercial market, so long as stations treat political candidates equally. If candidate demand for airtime is less elastic than local commercial demand, then rate regulation should induce a large drop in prices. I document a 40% drop in prices around the cutoff based on a sample of 5,161 purchases during the 2012 presidential election. Prices fall most dramatically for daytime and graveyard airtime, slots commercial advertisers traditionally eschew. Lowest unit rate regulation seems to have bite across the board.

While campaign prices decline following the institution of LUR regulation, they may not fall to the unregulated lowest unit rates (the lowest price a commercial advertiser would have paid, absent regulation). Rate regulation inhibits price discrimination between commercial and campaign advertisers, so that a profit-maximizing station sets lowest commercial rates with campaign demand in mind. The lowest commercial price might therefore rise to meet

---


the campaign price. This constitutes a potentially efficient reallocation of airtime from the commercial to the campaign market, since uniform prices allocate a fixed amount of time efficient. But there is a second efficiency consequence: if LURs rise enough, total quantity might fall compared to a counterfactual without regulation. Stations may sell less total airtime in order to buoy lowest unit rates. Quantity withholding constitutes a cost of LUR regulation borne both by stations and commercial advertisers. The second contribution of this paper is to provide an empirical assessment of this externality.

Understanding the extent of quantity withholding in advertising markets can shed light on similar policies that anchor rates paid by a protected class to the wider market. As an example, Medicaid reimburses pharmaceutical companies at the average rate for the same drug among private insurers. Duggan & Scott Morton (2006) find evidence that these companies increase prices for prescription drugs with a large Medicaid audience. One contribution of this paper is to consider the quantity effects, which are first-order in assessing efficiency.

The dearth of detailed data on commercial advertising rates poses a challenge in estimating quantity withholding. The econometrician cannot directly measure the decline in total airtime or the increase in commercial lowest rates. My approach is to infer the decline in airtime based on observed outcomes on the campaign market and a model of station conduct. Stations' commercial and campaign sales are linked through two channels: capacity constraints and lowest unit rate regulation. Stations set lowest rates to maximize profits given these two constraints. I find the commercial airtime quantities that best rationalize the data, given a model of station optimization, campaign and commercial demand.

The identification strategy exploits the misalignment of political and media markets as in Strömberg (2008). Some media markets broadcast both to viewers who live in pivotal states and viewers who live in states where the presidential election is a foregone conclusion. Uncontested viewership influences the willingness-to-pay of commercial advertisers, but should not affect campaign demand, which is sensitive to state borders. This aids in tracing out separate campaign and commercial demand parameters.

I estimate withholding using a Tobit-style structural model and Bayesian MCMC techniques to avoid difficulties in gradient-based optimization of non-smooth functions. Using parameter estimates and a model of station pricing, I extrapolate the extent to which stations engage in quantity withholding. The results suggest a 10% reduction in advertising airtime.

This cost is incurred in an effort to broaden access to airwaves by keeping prices low. However, the lowest unit rate subsidy may accrue unequally across candidates. A candidate's benefit from regulation depends on the his target audience, and in particular, commercial advertisers' demand for that same audience. The Obama and Romney Campaigns' pur-

\[5\text{If viewers dislike advertising, then this constitutes an upper bound on the welfare loss from regulation.}\]
chasing patterns suggest a scope for their valuing demographics differently. Further work is necessary to determine the extent to which political advertising regulation has perverse distributional consequences by differentially subsidizing certain candidates over others. These findings highlight the importance of considering market imperfections in crafting government pricing policies.

The paper proceeds in the following order: in section 1, I describe the data; in section 2, provide evidence on how LUR regulation affects campaign prices. In section 3, I develop a model of station optimal airtime allocation between the commercial and campaign markets. Section 4 maps this to an empirical demand specification. Section 5 delineates a Bayesian MCMC estimation procedure to back-out the demand parameters. Results are presented in Section 6, including estimates of quantity withholding for the 2012 election cycle. I conclude in section 7.

2.2 Advertising in the 2012 Presidential Election

The 2012 presidential election saw record high levels of spending. The Obama and Romney campaigns spent $775 and $460 million respectively, the lion’s share on television advertising (Ashkenas et al. 2012, Mooney & Ailworth 2012). Political Action Committees (PACs) played an unprecedented role, buoyed by a series of 2010 judiciary rulings easing donation restrictions to these groups. Although PACs enjoy more freedom in fundraising, in the ad-buying arena, regulation favors official campaigns over PACs. Since the 1934 Telecommunications Act, the FCC mandates that stations treat two candidates to the same office equally and provide them with reasonable access to the airwaves. The Federal Election Campaign Act of 1971 introduced lowest unit rate regulation (LUR), which requires TV stations charge candidates lowest unit rates – the lowest rate paid by any commercial advertiser for comparable airtime – within sixty days of the general election and forty-five days of the primary.

These regulations aim to balance freedom of speech and competitive elections (Lipsitz 2011). The First Amendment protects paid political advertising as a form of speech, and advertising has become an integral part of American elections (Haberman 2014). Candidates spend approximately 75% of their budget on TV advertising (almost 500 million in the 2012 Presidential race) (Lipsitz 2011). Opponents of paid political advertising fear that the American reliance on airtime disadvantages challengers. Airtime is pricey, and incumbents

---


7 Obama for American and Romney for President were the official campaigns in 2012. Examples of PACs include Priorities USA (Obama) and Restore Our Future (Romney).
are thought to have a fundraising advantage. Other Western democracies ban paid political advertising for just this reason.\(^8\)

However, there is little evidence on the extent to which LUR regulation impacts advertising markets, particularly since stations have been left to interpret the mandate. Lipsitz (2011) postulates “in practice, however, the lowest unit rate is meaningless. Broadcast rates are in constant flux depending on the nature of the market for airtime. Stations charge different rates for various times of the day and seasons of the year.” Law firms catering to stations suggest their clients define lowest unit rates narrowly, for instance, at the program or week level, but not so finely that campaigns purchase a unique category of time.

However, if regulation carries weight, then profit-maximizing stations may sell less total advertising time than they otherwise would. A case study of five states during the 2006 midterm elections suggests that crowd-out of commercial advertising is more than one-for-one with political demand. The University of Wisconsin Media Project monitored the amount of political advertising and total advertising on news broadcasts (Midwest News Index) in the two months preceding the election. Milwaukee averaged 10 more seconds of political advertising compared to Madison (on a thirty minute broadcast), but 120 seconds less total advertising. The time surplus went chiefly to additional crime coverage. Although the sample is small, the negative correlation of total and political advertising suggests that regulation may reduce efficiency in advertising markets.

It is unsurprising that stations jockey to minimize impact of LUR regulation. Political demand represents big money to local stations because of its targeted nature. Network buys make little sense for presidential candidates because they broadcast in states where the election outcome is a forgone conclusion. Network buys make candidates pay for incidental viewers in cities like New York, Los Angeles or Houston. Buying on the local or national spot market provides a higher bang for the buck, and means a windfall for local affiliates. As an example, WHO-tv, the Des Moines NBC affiliate, devoted 70% of its advertising time to political ads in the run-up to the 2014 midterms (Kang & Gold 2014). For such stations, a strategic response to rate regulation is big money.

### 2.3 FCC, CPS, and Simmons Household Data

I merge three data sources to create a salient picture of airtime purchases and advertising views from the 2012 presidential race. The first source is invoices, order forms, and contracts collected from the Federal Communication Commission. In 2012, the FCC mandated stations post their political advertising sales to an online database, rather than maintain print copies

\(^8\)Belgium, France, Ireland, and the UK, among others.
in-house. In response to station concerns about compliance costs, the rule was phased in gradually. For the first two years, only stations broadcasting to the fifty largest media markets need to comply. An observation in the FCC data is a single ad slot (typically thirty seconds). Slot attributes include price, station, date, and program. This detail is essential in unpacking the effects of regulation, which might be camouflaged in the aggregate data commonly employed in the political advertising literature. There are 90,748 spots in total.

The second source of data is the Current Population Survey from 2010, which contains information about population at the zip code level. I aggregate the following demographics at the DMA level: men, women, seniors, minors, blacks, whites, and hispanics. These groups constitute the potential audience for each ad slot. Actual viewership depends on the proclivity of different groups to watch particular programs.

The final source of data speaks to those viewing habits: the Simmons 2010 household survey, which contains a sample of 25,000 American households. Respondents report their demographic information and also their TV viewing. Based on their responses, I construct the expected audience purchases for the Obama and Romney campaigns. As an example, the expected female viewership for a particular show $s$ in DMA $d$, $F_{ds}$, is a function of the number of women in that market, $N_{Fd}$, the number of female respondents in the household survey, $N_f$, and the number of those respondents who confirm watching the program in the preceding week, $N_{fs}$. I estimate the expected female viewership for show $s$ in DMA $d$ as

$$E[F_{ds}] = N_{Fd} \cdot P\{S|F\}$$

$$= N_{Fd} \frac{N_{fs}}{N_f}.$$ 

Table 1 reports summary statistics on the characteristics of ad purchases for the Romney and Obama campaigns. On average, the Obama campaign purchased cheaper programs, with smaller audiences. Compared to the Romney campaign, the average Obama purchase slanted towards a more Black, male, and young (under 65) audience. By law, stations must present the campaigns receive the same menu of choices, so these purchasing differences suggest underlying preference heterogeneity across candidates.

For the demand analysis, I normalize demographic variables by total viewership in contested states. As an example, I calculate the proportion viewers who are women and also live in Virginia for each ad spot. I then weight groups according to their importance in the electoral college, since some states carry more weight than others. I borrow Nate Silver’s estimates of state pivotality, the probability that a particular state decides the national election. Silver (2012) calculates the probability that each state is the least favorable state a

---

9 https://stations.fcc.gov  
candidate must win in order to earn 270 electoral college votes.\textsuperscript{11}

For the structural analysis, products are daypart and day type (weekday or weekend) combinations, rather than ad spots. "Dayparting" is the industry stratification method for grouping similar classes of airtime. It is composed of six bins: early (6 am - 9 am), daytime (10 am - 3 pm), early fringe (4 pm - 6 pm), primetime (7 pm - 10 pm), late fringe (11 pm - 1 am), and graveyard (2 am - 5 am).\textsuperscript{12} This stratification translates to twelve advertising products per station for each of ten weeks, leaving 3,720 products in total (31 stations).

Aggregation facilitates estimation by standardizing products across stations, and creating a schedule that allows me to reconstruct the product set candidates face. The raw data is based on transactions, so that omitting unpurchased products leads to traditional selection bias as in Heckman (1979). Product characteristics, such as the number of viewers, are averages across programs within each daypart. Since dayparts are constructed to bin similar programs (indeed airtime is sometimes sold at the daypart level), this simplification does not mask significant variation in the data. To confirm that aggregation retains key variation in the data, in table 2, I report the coefficient of variation for several key ad slot characteristics, both across and within dayparts for each station and week. For all characteristics, including price and viewership, the coefficient of variation across is twice as large as within.

Figure 1 shows the distribution of purchases for the presidential campaigns across dayparts. Compared to the Romney campaign, the Obama campaign favored less conventional ad spots like daytime television. This strategy, and his subsequent victory, was touted as the success of a "money-ball" approach to politics. Daytime slots are among the cheapest in terms of price per viewer, while primetime is the priciest.

\subsection*{2.4 Effects of LUR on Campaign Prices}

In this section, I investigate the effect of lowest unit rate regulation on campaign prices. As a first step, I exploit the timing of lowest unit rate regulation to test whether and how much regulation affects pricing. Lowest unit rate regulation comes into effect sixty days prior to the general election. Ideally, variation in regulation across races would identify the effect of regulation on advertising. However, LUR regulation applies to all races (even local ones). Instead, I exploit time series variation.

The goal of this exercise is similar in spirit to Duggan & Scott Morton (2006), who document price changes in response to Medicaid reimbursement rules: Medicaid pays the average price in the private market. Their data only permits them to examine market outcomes under the average-pricing rule, so they cannot explore the reaction of prices to

\textsuperscript{11}See Stromberg for a discussion of election pivotality.
\textsuperscript{12}On the weekend, bins are different.
the institution of the Medicaid pricing policy directly. Instead, they rely on cross-sectional variation in prices based on the relative size of Medicaid to private insurer demand for a product. In this setting, prices are observed under the regulated and unregulated regimes, since the policy stands only within sixty days of the election.

The average price of an ad spot falls 40% the week following the institution of lowest unit rate regulation, from $6.01 to $3.41 per thousand viewers. A t-test for equality of means rejects the null hypothesis of no price change at the 1% significance level. Although data on commercial rates are scarce, pre-LUR campaign rates appear higher than their commercial counterparts. SQAD, a media analytics company, puts the prime access cost per thousand households (CPM) at $16.68 in 2012. Campaign CPM averaged $19.79 for comparable time the week before LUR came into effect.\(^{13}\) This back-of-the-envelope calculation suggests that unregulated campaign prices were higher even than average commercial rates.

The price decline during the first week of September contrasts with an overall trend of rising prices through the election cycle, evident in figure 2. Several hypotheses in the advertising literature are consistent with this trend. First, a rapid decay in advertising effects may drive candidates to value time nearest the election (Hill \textit{et al.} (2013)). Second, since candidates must pay upfront for airtime (Nelson (2015)), last-minute donations may also bolster demand (and prices) as the election nears. Finally, candidates may delay purchases until election uncertainty resolves.

I hesitate to interpret the 40% decline as the causal effect of regulation on campaign prices because the institution of LUR regulation is anticipatable. Adjustment by market participants could confound our estimates. As an example, campaigns may postpone purchasing until they are guaranteed LURs. If the elasticity of substitution across weeks is high, then regulation might induce a decline in pre-period prices. In that case, we should interpret these estimates as strong evidence that LUR regulation affects market outcomes. A null result would not have been enough to rule out an effect, but a 40% drop rather confirms one.

A second concern is that the fall in average prices in the first week of September may represent a shift in the product mix purchased by campaigns, rather than in the prices for particular products. To be clear, campaigns might switch from expensive to inexpensive programming (on a per-viewer basis) at the sixty-day cutoff. For example, they might purchase more daytime slots or slots in cheaper media markets after LUR comes into effect. Figure 3 plots the Laspeyres and Paasche price indices over the three month period leading to the election. To control for potential product selection bias, these indices evaluate the relative cost of a fixed bundle of ads as prices change over time. The Laspeyres index uses the initial bundle observed in the data, the ad spots purchased in the first week of August, while the Paasche index uses the last bundle, or spots purchased in the first week of November.

\(^{13}\)Based on $7.62 cost per thousand viewers, and assuming 2.6 people per household. [http://www.tvb.org/trends/4718/4714]
Both indices show a sharp drop in relative prices (on the order of 50%) in the first week of September, when LURs come into effect. If anything, changes in the product mix mask changes in prices.

Prices decline more sharply for certain program types. Figure 4 plots the relative price change the week before versus after the cutoff separately by daypart. While prices fell across the board, graveyard and daytime programs experienced the largest percentage-point decline. These segments attract less affluent viewers since they air during standard sleeping and working hours, respectively, and are eschewed by commercial advertisers. One explanation for the precipitous decline for these programs is that presidential candidates might actually favor this less wealthy demographic. Stations then want to charge them relatively high prices compared to the lowest commercial rate. When LUR regulation comes into effect, the campaign price must fall considerably to meet its commercial counterpart. To the extent that candidates differentially value these untraditional demographics, they benefit unequally from LUR pricing. Figure 1 shows that the Obama campaign purchased more graveyard and daytime viewers (as a proportion of total purchases). While this pattern might be the artifact of higher price sensitivity (which seems unlikely as the Obama campaign was wealthier), it is also consistent with a preference difference between the Democrat and Republican candidates. If so, it suggests that LUR regulation favored the Obama campaign in 2012.

2.5 Station Quantity Decisions

In this section, I develop a model of TV station pricing behavior to illustrate how LUR regulation affects commercial advertisers. The model shows how the pricing restriction induces a tradeoff between exhausting advertising capacity and capturing rent from political campaigns. In subsequent sections, I bring this theoretical model to the data to quantify the externality regulation imposes on commercial advertisers. This sort of structural approach is well-suited to study quantity withholding using only campaign data. Commercial market data would allow a direct test of whether commercial prices rise at the sixty-day cutoff, which would correspond to the campaign price decline documented in the preceding section. The dearth of commercial data, which is considered proprietary, precludes this sort of analysis and leads me to a structural approach.\(^\text{14}\)

In determining how to set LURs, a TV station considers both campaign \(\hat{P}(\hat{Q})\) and commercial demand \(P(Q)\) (that might include PACs). The station only has a limited inventory of advertising time to sell. In the long-run, the hard constraint on time is the number of minutes per hour. In the short-run, however, advertising time is constrained by program

\(^\text{14}\)SQAD provides quarterly data by DMA, but the quarters poorly align with the election calendar, so that the effect of LUR would average out.
length. Stations can adjust total advertising time by running ads promoting their upcoming shows, trimming programs, or substituting local news for syndicated shows.\footnote{Lowest unit rates are potentially complicated by inventory exchange between networks and local affiliates. Typically networks and stations negotiate spot allocations in contracts that license programming. Inventory exchange systems are new, and evidence on their success and utilization is thin.} I assume that there is a maximum of $T < 60$ units of advertising per hour available for sale.

Stations act as single-product monopolists. This assumption requires that stations do not compete with each other (that the local FOX and NBC affiliates are poor substitutes), and they do not compete with themselves. The fast-pace of politics, coupled with the definition of products as daypart-week cells, suggests that the latter is not very restrictive. Major networks are few (ABC, CBS, FOX, and NBC), and viewers can watch at most one program at a time. Further, TV advertising is not a posted price market. Stations negotiate sales with commercial advertisers directly or through media agencies (McDowell (2006)). Although evidence on station pricing is thin – detailed data is proprietary – stations are suspected of setting different rates for different advertisers (Bagwell (2007)). This literature motivates modeling stations as perfectly price discriminating.

Absent LUR regulation, the station solves the following problem in deciding how to allocate time across political campaigns and commercial advertisers

$$\max_{\hat{Q}, Q} \pi(\hat{Q}, Q) = \int_{0}^{Q} P(q)dq + \hat{Q}\hat{P}(\hat{Q})$$

subject to:

$$T \geq Q + \hat{Q} \quad \text{(Capacity Constraint)}$$

In this counterfactual, stations are still precluded from discriminating across candidates. Instead, the station chooses a single price $\hat{P}$ to charge for all units sold to campaigns. If the profit-maximizing station exhausts its capacity, the commercial-political split equates the marginal revenue from the two markets according to the first order condition:

$$P(T - \hat{Q}) = \hat{P}(\hat{Q}) + \hat{Q}\hat{P}'(\hat{Q}) \quad (2.1)$$

On some programs, the value of airtime may be sufficiently low so that the station does not exhaust capacity. In that case, the station sells chooses quantities for campaigns and commercial advertisers such that their respective marginal revenue is zero.

$$P(Q) = 0 \quad (2.2)$$

$$\hat{P}'(\hat{Q})\hat{Q} + \hat{P}(\hat{Q}) = 0$$
Regulation imposes an additional constraint on TV stations: campaigns are entitled to "reasonable access" at the lowest unit rate. LUR regulation renders the allocations in (2.1) and (2.2) infeasible, since the profit-maximizing campaign price is strictly higher than the lowest commercial rate. In consequence, stations might reduce the total airtime they sell.

In sum, the station faces the following constrained optimization problem:

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

st: $T \geq Q + \tilde{Q}$  
$P(Q) \geq \tilde{P}(\tilde{Q})$.

(Capacity Constraint)  
(LUR)

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

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.

$$P^{LUR} = P(T) \geq \tilde{P}(0)$$

$$\implies \pi^* = \int_0^T P(q) dq$$

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.

$$P(Q^*) = \tilde{P}(T - Q^*)$$

$$\implies \pi^* = \int_0^{Q^*} P(q) dq + (T - Q^*) P(Q^*)$$

(2.3)

3. Only the LUR binds:

$$\max_P P \tilde{Q}(P) + \int_0^{Q(P)} P(q) dq$$

FOC: $\tilde{Q}(P) + P \tilde{Q}'(P) + PQ'(P) = 0$.  

(2.4)

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).
2.6 Empirical Demand Specification

In this section, I specify a structural model of campaign and commercial demand for advertising spots. Coupled with the preceding model of firm optimization, I can then estimate quantity withholding. I can also conduct other counterfactual analyses, including predicting station prices absent lowest unit rate regulation.

Each market is a week \( w \), station \( s \), daypart \( d \) combination (for example, primetime on WCCO-TV October 1st-7th).\(^{17}\) This leaves approximately 2,480 markets, since sales more than sixty days before the election are excluded. \( M_{wds} \) is the number of thirty-second intervals in market \( wds \), of which \( T_{wds} \) are potentially available for advertising. Campaign demand for slots depends on price \( p_{wds} \) and audience pivotality (state pivotality scaled by state population). Recall that stations may broadcast to audiences in different states, so viewer demographics are calculated separately by state, indexed by \( l \). \( t_l \) denotes the pivotality of viewers in state \( l \). I follow Moshary (2015) in the parametrization of viewer demographics in advertising demand. She delineates a model of ad impacts on the probability of winning the election that maps to this empirical specification. The key characteristics are viewer demographics \( f_{sdlg} \), the proportion of daypart \( d \), station \( s \) viewers who live in state \( l \) and are members of demographic group \( g \). The value of each slot is

\[
v_{wdsk} = \beta_0 \sum_{l \in L} t_l f_{sdl} + \sum_{g \in G} \beta_g \sum_{l \in L} t_l f_{sdg} - \alpha p_{wds} + \gamma_w + \beta_H H_{wds} + \xi_{wds} + \epsilon_{wds}.\]

The set of demographics included in \( G \) are: female, black, white, Hispanic, and seniors. \( H_{wds} \) is a dummy variable for whether the purchase was high priority, which proxies for the probability of preemption. The component of ad quality unobserved by the econometrician, but known to advertisers and stations, is \( \xi_{wds} \). There is also an unobservable taste shock \( \epsilon_{wds} \), known only to the advertiser, presumed to distribute type I extreme value.\(^{20}\) The share of total broadcast time demanded by campaigns in market \( wds \) is modeled as

\[
\tilde{s}_{wds} = \frac{\exp \left\{ \beta_0 \sum_{l \in L} t_l f_{sdl} + \sum_{g \in G} \beta_g \sum_{l \in L} t_l f_{sdg} - \alpha p_{wds} + \gamma_w + \beta_H H_{wds} + \xi_{wds} \right\}}{1 + \exp \left\{ \beta_0 \sum_{l \in L} t_l f_{sdl} + \sum_{g \in G} \beta_g \sum_{l \in L} t_l f_{sdg} - \alpha p_{wds} + \gamma_w + \beta_H H_{wds} + \xi_{wds} \right\}}.
\]

\(^{17}\) Dayparts include early (5am-9am), daytime (9am-5pm), news (5pm-7pm), primetime (7pm-11pm) and late night (11pm-5am).

\(^{18}\) I assume 27 slots are available per hour. 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

\(^{19}\) 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.

\(^{20}\) 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.
The corresponding campaign demand function is simply \( \hat{Q}_{\text{wds}}(P_{\text{wds}}) = M_{\text{wds}}\hat{s}_{\text{wds}} \).

Commercial advertisers include PACS, who are likely to value airtime similarly to campaigns. Therefore, the commercial demand function includes all of the covariates listed above. However, commercial value a larger set of covariates than their campaign counterparts. In particular, they ought to value viewers in uncontested states, so I include demographics unscaled by pivotality \((t_i)\). Commercial advertisers also to have a different unobserved utility component \(\omega_{\text{wds}}\) and a separate logit shock. The commercial and campaign unobservables \((\xi_{\text{wds}}, \omega_{\text{wds}})\) distribute bivariate normal with variances \(\sigma^2_\xi, \sigma^2_\omega\) and covariance \(\rho\), which I estimate from the data. The additional covariates in commercial demand help disentangle covariance between unobservables and common taste for observable characteristics. The share purchased by commercial buyers is

\[
\begin{aligned}
\hat{s}_{\text{wds}} &= \frac{\exp\{\delta_{\text{wds}}\}}{1 + \exp\{\delta_{\text{wds}}\}} \\
\delta_{\text{wds}} &= \sum_{l \in \mathcal{L}} (f_{sdl}(\gamma_0 + \gamma_1 t_l)) + \sum_{g \in \mathcal{G}} \left( \gamma_g \sum_{l \in \mathcal{L}} t_l f_{sdlg} + \delta_g f_{sdg} \right) \\
&\quad - \gamma_p P_{\text{wds}} + \gamma_w + \gamma_H H_{\text{wds}} + \omega_{\text{wds}}.
\end{aligned}
\]

In this model, given a price \(P_{\text{wds}}\), commercial and campaign demand may exceed capacity (the shares need not sum to one). However, stations set prices after observing \((\xi, \omega)\), so that capacity constraints are never violated in equilibrium.

A guess of the parameter vector \(\theta = (\tilde{\beta}, \tilde{\gamma})\) coupled with data on attributes and the campaign share maps to single value of unobserved campaign taste

\[
\xi_{\text{wds}} = \ln \left( \frac{\hat{s}_{\text{wds}}}{1 - \hat{s}_{\text{wds}}} \right) - \beta_0 \sum_{l \in \mathcal{L}} t_l f_{sdt} - \sum_{g \in \mathcal{G}} \beta_g \sum_{l \in \mathcal{L}} t_l f_{sdlg} + \alpha p_{\text{wds}} - \beta_w - \beta_H H_{\text{wds}}. \tag{2.6}
\]

I use this inversion to construct a likelihood function.

### 2.6.1 Correcting for Unobserved Commercial Quantity

The target likelihood function evaluates the probability of a joint draw of commercial and campaign demand shocks implied by a candidate parameter vector \(\theta\). The mapping from data to \(\omega_{\text{wds}}\) is tricky because commercial quantities are unobserved. Given campaign price \(P_{\text{wds}}\) and demand parameters \(\theta\), 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 (2.4). I do not know which of these two conditions gave rise to any particular market \(wds\) observed in the data, but I know that exactly one did.

There are two potential commercial shares that correspond to the observed campaign
1. If residual supply is exhausted (the station hits its capacity constraint as in (2.3)), then I can back-out the commercial quantity demand from a boundary solution \( B, s_wds \), simply through an adding up constraint:

\[ s_{wds} = \frac{T_{wds}}{M_{wds}} - \tilde{s}_{wds}. \]

Inverting the logit share function gives the mean commercial utility implied by a boundary solution, \( \delta^B_{wds} = \ln \left( \frac{s^B_{wds}}{1 - s^B_{wds}} \right) \). Coupled with a candidate preference parameter vector, mean utility maps to a demand taste shock for the commercial market, \( \omega^B_{wds} \).

\[
\omega^B_{wds} = \delta^B_{wds} - \sum_{l \in L} (f_{sl}(\gamma_0 + \gamma_t t_l)) - \sum_{g \in G} \left( \gamma_g \sum_{l \in L} t_l f_{sdg} + \delta_g f_{sdg} \right) \tag{2.7}
\]

\[
+ \gamma_p p_{wds} - \gamma_w - \gamma H_{wds}
\]

The likelihood of the observed price and campaign share at a boundary optimum is then a transformation of the likelihood of \( (\xi_{wds}, \omega^B_{wds}) \).

2. If, on the other hand, the capacity constraint does not bind, then I recover the share and demand shock that correspond to an interior solution to the firm’s optimization problem. From the first order condition (2.4), the interior solution commercial share is given by the quadratic equation:

\[
M_{wds} \tilde{s}_{wds} - p_{wds} M_{wds} \tilde{s}_{wds} \alpha (1 - \tilde{s}_{wds}) = \gamma_p M_{wds} p_{wds} s^I_{wds} (1 - s^I_{wds})
\]

The corresponding interior mean utility \( \delta^I_{wds} = \ln \left( \frac{s^I_{wds}}{1 - s^I_{wds}} \right) \) can then be inverted to recover a second value for the commercial demand shock:

\[
\omega^I_{wds} = \delta^I_{wds} - \sum_{l \in L} (f_{sl}(\gamma_0 + \gamma_t t_l)) - \sum_{g \in G} \left( \gamma_g \sum_{l \in L} t_l f_{sdg} + \delta_g f_{sdg} \right) \tag{2.8}
\]

\[
+ \gamma_p p_{wds} - \gamma_w - \gamma H_{wds}
\]

Evaluating \( (\xi_{wds}, \omega^I_{wds}) \) using the bivariate normal, a posited covariate matrix, and a Jacobian for the change-of-variables allows me to construct the likelihood of the observed commercial share and price at an interior optimum.

The goal of the empirical exercise is to discern how often stations price according to the interior rather than boundary solution. Equations (2.7) and (2.8) map \( \theta \) to two potential
draws of the unobservable for each observed market \( wds \). The likelihood of observing price \( p_{wds} \) and campaign purchase \( s_{wds} \) is then the probability of observing either of these outcomes. I can then use the likelihood function to evaluate the relative probability of (2.8) versus (2.7), which is probability of quantity withholding.

### 2.6.2 Zero Shares

The data contains many instances of zero campaign shares (1,330 zero shares in 2,480 markets), which hinders inversion to find taste shocks \( (\xi_{wds}) \) in (2.6). If airtime were perfectly divisible, then the campaign share is strictly bounded away from zero under the logit specification. Because of the logit shock \( \epsilon_{wds} \), the first infinitesimal amount of airtime is infinitely valuable to campaigns, so a station always allocates them airtime. Rather than faulting the expected logit share as a poor approximation to its empirical counterpart, I approach this as a missing data problem. I do not observe the true quantity when it falls below the single-unit threshold:

\[
\tilde{Q}_{wds} = \begin{cases} 
M_{wds}s_{wds} & \text{if } M_{wds}s_{wds} \geq 1 \\
0 & \text{if } M_{wds}s_{wds} < 1 
\end{cases}
\]

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 markets with few consumers, a realized zero share might mask a very high expected share. Rather, the difficulty here is that the data is 'binned' after it is generated. A zero share, in my setting, rules out the possibility that the expected share was higher than \( \frac{1}{M_{wds}} \).

### 2.7 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 implies campaigns care about viewers in so much as they are pivotal in the 2012 presidential election. I borrow estimates of state pivotality from Nate Silver.\(^{21}\) His estimates align with campaign ad choices. As an example, he pegged Ohio at 50% odds of playing the pivotal role, and approximately 30% of political ads in my dataset air in Cincinnati, Cleveland and Columbus. In contrast, commercial advertisers value

\(^{21}\)Available on his blog 538.com.
viewers who live in uncontested states, so that cross-media market variation in uncontested viewership helps trace out campaign demand for airtime.

Ideally, a second exclusion restriction would identify commercial demand parameters. Unfortunately, the emergence of Political Action Committees makes finding a covariate that affects campaign demand, but not commercial demand, difficult. PACs are included in the commercial market because they are not protected by LUR regulation, but they are likely to value the same viewership as campaigns. Instead, I take a control function approach that relies on the model of station conduct in section 4. To see how the behavioral assumption aids in identification, consider the station’s first order condition with respect to the lowest unit rate $\hat{P}$:

\[
\frac{\hat{P}Q'(\hat{P}) + \hat{Q}(\hat{P})}{\text{campaign marginal revenue}} = \frac{-\hat{P}Q'(\hat{P}) + \lambda(\hat{Q}'(\hat{P}) + Q'(\hat{P}))}{\text{commercial marginal revenue + shadow value of capacity}}.
\]

There are three separate components: (I) the marginal revenue from campaigns $\hat{M}R$, (II) the marginal revenue from commercial advertisers, $MR$ and (III) the shadow value of the capacity constraint. The marginal revenue from campaigns is identified by the exclusion restriction discussed above, but what moments in the data tell us about the parameters of commercial demand that appear in (II) and (III)?

First, notice that the sign of (II) is always positive because the station can perfectly price discriminate. Setting a lower lowest unit rate does not cannibalize profits on inframarginal units sold to commercial advertisers. Where the station faces negative marginal revenue from campaigns – in instances where the inframarginal loss is relatively small – then the station must be at the capacity constraint for the FOC to hold. If the marginal revenue from campaigns (I) is negative, then the station unambiguously wants to lower price to sell more units, but cannot do so; both commercial and campaign advertisers prize those ad spots. Marginal revenue from campaigns, $\hat{s}(1 - \hat{\alpha}\hat{P}(1 - \hat{s}))$, is small whenever price $\hat{p}$ is large and campaign demand $\hat{s}$ is low. Characteristics that covary positively with large $\hat{p}$ and small $\hat{s}$ (which are both observed) must be valued highly by the commercial market.

Second, notice that there is an upper bound on the RHS of (2.9):

\[-\hat{P}Q'(\hat{P}) + \lambda(\hat{Q}'(\hat{P}) + Q'(\hat{P})) \leq \alpha\hat{P}s(1 - s) \leq \alpha\hat{P}\left(\frac{T}{M} - \hat{s}\right) \left(\frac{M - T}{M} + \hat{s}\right).
\]

Instances where $\hat{M}R$ is positive and large provide a lower bound for the commercial price...
coefficient, \(a\). In order to rationalize these cases, commercial demand must be sufficiently elastic.

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 consistency requires the number of simulations grow faster than the number of observations (Train (2009)). Since simulations are computationally expensive, in practice, the number of draws per observation is constrained. A more serious concern is that the likelihood function is not smooth in the parameter space.\(^{22}\) 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.\(^{23}\) 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 sample is

\[
A = \sum_{w,d,s} T_{wds}.
\]

For each market, \(wds\), the probability of quantity withholding is the relative likelihood of an interior rather than boundary solution. The amount of inventory withheld at the interior solution is simply the residual airtime unsold on either the campaign or commercial market. I calculate the fraction of airtime unsold, \(L\), for each MCMC step

\[
L = \frac{1}{A} \sum_{w,d,s} \left( T_{wds} - M_{wds} s_{wds} - M_{wds} s_{wds}^I \right),
\]

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

\subsection*{2.8 Campaign versus Commercial Preferences}

Table 4 presents parameter estimates and credible intervals from a random-walk metropolis chain of 500,000 draws with a burn-in of 50,000 draws. The estimated correlation between the commercial and campaign taste shocks is positive and large (0.77). This correlation coefficient is consistent with an unobserved quality dimension valued by both groups of advertisers. As an example, both commercial and campaign advertisers might prize primetime

\footnote{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 \(10^{-\delta}\) where \(\delta = 10 + 10 \cdot \text{fraction unrationalizable}\). In practice, this amounts to approximately 3\% of observations.}

\footnote{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.}
shows, which may attract viewers who are otherwise hard to reach (Phillips & Young (2012)). The estimated variance parameters are large compared to the mean utility of ad products. These parameters rationalize the variation in campaign shares across ad products in the data. Campaigns have no observed purchases for many 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 (-26.04) than its commercial counterpart (-91.51). The campaign price elasticity corresponds to an average demand elasticity of -0.37; inelastic demand in equilibrium is consistent with capacity constraints. A shallower demand curve is consistent with campaigns’ having limited alternative advertising opportunities relative to commercial advertisers. Campaigns prioritize tipping-point DMAs (the markets studied here) 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 commercial mean utility is higher than its campaign counterpart. High mean utility reconcile the price coefficients with the empirical regularity that commercial advertisers still purchase the lion’s share of airtime.

Apart from the price coefficient, parameter estimates suggest campaigns value viewers in states more likely to play the tipping-point roll. Since campaign demand encompasses both the Obama and Romney campaigns, it is hard to interpret the campaign preferences for racial groups. However, there is a strong preference for older viewers, who are more likely to turn out the polls. The coefficient on weeks to the election is negative and statistically significant, consistent with advertising being most valuable near election day. High-priority advertisements, which are less likely to be preempted, are also more valuable, both to commercial and campaign advertisers.

### 2.9 Evidence on Quantity Withholding

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 quantity withholding is 10.5% of total available advertising time. The credible set (analogous to a 95% confidence interval) extends from 9.8% to 11%, so the estimate is fairly precise. In other words, regulation reduces advertising by 1.3 thirty-second spots each hour. This estimate combines both the probability that withholding occurs and the quantity withheld at the interior solution. Table 5 contains separate estimates of these two components. To be clear, these estimates are for time slots where campaigns purchase. Stations have no incentive to withhold on slots that contain no campaign advertising. The average probability of withholding is 75.2%, and average quantity withheld at an interior solution is 16.4% of advertising time. These findings suggest the distortionary effects are of first-order importance in evaluating LUR regulation.
The burden of regulation falls unequally across media markets, even among swing states. Stations in Boston do not distort their sales, while those in Las Vegas sell two fewer spots per hour (just above 15% of advertising time). Figure 5 shows the distribution of withholding across DMAs. These estimates suggest that distortions are borne unequally across stations and commercial advertisers. In particular, markets with a relatively large population of incidental (politically irrelevant) viewers have low levels of estimated withholding. This finding is consistent with legal advice to stations to withhold quantity “when there is a heavy volume of political business on the air.”

2.9.1 How much does regulation inflate commercial rates?

The structural demand model allows me to simulate outcomes in a counterfactual without lowest unit rate regulation, and in particular, to estimate the change in commercial lowest unit rates. I find the expected unconstrained price for each program where campaigns purchased time in 2012. The optimal price depends on the draw of the taste shocks \((\xi, \omega)\), and therefore I can only find the expected counterfactual price, since estimation provides only a distribution over \((\xi, \omega)\) for each program. The unconstrained firm sets prices according to (2.1). I find the optimal price separately under the shocks implied by interior solution and a boundary solution.

This counterfactual exercise illustrates how prices would change absent LUR regulation, but it does not incorporate certain general equilibrium effects. As an example, an increase in campaign prices could affect donations. Supporters might substitute donations away from campaigns toward Political Action Committees or lobbyists. The counterfactual I simulate holds constant the marginal utility of money \((\bar{\alpha} \text{ and } \alpha)\), precluding income effects under alternative pricing schemes. Gordon & Hartmann (2013) also employ this assumption. They argue that the marginal value of money to political donors is invariant to the price of advertising.

To simulate the counterfactual with LUR regulation, I calculate the expected increase in the prices paid by campaigns. I calculate the price increase as:

\[
E[P(\theta, x_{\text{wds}}, \bar{x}_{\text{wds}}, p_{\text{wds}}, s_{\text{wds}})] = P^{\ast}(\omega^I_{\text{wds}}, \xi_{\text{wds}}) \cdot \frac{\mathbb{P}\left\{\omega^I_{\text{wds}}, \xi_{\text{wds}}\right\}}{\mathbb{P}\left\{\omega^I_{\text{wds}}, \xi_{\text{wds}}\right\} + \mathbb{P}\left\{\omega^B_{\text{wds}}, \xi_{\text{wds}}\right\}} + P^{\ast}(\omega^B_{\text{wds}}, \xi_{\text{wds}}) \cdot \frac{\mathbb{P}\left\{\omega^B_{\text{wds}}, \xi_{\text{wds}}\right\}}{\mathbb{P}\left\{\omega^I_{\text{wds}}, \xi_{\text{wds}}\right\} + \mathbb{P}\left\{\omega^B_{\text{wds}}, \xi_{\text{wds}}\right\}}.
\]

Results, presented in table 5, suggest that the expected increase is on the order of 80%. Importantly, I can compare this estimate to the observed price increase at the 60-day cutoff.


72
when lowest unit rate regulation comes into effect. Since campaign purchase data from this period is not used to estimate the demand parameters, this comparison allows me to cross-validate the model. Campaign prices are approximately 60% higher before LUR regulations are instituted, which is within the credible set for the model-predicted price increase. That the model predictions are consistent with observed outcomes suggests the model captures the salient features of the political advertising markets. The model also predicts a 53% decline in lowest unit rates for commercial advertisers in the switch to an unconstrained regime. This effect is for programs where campaigns purchased airtime in 2012, so the effect on national lowest unit rates should be much smaller, this estimate suggests a considerable distortion in commercial advertising rates.

2.10 Conclusion

To foster competition in American elections, the FEC mandates that legitimate candidates to political office receive the lowest advertising rates stations charge their commercial clients. The goal is to reduce the cost of campaigning, but this regulation has consequences outside of politics. In marrying the campaign price to the lowest price paid by any other advertiser, regulation incentivizes stations to keep lowest unit rates high. Thus, to extract rent from campaigns, stations may reduce airtime sales to commercial advertisers.

Both the efficacy and distortion of this regulation depend on the extent to which local stations price discriminate across buyers. Were local broadcast markets perfectly competitive, then this most favored nation would be inconsequential, since all advertisers would pay the same rate. The first contribution of this paper is to show that rate regulation affects market outcomes. I exploit the timing of regulation to test whether LUR regulation affects advertising markets. Rate regulation comes into effect 60 days before the general election. Using a novel database of campaign advertising purchases, this paper documents evidence that campaign rates drop by approximately 40% the week after this cutoff.

Findings suggest that the quantity distortion is also large, on the order of 10% of total advertising time. Recovering the decline in airtime sales is challenging because detailed data on commercial advertising is considered proprietary and not available to researchers. These estimates of quantity withholding are based on a structural approach, which relies on a model of station conduct for identification. The key inputs to the model are commercial and campaign demand for advertising, which are connected through the LUR regulation and a capacity constraint on total airtime per hour. The model permits recovery of unobserved commercial sales based on observed optimal campaign sales. The estimates suggest that quantity withholding is concentrated in media markets with large share of contested voters, and raises commercial rates approximately twofold. The model predictions are consistent with the observed price decline at the sixty-day mark when LUR come into effect.
If advertising poses only a small negative externality on viewers, then the estimated quantity withholding constitutes a large loss in efficiency. The recent influx of money into politics, such as the 2014 Supreme Court ruling to abolish limits on total campaign donations, will likely exacerbate this externality. These estimates also suggest 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.

\footnote{McCutcheon v. FEC, 134 S.Ct. 1434 (2014), No. 12-536.}
2.11 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 \( \hat{\delta}_i \), and the implied demand shock: \( \xi_i = \hat{\delta}_i - \bar{x}_i\beta + \alpha p_i \). 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 \( P\{p_i, \bar{s}_i\} \) be the probability of observing price \( p_i \) and campaign demand \( \bar{s}_i \) (with corresponding campaign mean utility \( \hat{\delta}_i \)):

\[
P\{p_i, \bar{s}_i\} = P\{p_i, \bar{s}_i, \text{ boundary solution}\} + P\{p_i, \bar{s}_i, \text{ interior solution}\}.
\]

Observed outcomes \( p_i \) and \( \bar{s}_i \) 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:

\[
\delta_i^B = \ln \left( \frac{\bar{t}_i/M_i - \bar{s}_i}{1 - \bar{t}_i/M_i + \bar{s}_i} \right).
\]

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 \( (\hat{\delta}, p) \) to \( (\xi, \omega) \).

\[
P\{\hat{\delta}, p, \delta = \delta^B\} = P\{\xi = \hat{\delta} - \bar{x}\beta + \alpha p, \omega = \delta^B - x\beta + \alpha p\} 1\{p = p^*(\omega, \xi)\}
\]

\[
\begin{vmatrix}
\frac{\partial \xi}{\partial \hat{\delta}} & \frac{\partial \xi}{\partial p} \\
\frac{\partial \omega}{\partial \hat{\delta}} & \frac{\partial \omega}{\partial p}
\end{vmatrix}
\]
The elements of the Jacobian for the change-of-variables between observed mean utilities \( \tilde{\delta} \) and prices \( p \) are:

\[
\begin{align*}
\frac{\partial \xi}{\partial \tilde{\delta}} &= 1 & \frac{\partial \xi}{\partial p} &= \tilde{\alpha} \\
\frac{\partial \omega}{\partial \tilde{\delta}} &= -1 & \frac{\partial \omega}{\partial p} &= \alpha
\end{align*}
\]

\[ \Rightarrow \text{det} = \alpha + \tilde{\alpha}. \]

Re-writing the probability:

\[
P\{\tilde{\delta}, p, \delta = \delta^B\} = P\{\xi = \tilde{\delta} - \tilde{x}\tilde{\beta} + \tilde{\alpha}p, \omega = \delta^B - x\beta + \alpha p\} \cdot 1\{p = p^*(\omega, \xi)\}|\alpha + \tilde{\alpha}|.
\]

**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:

\[
s' = \frac{1}{2} \pm \sqrt{\frac{\frac{1}{4} - \tilde{s} - \tilde{\alpha}\tilde{s}(1 - \tilde{s})}{\alpha p}}.
\]

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:

\[
P\{\tilde{\delta}, p, \delta = \delta'\} = P\{\xi = \tilde{\delta} - \tilde{x}\tilde{\beta} + \tilde{\alpha}p, \omega = \delta^{(1)} - x\beta + \alpha p\} \cdot 1\{s^{(1)} \in [0, \frac{T}{M} - \tilde{s}]\} \cdot \begin{vmatrix} \frac{\partial \xi}{\partial \tilde{\delta}} & \frac{\partial \xi}{\partial p} \\ \frac{\partial \omega}{\partial \tilde{\delta}} & \frac{\partial \omega}{\partial p} \end{vmatrix}.
\]

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

The Jacobian is not the same as in the boundary solution case, since the relationship between \( \omega_i \) and \( p_i \) is now given by the FOC.
\[
\frac{\partial \omega}{\partial \delta} = \frac{\partial \delta \partial s \partial \delta}{\partial s \partial \delta \partial \delta} = \frac{1}{s(1-s)} \frac{\partial \delta \partial s \partial \delta}{\partial s \partial s \partial s} = \frac{1}{2s(1-s)} \left( \frac{1}{4} - \frac{s - \alpha p s(1-s)}{\alpha p} \right)^{-\frac{1}{2}} \frac{\alpha p (1-2\tilde{s}) - 1}{\alpha p} \frac{\partial s}{\partial \delta}
\]

\[
= \pm \frac{s}{2s(1-s)} \left( \frac{1}{4} - \frac{s - \alpha p s(1-s)}{\alpha p} \right)^{-\frac{1}{2}} \frac{\alpha p (1-2\tilde{s}) - 1}{\alpha p}
\]

\[
= \left\{ \begin{array}{ll}
\frac{s(1-s)(\alpha p s(1-2s)-1)}{\alpha p s(1-s)(2s-1)} & \text{if pos root} \\
\frac{s(1-s)(\alpha p s(1-2s)-1)}{\alpha p s(1-s)(1-2s)} & \text{if neg root}
\end{array} \right.
\]

\[
\frac{\partial \omega}{\partial p} = \alpha + \frac{\partial \delta \partial s}{\partial s \partial p} = \alpha + \frac{1-s}{s} \left( \frac{1}{1-s} + \frac{s}{(1-s)^2} \right) \frac{\partial s}{\partial p}
\]

\[
= \alpha + \frac{1}{s(1-s)} \left( \pm \frac{\alpha p^2 s(1-s)(2s-1)}{2 \sqrt{1 - \frac{\tilde{s} - \alpha p s(1-s)}{\alpha p}} \right)
\]

\[
= \left\{ \begin{array}{ll}
\alpha + \frac{s}{\alpha p^2 s(1-s)(2s-1)} & \text{if pos root} \\
\alpha + \frac{s}{\alpha p^2 s(1-s)(1-2s)} & \text{if neg root}
\end{array} \right.
\]

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 \(x, \bar{x}\). For tractability, split this piece of the likelihood into two components, depending on whether draws of \((\xi, \omega)\) imply an interior or boundary solution.

**Interior optimum (unobserved)**

Integrating over \((\xi, \omega)\) space, the likelihood of an unobserved interior optimum corresponds to:

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

Unfortunately, the domain is not closed-form, and sampling from the full distribution of \((\omega, \xi)\) might require a large number of simulations to produce draws within the bounds. Instead, consider integration over mean utility \((\tilde{\delta}, \tilde{\delta})\) 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 $\delta(p') \leq \ln \frac{1}{M-1}$.

Using the change-of-variables:

$$P\left\{ \tilde{s} \leq \frac{1}{M}, \delta = \delta' \right\} = \int_{-\infty}^{\ln \frac{1}{M-1}} \int_{-\infty}^{\infty} f\left( \tilde{x}(\tilde{\delta}, \delta), \omega(\tilde{\delta}, \delta) \right) \left| \frac{\partial \xi}{\partial \delta} \frac{\partial \xi}{\partial \omega} \right| d\delta d\tilde{\delta}$$

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

$$\tilde{s}_s(\tilde{\delta}) = \frac{\exp(\tilde{\delta})}{1 + \exp(\tilde{\delta})}$$
$$p_s' = \frac{\tilde{s}_s}{\tilde{s}_s (1 - \tilde{s}_s) + \alpha s_s (1 - s_s)}$$

This implies values of the unobservable:

$$\xi_s = \tilde{x}_s - \tilde{x}_i \beta + \tilde{\alpha} p_s'$$
$$\omega_s = \delta_s - \tilde{x}_i \beta + \alpha p_s'$$

I draw $\tilde{\delta}_s$ from a normal distribution with mean $\ln \frac{1}{M-1}$, variance $\tilde{\sigma}_\xi^2$, truncated above at $\ln \frac{1}{M-1}$. I draw $\delta_s \sim N(x_\beta, 10\tilde{\sigma}_\omega^2)$. $\tilde{\sigma}_\omega^2$ and $\tilde{\sigma}_\xi^2$ 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:

$$\hat{P}\left\{ \tilde{s} \leq \frac{1}{M}, \delta = \delta' \right\} = \frac{1}{S} \sum_{s=1}^{S} \frac{F(\xi_s, \omega_s) \left| \frac{\partial \xi}{\partial \delta} \frac{\partial \xi}{\partial \omega} \right|}{\phi\left( \frac{\tilde{\delta}_s - \tilde{\alpha} \tilde{x}_i \beta}{\sqrt{10\tilde{\sigma}_\omega^2}} \right) \phi\left( \frac{\tilde{\delta}_s + \ln(M-1)}{\sqrt{10\tilde{\sigma}_\xi^2}} \right)}$$

Constrained optimum (unobserved)

It is also difficult to sample $(\xi, \omega)$ 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 $(p')$, 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:

78
\[ p_s^I = \frac{\bar{s}(\bar{\delta}_s)}{\bar{\alpha} \bar{s}(\bar{\delta}_s)(1 - \bar{s}(\bar{\delta}_s)) + \alpha \bar{s}(\bar{\delta}_s)(1 - s(\delta_s))} \]

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

\[
\begin{align*}
\xi_s &= \bar{s} - \bar{x} - \bar{\alpha} p^I \\
\omega_s &= \delta_s - x\beta + \alpha p^I
\end{align*}
\]

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:

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

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

\[
\mathbb{P} \left\{ \bar{s} \leq \frac{1}{M}, \delta = \delta^B \right\} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(\xi, \omega) \left| \frac{\partial \xi}{\partial \bar{s}} \frac{\partial \xi}{\partial p} \frac{\partial \omega}{\partial \bar{s}} \frac{\partial \omega}{\partial p} \right| d\delta d\bar{s}
\]

\[
\mathbb{P} \left\{ \bar{s} \leq \frac{1}{M}, \delta = \delta^B \right\} = \frac{1}{S} \sum_{s=1}^{S} \frac{F(\xi_s, \omega_s) \left| \frac{\partial \xi}{\partial \bar{s}} \frac{\partial \xi}{\partial p} \frac{\partial \omega}{\partial \bar{s}} \frac{\partial \omega}{\partial p} \right|}{\phi \left( \frac{\bar{s} - x\beta}{\sqrt{10\bar{s}\omega_x}} \right) \phi \left( \frac{\bar{s} + \ln(\frac{T}{M}) - 1}{\sqrt{10\bar{s}\sigma_\omega}} \right)} \cdot \left\{ \bar{s}(p^B) \leq \frac{1}{M} \right\}
\]
FIGURE 2.1: DISTRIBUTION OF CAMPAIGN PURCHASES ACROSS DAYPART 2012

Notes: Figure 1 shows the distribution of campaign purchases across time slots.
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 2.3: Price Indices Leading Up to Election Day 2012

Notes: Figure 2 shows the Laspeyres (fixed weight) price index over the election cycle for presidential campaigns. The index is calculated relative to the ad spots purchased in the second week of August, 2012. Products are DMA-daypart combinations (80 in total).
Notes: Figure 2 shows the average % decline in prices the week before/after LUR regulation went into effect, separately by daypart. The sample includes all ad spots purchased by campaigns August 27, 2012-September 2, 2012 or September 3, 2012-September 9, 2012.
Figure 2.5: Quantity Withholding across DMAs

(A) Distribution of Quantity Withholding by DMA

(B) Quantity Withholding versus Incidental Viewers

Notes: Figure
## Table 1
Summary Statistics on Ad Purchases by Campaign

<table>
<thead>
<tr>
<th></th>
<th>Obama for America</th>
<th>Romney for President</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Price Per 1,000</strong></td>
<td>5.12</td>
<td>7.37</td>
<td><strong>-2.25</strong>*</td>
</tr>
<tr>
<td><strong>Viewers</strong></td>
<td>(8.14)</td>
<td>(11.90)</td>
<td><strong>-32.89</strong></td>
</tr>
<tr>
<td><strong>Hour Aired</strong></td>
<td>13.73</td>
<td>14.35</td>
<td><strong>-0.62</strong>*</td>
</tr>
<tr>
<td></td>
<td>(6.33)</td>
<td>(6.31)</td>
<td><strong>-13.43</strong></td>
</tr>
<tr>
<td><strong>Number of Viewers</strong></td>
<td><strong>169,816</strong></td>
<td><strong>181,253</strong></td>
<td><strong>-11382</strong>*</td>
</tr>
<tr>
<td></td>
<td>(128,376)</td>
<td>(129,208)</td>
<td><strong>-12.2</strong></td>
</tr>
<tr>
<td><strong>Ratio of uncontested to contested viewers</strong></td>
<td>0.23</td>
<td>0.21</td>
<td><strong>0.024</strong>*</td>
</tr>
<tr>
<td></td>
<td>(0.77)</td>
<td>(0.65)</td>
<td><strong>4.61</strong></td>
</tr>
<tr>
<td><strong>Demographics</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>% White</td>
<td>78.92</td>
<td>80.39</td>
<td><strong>-1.46</strong>*</td>
</tr>
<tr>
<td></td>
<td>(0.04)</td>
<td>(0.05)</td>
<td><strong>-20.4</strong></td>
</tr>
<tr>
<td>% Black</td>
<td>21.18</td>
<td>19.34</td>
<td><strong>1.84</strong>*</td>
</tr>
<tr>
<td></td>
<td>(0.07)</td>
<td>(0.07)</td>
<td><strong>16.6</strong></td>
</tr>
<tr>
<td>% Female</td>
<td>53.55</td>
<td>54.38</td>
<td><strong>-0.82</strong>*</td>
</tr>
<tr>
<td></td>
<td>(0.03)</td>
<td>(0.05)</td>
<td><strong>-14.07</strong></td>
</tr>
<tr>
<td>% Age 65+</td>
<td>17.51</td>
<td>18.36</td>
<td><strong>-0.84</strong>*</td>
</tr>
<tr>
<td></td>
<td>(0.02)</td>
<td>(0.03)</td>
<td><strong>-20.13</strong></td>
</tr>
<tr>
<td><strong>Observations</strong></td>
<td>63625</td>
<td>27123</td>
<td></td>
</tr>
</tbody>
</table>

Notes: Sample of ads purchased from August 1 - November 6, 2012. Standard errors in parentheses, t-statistics below difference in means.
<table>
<thead>
<tr>
<th>Variable</th>
<th>Within</th>
<th>Across</th>
</tr>
</thead>
<tbody>
<tr>
<td>Price Per 1,000 Viewers</td>
<td>0.70</td>
<td>1.19</td>
</tr>
<tr>
<td></td>
<td>0.40</td>
<td></td>
</tr>
<tr>
<td>Hour Aired</td>
<td>0.12</td>
<td>0.46</td>
</tr>
<tr>
<td></td>
<td>0.13</td>
<td></td>
</tr>
<tr>
<td>Number of Viewers</td>
<td>0.39</td>
<td>0.72</td>
</tr>
<tr>
<td></td>
<td>0.23</td>
<td></td>
</tr>
<tr>
<td>Ratio of uncontested to</td>
<td>0.00</td>
<td>2.40</td>
</tr>
<tr>
<td>contested viewers</td>
<td>0.00</td>
<td></td>
</tr>
</tbody>
</table>

Notes: Each ad product is a daypart-week-station combination. Dayparts combine programs within a set block of hours. The coefficient of variation is the standard deviation divided by the mean. Standard deviations reported below means.
Table 3
Candidate Prices Before and After LUR Come into Effect

<table>
<thead>
<tr>
<th></th>
<th>Log Price</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
</tr>
<tr>
<td></td>
<td>(2)</td>
</tr>
<tr>
<td>LUR</td>
<td>-0.45***</td>
</tr>
<tr>
<td></td>
<td>(0.014)</td>
</tr>
<tr>
<td>Days to Election</td>
<td>0.060***</td>
</tr>
<tr>
<td></td>
<td>(0.0015)</td>
</tr>
<tr>
<td>Days to Election^2</td>
<td>0.0090***</td>
</tr>
<tr>
<td></td>
<td>(0.0010)</td>
</tr>
<tr>
<td>Fixed Effects</td>
<td></td>
</tr>
<tr>
<td>Station</td>
<td>X</td>
</tr>
<tr>
<td>Show</td>
<td>X</td>
</tr>
<tr>
<td># Observations</td>
<td>90748</td>
</tr>
</tbody>
</table>

Notes: Heteroskedastic-robust standard errors in parentheses. Coefficients are significant at the *10%, **5%, ***1% level. LUR comes into effect within 60 days of general election (November 6, 2012).
Table 4
Bayesian Parameter Estimates for Commercial and Campaign Demand

<table>
<thead>
<tr>
<th></th>
<th>Commercial Parameters</th>
<th></th>
<th>Campaign Parameters</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Initial Guess</td>
<td>Posterior Mean</td>
<td>Credible Interval</td>
<td>Initial Guess</td>
</tr>
<tr>
<td>Pivotality × Fraction White</td>
<td>-0.58</td>
<td>1.42</td>
<td>1.24 1.63</td>
<td>-0.84</td>
</tr>
<tr>
<td>Pivotality × Fraction Black</td>
<td>-0.34</td>
<td>3.36</td>
<td>2.02 4.29</td>
<td>-0.71</td>
</tr>
<tr>
<td>Pivotality × Fraction Old</td>
<td>0.34</td>
<td>-0.25</td>
<td>-0.52 0.07</td>
<td>2.04</td>
</tr>
<tr>
<td>Pivotality × Fraction Female</td>
<td>-0.23</td>
<td>-0.91</td>
<td>-1.32 -0.73</td>
<td>0.99</td>
</tr>
<tr>
<td>Fraction White</td>
<td>0.21</td>
<td>-0.44</td>
<td>-0.99 -0.25</td>
<td></td>
</tr>
<tr>
<td>Fraction Black</td>
<td>0.10</td>
<td>-3.42</td>
<td>-3.87 -2.00</td>
<td></td>
</tr>
<tr>
<td>Fraction Old</td>
<td>-0.27</td>
<td>0.11</td>
<td>-0.30 1.33</td>
<td></td>
</tr>
<tr>
<td>Fraction Female</td>
<td>-0.03</td>
<td>1.57</td>
<td>0.88 1.95</td>
<td></td>
</tr>
<tr>
<td>Pivotality</td>
<td>0.10</td>
<td>0.28</td>
<td>0.17 0.77</td>
<td>0.00</td>
</tr>
<tr>
<td>Ratio of Uncontested to Contested Viewers</td>
<td>0.59</td>
<td>-0.19</td>
<td>-0.37 0.12</td>
<td></td>
</tr>
<tr>
<td>Ratio of Viewers in Contested Senate Races to Contests</td>
<td>-0.01</td>
<td>0.25</td>
<td>0.06 0.61</td>
<td></td>
</tr>
<tr>
<td>Week</td>
<td>-0.02</td>
<td>-0.17</td>
<td>-0.18 -0.16</td>
<td>-0.12</td>
</tr>
<tr>
<td>High Priority</td>
<td>-0.03</td>
<td>1.68</td>
<td>0.79 2.15</td>
<td>0.89</td>
</tr>
</tbody>
</table>

Error Covariance

<table>
<thead>
<tr>
<th></th>
<th>Variance (σ²u)</th>
<th>Correlation (σu/σu)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.22</td>
<td>-1.00</td>
</tr>
<tr>
<td></td>
<td>6.40</td>
<td>0.77</td>
</tr>
<tr>
<td></td>
<td>5.44</td>
<td>0.75</td>
</tr>
<tr>
<td></td>
<td>10.85</td>
<td>0.82</td>
</tr>
<tr>
<td></td>
<td>44.69</td>
<td>3.17</td>
</tr>
<tr>
<td></td>
<td>2.87</td>
<td>3.78</td>
</tr>
</tbody>
</table>

Notes: Table 2 shows estimates for the parameters of the demand model outlined in section 5. 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.
Chapter 3

Estimating the Effect of Potential Entry on Market Outcomes Using a Licensure Threshold

with Gaston Illanes

3.1 Introduction

Understanding how firm entry affects competition is a central question in the Industrial Organization literature, and an important input for antitrust policy. The theoretical literature suggests that both the threat of entry and entry itself discipline firm behavior. This paper exploits a natural experiment in firm licensure from Washington state’s 2012 privatization of liquor sales to identify causal effects of these entry forces. The key ingredient to our estimates is exogenous variation in the number of eligible licensees in local liquor markets, generated by a licensure threshold requirement on store size. Although privatized liquor markets in Washington state average fewer than five stores, we find that widening the pool of potential entrants has a small effect on pricing, but a significant effect on product offerings.

Until June 2012, Washington state held a local monopoly over all spirit sales, administered through the Washington State Liquor Control Board (WSLCB). The WSLCB oversaw approximately 320 liquor outlets, each with standardized inventory and uniform prices. This regime is similar to the infrastructure in other Alcohol Beverage Control (ABC) states. In November 2011, voters approved a ballot initiative (I-1183) to privatize liquor sales. Within a year, the WSLCB sold its inventory and the rights to take over existing liquor outlets at

1Beer and wine less than 24% ABV were excluded.
auction. Apart from former state liquor outlets, establishments with at least 10,000 square feet of retail space were allowed to sell liquor. We exploit this threshold rule to estimate the impact of potential entry on market outcomes, in the spirit of a regression discontinuity design. Comparisons of markets with existing supermarkets just above and below the 10,000 square foot cutoff allow us to recover the effect of potential entry on prices and product variety.

We find that grocery stores just above the threshold are 30% more likely to sell liquor than those just below. However, large supermarkets (12,000+ square feet) are less likely to enter in markets with a grocery store just above versus just below the cutoff. On net, these forces combine so that we cannot distinguish a positive effect of potential entry on realized entry in liquor markets. In terms of conduct, we find that shifting a store above the threshold leads to a 3% decrease in transacted liquor prices. This effect is driven by differences in the product mix across markets, as within-product price comparisons show no effect on prices. That is, markets with an additional grocery store above the 10,000 square foot cutoff exhibit a shift towards cheaper products, rather than lower prices for a fixed set of goods. Since eligible stores need not obtain licensure, we interpret these results as the net effect of potential entry. This effect combines two mechanisms: the effect of entry on market outcomes and the effect of deterrence on outcomes.

This paper complements the existing empirical literature on entry, which chiefly adopts a structural approach to tackle endogeneity concerns. Structural models allow the authors to back-out market and firm primitives from observed equilibrium outcomes. These primitives are then used to simulate the effect of entry on market outcomes. In a seminal paper, Bresnahan & Reiss develop a structural model to study the effects of entry on competitive conduct. Focusing on a cross-section of geographically segregated markets, they provide evidence of sharply diminishing effects of entry on price levels. In a similar vein, Berry & Waldfoleg (1999) conclude that free entry leads to an excessive number of entrants in radio broadcasting. They find that marginal firms provide little variety in music genres, but incur large operating costs. On the other hand, Syverson (2004) finds that average production efficiency is higher in markets with more competitors. Berry (1992) argues that heterogeneity across entrants can explain the relationship between profitability and number of firms. Other papers that focus on heterogeneity across entrants include Ciliberto & Tamer (2009) and Jia (2008).

A recent empirical literature on entry deterrence adopts a less structured approach. As an example, Ellison & Ellison (2011) find evidence of strategic investment by testing predictions from a model of the pharmaceutical industry. Their test of entry deterrence boils down to a test of non-monotonicity in strategic investment as a function of market size. Goolsbee &

---

3 http://liq.wa.gov/transition/retailers
Syverson (2008) study the incumbent price response to potential entry by Southwest Airlines. They construct an event study of rivals’ responses to Southwest’s incorporation of new cities into their flight network. They find that rivals lower prices substantially when Southwest expands into the two airports that define the incumbent’s route, even when Southwest has not announced any intention of serving that route.

Our results are significantly smaller than those found by Ellison & Ellison (2011) and Goolsbee & Syverson (2008). This discrepancy hints at the importance of barriers to entry in the airline and pharmaceutical industries compared to liquor markets. As an example, Goolsbee & Syverson (2008) point out that the airline industry is fraught with dynamic demand considerations (such as frequent flyer programs), which are absent from liquor markets and might make entry deterrence differentially profitable. However, our results are consistent with earlier work. In particular, if we ignore the entry deterrence mechanism, we can construct an 2SLS estimate for the effect of entry on prices. Our estimates suggest that a mid-sized entrant leads to a 10% decline in transacted prices. As before, this effect is driven by differences in the product mix, as we find no effect on prices in within-product comparisons. These results are consistent with Bresnahan & Reiss’s finding that the main impact of entry on pricing comes from moving from monopoly to duopoly, as the markets we consider average 4.3 firms. More work on entry is needed to understand where and when entry and strategic investment loom large.

This setting also offers an opportunity to investigate the operating goals of state liquor control boards. Seim & Waldfogel (2013) and Miravete et al. (2014) suggest the Pennsylvania Liquor Control Board (PSCLB) expressly tries to reduce alcohol consumption through outlet location decisions and by setting markups above the profit-maximizing level. Their conclusions are based on demand estimates coupled with structural models of profit-maximizing monopoly behavior. Like the PSCLB, the WSLCB chose store locations and set uniform markups, so that Washington’s deregulation provides an event study we can use to benchmark their estimates. As an example, their results suggest that moving from the standard ABC uniform markup pricing rule to a monopolist setting product-specific markups leads to high price increases for rum and gin compared to vodka, whiskey and tequila. The outcome of liberalization in Washington state was, in fact, the opposite: the price of tequila rose most dramatically. However, our findings are consistent with their determination that the returns to third-degree price discrimination (tailoring markups to local demand conditions) are small compared to second-degree discrimination (tailoring markups to products). They suggest that administrative costs might make such complex pricing strategies unprofitable, and we find that the average coefficient of variation across products is slim. In other words, there is little within-product variation in prices across markets. This

\footnote{Which Washington state also followed, albeit with a higher markup than Pennsylvania (51.9\% vs 30\%).}
result foreshadows our results on the returns to entry.

The remainder of the paper is structured as follows: Section 2 discusses the data used in this study, Section 3 provides an overview of the liberalization policy and its effects, Section 4 describes our empirical methodology and presents our main results, and Section 5 concludes.

3.2 Descriptive Evidence on Deregulation

3.2.1 Background on Liberalization

Washington privatized liquor sales on June 1, 2012. It is the first (and so far, only) control state to deregulate since the end of Prohibition.\(^5\) Costco spent over $20 million promoting the reform, which passed with 57% of the vote in a November 2011 referendum.\(^6\) The reform was marketed as a win-win. Consumers expected lower prices and greater product variety, while the state levied new taxes to compensate for their forgone profits from selling liquor themselves: a 17% tax on spirit retailers and a 10% tax on spirit distributors. On net, the initiative was touted as a means to increase state revenue.

Before liberalization, the WSLCB operated 166 stores (called State Liquor Stores, or SLS) and licensed an additional 162 contract stores (Contract Liquor Stores, or CLS). Contract stores were run by private citizens, but their actions were tightly circumscribed by the WSLCB. In particular, all stores maintained the same prices and product selection. The state acted as distributor and retailer, and charged a uniform markup of 51.9% on all liquor products. In addition, it charged a 20.5% alcohol sales tax and a $3.7708 per liter tax.

Several reform attempts preceded I-1183, most notably I-1180 in 2010. Also funded by Costco, it was defeated at the polls, 53% to 47%. There are two key differences between I-1183 and the unsuccessful I-1180. First, initiative 1183 added a 10% tax on distributors and a 17% tax on retailers, aimed at bolstering government revenue.\(^7\) Nonetheless, proponents of reform argued “some prices are expected to drop, though not as low as in California, because Washington will keep its high liquor taxes.”\(^8\) These proponents hoped that competition would drive markups down enough to compensate for the tax hikes. A second difference was the I-1183 size restriction for licensure, as I-1180 had no such restriction. One central argument against I-1180 was the fear it would allow convenience stores to sell liquor, increasing the availability of cheap products and “spark(ing) an increase in alcohol-related crime and

\(^{7}\)See:
http://liq.wa.gov/stores/liquor-pricing
underage drinking". As a response, the 10,000 square foot requirement, aimed at excluding small retailers, was introduced in 2011. The next section presents an identification strategy based on this discontinuity.

3.2.2 Liberalization and Prices

First, we investigate whether privatization led to higher prices for consumers. This descriptive evidence not only helps us understand what happened in Washington at regulation, but also contributes to the literature on the merits of state-owned enterprises versus deregulation. We propose several indices to measure changes in the overall price level, and then explore whether demographics help explain variation in these changes across the state and across product categories.

From a theoretical perspective, it is unclear whether liberalization leads to higher or lower prices, and the empirical evidence is mixed. As an example, in a case study of deregulation in Mexico, La Porta & López-De-Silanes (1999) document a 5% increase in prices. In our context, there are several forces whose combined effect on prices is ex ante ambiguous. First, if private firms hold market power, then deregulation might lead to price increases. If competition is strong, however, we would expect prices to fall. Indeed, proponents of reform in Washington argued the private firms would charge markups far below the WSLCB’s 51.9% level. Second, the new taxes implemented at deregulation ought to increase prices. Finally, as the state monopolist, the WSLCB contracted directly with distillers (rather than purchasing from distributors), and might have paid lower acquisition prices than retailers in the new private system. Indeed, local papers are rife with small retailer complaints that they lose out to monopsonistic firms like Costco.

We utilize price indices to compare prices before and after the reform. One challenge in price comparisons is that the state and the private market offer different product selections. The WSLCB data contain prices for all products, but the scanner data contain prices only for transacted goods. If consumers substitute away from expensive products, then the missing prices are not randomly selected. A naive comparison would therefore understate increases in offered prices. To deal with these issues, we follow the discussion in Chevalier & Kashyap (2014) and employ the Törnqvist price index (Törnqvist (1936)) to measure changes in price levels.

The Törnqvist index formula for a comparison of prices between $t$ and $t - 1$ is

\[ P_{t} = \frac{\sum_{i} p_{i t} q_{i t}}{\sum_{i} p_{i t-1} q_{i t-1}} \]


10 Note that while the text of the law allows for exceptions to this rule in “under-served areas”, with the definition of this concept left to the judgment of the WSLCB, as of 2015 no store with less than 10,000 square feet has received a liquor license.
\[ I_t = \prod_{j=1}^{N} \left( \frac{p_{j,t}}{p_{j,t-1}} \right) \left( \frac{\sum_{i=1}^{N} p_{i,t} q_{i,t} + \sum_{i=1}^{N} p_{i,t-1} q_{i,t-1}}{\sum_{i=1}^{N} p_{i,t} q_{i,t} + \sum_{i=1}^{N} p_{i,t-1} q_{i,t-1}} \right) \].

This index is a weighted average of the relative prices across products, where the weights are the average expenditure shares across the two periods. If consumers substitute away from goods that have high price increases, those products still receive substantial weight in the index if they were frequently purchased in the pre-period. Caves et al. (1982) show that this index approximates the ideal price index for a representative consumer with homothetic utility.

Figure 3.1 shows the monthly Törnqvist for all products from October 2010 to December 2012, and it is easy to see the dramatic level shift in prices at privatization, as the first week of June 2012 saw a 13.2% price increase relative to May. To be clear, the figure shows the month-to-month change, so this spike reads as an immediate increase that is sustained through the end of our dataset.

We observe 1,220 products sold in the last month before liberalization, and 721 products sold in the first week after (the frequency of our data changes from monthly to weekly at liberalization). As a result, the previous price change includes the effect of the dramatic drop in product variety, as products that are transacted in period \( t - 1 \) but are not transacted in period \( t \) will be included in the Törnqvist calculation. One might be interested in a calculation taking only into account products that are sold both before and after liberalization. Figure 3.2 repeats the previous exercise for products that are sold every week during our sample period. For these 354 products, which we call the “State-balanced Panel”, prices increase by 19.6%.

Second, we examine whether this average price increase masks heterogeneity either across the state or across products. Washington’s deregulation provides an event-study counterpoint to Miravete et al. (2014), who find that state uniform markup rules subsidize poorer clientele. They also find that wealthy consumers have less elastic demand for liquor, so that a monopolist ought to charge higher markups for products that are more demanded by wealthier individuals. As an example, cheaper goods have a higher consumption share among the poor in Pennsylvania, and Miravete et al. (2014) predict a large increase in the markup of these products if the PSLCB behaved as a monopolist. We find a similar trend for liquor consumption under the WSLCB. Following their lead, we categorize products as “cheap” (“expensive”) if the product is priced below (above) median for the its category. Panel d in Figure 3.6 shows that the share of expensive products increases with income. However, we do not find that liberalization leads to to disproportionate price increases in poorer areas or for cheaper products. Figure 3.3 presents a scatter plot of median income.
of the zip code and the Törnqvist price index at liberalization, while Figure 3.4 presents the same plot for the Törnqvist price index from liberalization to the end of 2012. While there is significant heterogeneity across the state, these indices appear uncorrelated with median income. Figures 3.12 and 3.13 in the Appendix repeat this exercise using zip code population, and again find no relationship.

We find some heterogeneity in price changes across liquor categories, but they do not correspond to the predictions in Miravete et al. (2014)'s monopoly model. These authors simulate the transition from a uniform markup rule to a product-specific markup for Pennsylvania, and predict higher price increases for rum and gin, and lower price increases for vodka, whiskey and tequila. In Washington, we observe the change from uniform markups to liberalization, where markups are tied to demand elasticities for the first time. Figure 3.5 reports the observed price changes in Washington state, by liquor category, for the state balanced panel. Figure 3.14 in the Appendix repeats this exercise for the unbalanced panel. We find that tequila experiences the highest price increase, while the remaining categories have roughly the same change.

We also document significantly less heterogeneity in markups across products than that predicted by Miravete et al. (2014). We infer product-level marginal cost using pre-liberalization prices net of the WSLCB's 51.9% markup. These prices are the WSLCB's acquisition costs. While the WSLCB, as a state monopsonist, ostensibly had access to low wholesale prices, the chain supermarkets in our sample are also likely to wield substantial bargaining power in upstream markets. The WSLCB's costs therefore ought to be a useful proxy for the acquisition cost faced by this set of retailers. However, we adjust the WSLCB wholesale prices to account for the new 10% distributor tax at liberalization, which may be passed through to retailers. We bound post-liberalization markups under the no- and perfect-pass through cases, and document how these bounds vary across product categories.

Table 3.1 reports percentage markups for the period between June 2012 and December 2012, and shows that markups are remarkably consistent across product categories. For example, the average markup for Whiskey is between 38% and 44%, while the average markup for Rum is between 37% and 43%. As before, there is no significant difference in markups for products classified as "cheap" or "expensive". These figures are in the ballpark of what Miravete et al. (2014) simulate a profit maximizing monopolist would charge in Pennsylvania (reported in the last columns of Table 3.1), with two key exceptions: Tequila should be priced more competitively (a mere 27% markup); and cheaper products should have higher markups (67% vs. 26%, with an average of 42%).

The discrepancies between the predictions in Miravete et al. (2014) and the facts we document in Washington state hint at the differences between monopoly and competitive second degree price discrimination. Differences in demand between Washington and Penn-


slyvania may also contribute to these disparities. As an example, Figure 3.6 reports market shares at the zip code level for the period between January 2012 and May 2012, broken down by observable characteristics. The first sub-figure reports shares by product category as the percentage who are a minority increases. Unlike the patterns found by Miravete et al. (2014) in Pennsylvania, there does not seem to be a significant gradient for any category across this dimension. The second sub-figure shows market shares by product category as the percentage college educated increases. Here we observe that whisky and rum consumption are negatively correlated with college education, while vodka and gin consumption are positively correlated with college education. These patterns are consistent with those found for Pennsylvania, albeit with shallower slopes in all cases. The third sub-figure repeats the analysis as the percentage of individuals who earn more than $50,000 a year (“High-income”) increases. Here we also find an increasing pattern of consumption of vodka and gin, and a weakly negative association for rum and whisky. However, we note that even when the underlying consumption patterns between the two states are similar, market outcomes diverge from their simulations.

3.2.3 Liberalization and Number of Stores

While deregulation brought higher prices for consumers, it has also meant increased availability of spirits. Indeed, opponents of reform feared an overabundance of liquor outlets, leading to hikes in underage consumption and driving accidents. Figure 3.7 shows that these fears have been partially realized: the number of liquor outlets state-wide increased from approximately 360 to over 1,400 stores within the first six months of privatization. Chamberlain (2014) documents a rise in neighborhood crime associated with liquor availability following deregulation.

Figure 3.8 is a scatterplot of liquor entry versus the number of supermarkets at the zip code level. There is a strong, positive relationship between the number of stores in the TDLinx data and the number of licensees recorded by the WSLCB. On average, there are 2.6 liquor outlets in each zip code, and 1.3 of these are grocers, superettes, or convenience stores. Under state control, zip codes averaged a mere 0.6 liquor stores. the dramatic increase in liquor outlets following deregulation hint at a central finding in Seim & Waldfogel (2013): the state monopolist restricted stores to curb alcohol consumption.
3.3 Data

3.3.1 Pre-liberalization: Price and Quantity Data from the WSLCB.

We obtained pre-liberalization data from the WSLCB’s public records. The WSLCB published monthly price lists, and each list contains the retail price, liquor taxes, liquor type, size, brand name and proof for every product offered in that month. We use data from these price lists from October 2010 to May 2012, for a total of 45,948 product-month observations. During this period 1,916 products were sold, and the average after-tax price was $21.70. Although the state sold malt beverages and wine, we focus on sales of hard liquor for tractability. Quantity sold is reported at the establishment level on a monthly basis, both for State Liquor Stores (SLS) and Contract Liquor Stores (CLS).\footnote{We drop instances of negative sales based on conversations with Melissa Norton at the WSLCB. These seem to be inventory adjustments triggered by state audits, and happens for 0.5\% of the observations.}

3.3.2 Post-liberalization: Grocery and Convenience Store Sizes and Licensure

Data on grocery and convenience store square footage comes from TDLinx, a subsidiary of Nielsen. For each establishment in Washington state, TDLinx provides store name, addresses, and square footage in January 2011 and December 2012. This data is crucial for constructing the set of eligible licensees based on the WSLCB’s threshold rule.

Figure 3.9 is a histogram of stores sizes near the threshold in 2011 and 2012. Importantly, the number of stores with reported square footage just above 10,000 square feet does not change across these periods. Since I-1183 passed in November 2011, stores had no incentive to manipulate square footage in January 2011. The stability of store sizes allays concerns that supermarkets might expand in order to gain licensure.

We match data on store sizes to data on liquor licensure from January 2013 using store name and addresses. The WSLCB maintains a list of off-premises licensees on their website. Historical licensure records are taken from the WayBackMachine. The licensure files contain information on licensee addresses, trade names, license type (beer, wine or spirit) and licensure date. Our final dataset is an establishment-level database of entry and square footage.

3.3.3 Post-liberalization: Grocery Store Liquor Prices

We collect data on post-reform price and quantities from the Nielsen Retail Scanner dataset, available through the Kilts Center at the University of Chicago’s Booth School
of Business. The Retail Scanner database tracks all transactions at a set of unnamed supermarkets across the United States. The data contain average weekly prices, quantities, sales information, and product descriptions for each anonymous establishment. Products are tracked at the UPC level. We focus on the 678 scanner stores in Washington that record at least one liquor transaction in 2012. The average after-tax price of a liquor product is $25.40 based on the sale of 1,525 unique UPCs in the six months following liberalization.

The Retail Scanner data allows us to investigate the pricing and product variety decisions of large supermarkets. All of these stores enter the liquor market only after privatization. Unfortunately, while Nielsen records prices at the establishment level, it obfuscates store identities. Only the FIPS county code is reported to researchers. This obfuscation poses a difficulty in measuring our left-hand-side variable. We cannot conduct a store-level analysis by matching establishment sales to square footage of licensees. Instead, we match Nielsen Scanner establishments to zip codes, and average across stores to find zip code-level prices for each product, each week. Our markets are therefore zip codes, and our main results consider the effect of entry on zip code outcomes.

We match stores to zip codes using shopping patterns from the Nielsen Panel dataset. This dataset tracks a panel of consumers, rather than stores. Each included household records all purchases by shopping trip, including store identifiers we match to the Retail Scanner dataset. We infer store locations based on the zip codes of households who shop there most often. We count the number of trips originating in each zip code and culminating in each store, and then assign stores the modal zip code across trips. Importantly, we only count trips before privatization, so that household choice of grocers should be independent of local liquor market competition. Stores with fewer than 10 trips are excluded to reduce noise. Figure 3.10 is a histogram of trips to each store, by zip code rank. Across stores, the modal zip code (most popular) originates 33 trips, while the second most-popular zip code originates a mere 12.5. This sharp decay in trips suggests the panelist data is informative about likely store locations.

Since we are working with transactions data, any product that is offered in a store but that is not sold during a certain week will be missing from our dataset. In fact, 53.8% of the UPCs we observe post liberalization are not sold in any store in the state for at least one week. This would be a potential concern if we were estimating demand, for example, as it would imply that products with a low unobserved preference value are less likely to enter our data. However, since our question of interest regards the effects of market structure on transacted prices, the fact that we are missing prices for goods that are not transacted is irrelevant. More concerning is the fact that we do not observe prices in stores outside the scope of the Nielsen database, and we will miss any goods that are sold by specialty liquor stores and not by supermarkets. We interpret the results that follow as identifying
the responses of grocery stores that sell liquor to market structure, and make no claims about external validity of these estimates to the response of other liquor retailers.

3.4 Empirical Strategy

3.4.1 10,000 Square Foot Licensure Requirement on Entry: Store Level

Our main empirical strategy is to compare entry and prices in markets with stores in the neighborhood of the 10,000 square foot licensure threshold. First, we look for evidence on whether stores just above the cutoff are more likely to acquire a liquor license than stores just below. While I-1183 allowed the WSLCB to make exceptions to the size requirement in underserved areas, in the time period we consider, the board had yet to exploit this loophole.\textsuperscript{12,13} In this analysis, therefore, we treat the threshold as a hard cutoff.

Our identification strategy requires that stores in the neighborhood of the threshold pose a threat of entry. If, for example, entry in liquor markets were blockaded (incumbents need not investment to deter entry), then this natural experiment would be uninformative about the effects of entry deterrence and entry on market outcomes. Our first test is therefore whether stores just above the threshold take advantage of their eligibility to sell liquor. We use the licensure requirement to construct a regression discontinuity design as in Imbens & Lemieux (2008), where $s$ denotes store:

$$L_s = \beta_0 + \beta_1 E_s + \beta_2 E_s \times (SQFT_s - 10) + \beta_3 (1 - E_s) \times (10 - SQFT_s) + \delta' X_s + \epsilon$$

$L_s$ takes a value of 1 if store $s$ is licensed, $SQFT_s$ is square footage, and $E_s$ is an indicator that the store is eligible for a license (has at least 10,000 square feet of space). We estimate (3.2) for three separate bandwidths: 2,000 square feet of the cutoff (73 stores), 5,000 square feet of the cutoff (246 stores), and the entire dataset (3,969 stores). Since there are relatively few stores near the threshold, we allow for linear trends in size only for the full sample. $X_s$ includes zip code level control variables, meant to capture characteristics about store $s$'s competitors. Our full specification includes the number of (licensed and unlicensed) supermarkets in the same zip code within, above, and below the bandwidth. We estimate the coefficients from (3.2) using OLS with heteroskedasticity-robust standard errors.


3.4.2 10,000 Square Foot Licensure Requirement on Entry and Prices: Market Level

The WSLCB square footage requirement is a treatment on individual supermarket eligibility to sell liquor, but also on the number of eligible liquor outlets at the market level. Some markets have more (fewer) potential entrants into liquor sales because the existing supermarkets were just above (below) the threshold. The identification assumption is that markets with the same number of mid-sized stores (stores with square footage within a fixed bandwidth around the cutoff), but different store size distributions within that bandwidth, are otherwise similar. The number of liquor entrants in market $m$ is a sum, across groceries, of individual establishment entry decisions. We aggregate (3.2) across the $S_m$ stores in market $m$ to model liquor entry at the market level:

$$NL_m = \sum_{s=1}^{S_m} L_s + \omega_m$$
$$= \sum_{s=1}^{S_m} (\beta_0 + \beta_1 E_s + \beta_2 E_s \times (SQFT_s - 10) + \beta_3 (1 - E_s) \times (10 - SQFT_s) + \delta X_s + \epsilon_s) + \omega_m$$
$$= \alpha_0 + \beta_0 S_m + \beta_1 S_m \sum_{s=1}^{S_m} E_s + \beta_2 \sum_{s=1}^{S_m} (E_s \times (SQFT_s - 10)) + \beta_3 \sum_{s=1}^{S_m} ((1 - E_s) \times (10 - SQFT_s)) + \delta' X_m + \omega_m$$

(3.3)

$$= \alpha_0 + \beta_0 S_m + \beta_1 NE_m + \beta_2 TSQFTA_m + \beta_3 TSQFTB_m + \delta' X_m + \eta_m$$

(3.4)

where $NL_m$ is the number of liquor outlets in market $m$, $S_m$ is the total number of grocery stores in the bandwidth, $NE_m$ is the number of eligible groceries in the bandwidth, and $TSQFTA_m$ and $TSQFTB_m$ are total square footage above and below the cutoff, respectively. $X_m$ includes market controls, such as the total number of groceries or the number of groceries above- and below the threshold, depending on the specification.

We extend this specification to examine how the licensure requirement affects liquor prices. We adopt (3.4) to a panel structure. The price of product $j$ in market $m$ in week $t$ is modeled as:

$$\log \text{Price}_{jmt} = \alpha_0 + \beta_0 S_m + \beta_1 NE_m + \delta' X_m + \gamma_j + \alpha_t + \nu_{jmt}$$

(3.5)

where $\gamma_j$ are product characteristics, including size or product-level fixed effects, and $\alpha_t$ are month fixed effects. We cluster standard errors at the market level, to account for correlation in pricing across products.
Finally, under additional assumptions, we can extend this strategy to identify the causal effect of entry on pricing. We estimate the following model by 2SLS:

\[
\begin{align*}
\text{LogPrice}_{jmt} &= \gamma_0 + \gamma_1 NL_m + \lambda'X_m + \gamma_j + \alpha_t + \nu_{jmt} \\
NL_m &= \alpha_0 + \beta_0 S_m + \beta_1 NE_m + \delta'X_m + \omega_m
\end{align*}
\] (3.6)

As before, the identifying assumption requires that conditional on the number of stores in the bandwidth, the number of these which are eligible is uncorrelated with \(\nu_{jmt}\). So long as eligibility affects licensure (a robust first stage), then we can exploit the threshold to estimate the causal effect of entry on prices.

However, we are skeptical of this identification argument for the two-stage least square estimates. Eligibility may affect prices not only through entry, but also through the threat of entry. For example, with demand or cost uncertainty, firms may attempt to deter entry by signaling market unprofitability (Milgrom & Roberts (1982)). In a limit-pricing model, potential entry (the number of eligible stores) might affect prices directly, and the number of eligible firms would constitute an omitted variable in the second-stage pricing equation (3.6). Only if potential entry affects prices solely through realized entry will the two-stage least squares estimates of (3.6) be the causal effect of entry on prices.

3.5 Results

3.5.1 Liquor Licensure by Square Footage

Figure 3.11a shows the relationship between square footage and the probability of licensure for stores between 4,500 and 19,499 square feet. The probability of licensure jumps approximately 30% between 9,000 and 10,000 square feet. This estimate likely understates the impact of eligibility on entry for two reasons. First, the TDLinx data bins store square footage, so that stores characterized as 10,000 square feet range, in fact, from 9,500-10,499 square feet. This means the 10s bin contains stores ineligible for licensure. Second, square footage is clearly measured with error. Since the 10,000 square foot requirement is strict, the probability of entry for smaller stores should be zero. The incidence of licensed small stores in our data points to errors in recording square footage, confirmed by Google Maps estimates of true store sizes. We therefore interpret the discontinuity at 10,000 square feet as a lower bound for the impact of eligibility on licensure.

As a robustness check, we construct the same threshold comparison for beer licensure.
Even before I-1183, the WSLCB licensed grocers to sell beer, without regard for store size. As a result, there should be no jump in the probability of licensure for beer at the 10,000 square foot mark. Figure ?? shows the standardized probability of entry by square footage for beer and liquor separately. If anything, at the threshold, the probability of beer licensure falls. These results suggest that the jump in liquor licensure is not driven by underlying differences in unobserved store characteristics.

Table 3.2 shows the store-level regressions corresponding to Figure 3.11a and model (3.2). The coefficient on the 10,000 square foot indicator is statistically significant and positive across all specifications. Adding controls for the market configuration, such as the number of other stores within the market and their sizes, affect neither coefficient magnitudes nor significance. Widening the bandwidth from 2,000 to 5,000 or to all square footage increases the point estimates from approximately 30% to 40%. This is consistent with larger supermarkets being more likely to sell liquor, even apart from the 10,000 foot threshold.

### 3.5.2 Effect of Licensure on Entry at the Market Level

Table 3.3 reports the estimates of licensure on entry at the market level, which corresponds to the model in equation (3.4). Columns (1) through (5) employ a 2,000 square foot bandwidth around the 10,000 square foot cutoff. The coefficient on the number of firms in the bandwidth, but just above the threshold, is large, positive, and statistically significant. The estimates imply an additional store just above the threshold corresponds to 1/3 more liquor licensees (within the bandwidth) at the market-level. This result is consistent with the store-level regression results in Table 3.2, and is robust to including controls for the composition of other supermarkets in the market.

We also consider the effect of eligibility on the number of small (less than 8,000 square feet) and large (12,000+ square feet) liquor licensees (columns (3) and (4)). Stores below 8,000 square feet are ineligible for licensure, so there should be no effect of entry by mid-sized stores on the number of small licensed stores. Column (4) confirms a null effect; the point estimate is small and statistically insignificant. Column (3) considers the effect of mid-sized store eligibility on large store licensure. If mid-sized stores enter, they may crowd-out larger stores from liquor markets. The point estimate is negative, but statistically insignificant. Reassuringly, the coefficient on the number of large supermarkets is large, positive and statistically significant, implying that almost all large supermarkets (80.8%) choose to sell liquor. Column (5) reports the coefficient of mid-sized store eligibility on the total number of licensed grocers. The coefficient is positive (on the order of 15%) but statistically insignificant. This attenuation (compared to columns (1) and (2)) comes from the crowd-out effect reported in column (3). Indeed, the coefficient estimate in column (5)
is the sum of the estimates in (2)-(4).

Since there are relatively few stores within the 2,000 square foot bandwidth, we estimate the effect of licensure at the market-level using all stores, but controlling for linear trends in square footage. Columns (6) and (7) report the coefficient on the number of groceries above 10,000 square feet in size on the total number of licensed grocers and total number of licensees, respectively. The estimates imply that moving a grocery from just below to just above the 10,000 square foot threshold leads to another .5 liquor-selling grocers, on average. The point estimate in column (7) suggests a similar increase in the total number of liquor outlets, but the standard errors are too large to reject a null effect. As a whole, these results suggest that markets with a grocery just above, rather than just below the threshold have 0.3 more licensed grocers.

3.5.3 Effect of Licensure on Prices

In this section, we present results on the effect of eligibility on liquor prices. In this specification, each observation is a UPC - week - zip code combination, and we cluster standard errors at the zip code level. Column (1) in Table 3.4 reports the baseline regression; an additional eligible supermarket reduces prices approximately 3%. The effect is marginally significant (the p-value is .054). Column (2) reports the coefficient estimates when product (UPC) fixed effects are included. The coefficient on eligibility loses both economic and statistical significance. While the fixed effects specification uses only within-product variation, the baseline specification constructs cross-product comparisons. These results suggest that pricing differences across markets with different numbers of eligible licensees are driven by differences in product offerings, rather than by differences in prices. In particular, supermarkets offer cheaper products in markets with more eligible licensees, rather than cheaper prices for common products.

To test this theory, we divide the sample according to product popularity. Estimates for popular products – those sold most widely across the state – should be invariant to the inclusion of fixed effects. Columns (3) and (4) report effects for the 10% most popular products, with and without fixed effects respectively. The estimates confirm our interpretation of the coefficients reported in columns (1) and (2); there is no discernible effect of eligibility on prices in either specification. In contrast, column (5) shows an effect for less popular products, which is eviscerated by including fixed effects.

As a second test, we consider the effect of eligibility on product variety directly. We calculate the number of products (unique UPC codes) sold in each market during the six months following deregulation. Results, presented in column 5 of table XX, suggest that an additional eligible firm leads to an increase of 60 products available in the market. This
amounts to a 7% increase in product variety. If consumers value variety, then within-product comparisons understate the returns to competition.

Table 3.5 shows the effect of eligibility across product categories. We find the largest effects for vodka and rum, and no evidence that eligibility affects prices for gin. Miravete et al. (2014) suggest that demand elasticities differ substantially across these product categories, and that profit-maximizing markups ought to be highest for products with low-income, low-education clientele. While the consumption patterns across demographics in Pennsylvania do not mirror those in Washington, we find substantial evidence of heterogeneity across product categories. As an example, eligibility has a large effect on vodka prices and a null effect for gin prices, although both of these track with college education.

We also consider the effects of entry on price dispersion and quantity sold. We calculate the standard deviation of prices for each UPC in each week after privatization, and test whether markets with more eligible grocers exhibit higher price dispersion. The results, presented in table 3.6, show no effect of potential entry on this measure of dispersion. Second, we test whether the average sales per store decline with potential entry. If entrants engage chiefly in business-stealing, then we would expect that quantities decline at chain stores in markets with additional potential entrants. However, we find no affect on sales, even including product fixed effects. Since we have already established that prices for fixed products do not decline with potential entry, this suggests that the additional products do not cannibalize sales of staple products.

If the effect of eligibility on prices operates only through the entry channel, then we can construct IV estimates for the effect of entry on prices. This interpretation precludes forces such as entry deterrence. Our estimates imply that an additional entrant leads to a decline in transacted prices on the order of 10%. IV estimates of entry on average price are presented in Table 3.7. Column (1) presents the OLS estimates of log prices on the number of mid-sized groceries selling liquor at the market-level. The coefficient estimate is statistically significant, and suggests that entry leads to higher prices. Of course, a central concern in the OLS estimates is that entry is endogenous, so that markets with more firms have higher demand. This omitted variable might drive the positive correlation between entry and prices. Column (2) presents the baseline IV estimate, using the 2,000 square foot bandwidth. The coefficient on the number of licensed grocers is now large and negative (suggesting an additional entrant leads to a 10% decline in price), but the standard errors are wide.

A null effect of entry on prices for a fixed set of products is consistent with Bresnahan & Reiss (1991)'s non-experimental evidence of entry on tire prices. They cannot reject that markets with three-, four-, and five- firms have the same prices. On average, markets with grocers just below the threshold have 4.15 liquor outlets, so that an additional liquor outlet
constitutes a shift from a four- to five- firm market. However, our results hint that entry affects product variety and therefore average prices. We find that markets with an additional eligible firm offer consumers more and cheaper choices.

3.6 Conclusion

Proponents of liquor deregulation in Washington state confronted a canonical challenge in designing institutions for a new private market: how to harness competition to improve market outcomes. Concern about alcohol-related crime prompted regulators to institute a 10,000 square foot licensure requirement to curtail entry, implicitly selecting the number and type of potential entrants into these fledgling markets across the state. We exploit this threshold rule to estimate causal effects of potential entry on market outcomes in the six months following privatization. Our findings that suggest an additional potential entrant lowers average prices by 3%. This effect represents a shift in the product mix towards cheaper goods; markets with an additional potential entrant have roughly 60 more products transacted. However, we find that prices for a fixed set of goods do not change, and that there are no effects of potential entry on price dispersion or on rivals’ average sales.

These results contrast with the recent empirical literature on potential entry, which find larger effects. As an example, Goolsbee & Syverson (2008) document large price changes in airlines or and Ellison & Ellison (2011) find large effects on advertising decisions in pharmaceuticals. Our results point to the importance of interactions between potential entry and other market features. As an example, liquor retail sales involve smaller fixed costs and have fewer dynamic considerations than these two comparison industries.

Our results also provide policy implications from Washington state’s experience with privatization of liquor retail. First, we find that the 10,000 square foot regulation appears binding (stores just above the cutoff are more likely to enter than those just below), but it does not significantly affect the overall number of liquor outlets within each market. Large supermarkets adjust their entry decisions depending on the eligibility of their mid-sized neighbors. Therefore, it appears the regulation chiefly affected the composition of liquor outlets in the market. While these findings suggest that extending the liquor franchise to smaller supermarkets would not dramatically increase liquor availability, the behavior of very small stores (for example, gas stations) need not conform to this result. Second, we find that markets with an additional potential entrant shift their product mix towards cheaper products. This confirms concerns that competition in liquor markets leads to greater availability of cheap alcohol, and suggests that regulation has an effect in limiting the availability of those types of products.
Exploiting credible exogenous variation to separately identify the effects of realized entry and entry deterrence is an important next step in unpacking the effects of potential entry we explore here. As an example, to test whether firms deter entry using limit pricing, researchers should study markets where (for exogenous reasons) potential entrants are differentially informed about market conditions. In a setting like ours, this might involve a comparison of chain versus independent stores. Researchers might also employ variation similar to ours to test the robustness of structural entry models, by comparing their predictions using pre-liberalization data to actual market outcomes. Such work would complement the growing entry games literature, in a fashion analogous to Peters (2006) and the merger simulation literature.
FIGURE 3.1: TÖRNQVIST PRICE INDEX, UNBALANCED PANEL

FIGURE 3.2: TÖRNQVIST PRICE INDEX, STATE-BALANCED PANEL
Figure 3.3: Törnqvist Price Index Change at Liberalization and Zip 5 Median Income, State-balanced Panel

Figure 3.4: Törnqvist Price Index Change from Liberalization to End of 2012 and Zip 5 Median Income, State-balanced Panel
Figure 3.5: Törnqvist Price Index for Liquor Categories, State-balanced Panel

Figure 3.7: Number of Licensees over Time
FIGURE 3.6: SHARES BY PRODUCT CATEGORY/TYPE AND ZIP CODE DEMOGRAPHICS

(a) % Minority and Shares by Product (b) % College Educated and Shares by Product Category

(c) % High-Income and Shares by Product Category (d) % High-Income and Shares by Product Type
Figure 3.8: Liquor Licensees vs. Supermarkets by Zip Code

Figure 3.9: Number of Stores by Size, Before and After Privatization
Figure 3.10: Zip Code Fuzzy Match Algorithm

![Graph showing average number of trips vs. zip code rank by number of trips.](image)
**Figure 3.11: Liquor Licensure at the 10,000 Square Foot Threshold**

(a) Liquor Licensure v. Store Square Footage

(b) Standardized Liquor & Beer Licensure
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>Mean</td>
<td>Mean</td>
<td>Mean</td>
</tr>
<tr>
<td>Mean</td>
<td>SD</td>
<td>SD</td>
<td>SD</td>
</tr>
<tr>
<td>Full Sample</td>
<td>44.0%</td>
<td>38.4%</td>
<td>42.7%</td>
</tr>
<tr>
<td>Whiskey</td>
<td>43.9%</td>
<td>38.2%</td>
<td>39.9%</td>
</tr>
<tr>
<td>Rum</td>
<td>42.7%</td>
<td>37.0%</td>
<td>59.6%</td>
</tr>
<tr>
<td>Tequila</td>
<td>46.9%</td>
<td>41.6%</td>
<td>21.8%</td>
</tr>
<tr>
<td>Vodka</td>
<td>43.9%</td>
<td>38.3%</td>
<td>40.7%</td>
</tr>
<tr>
<td>Gin</td>
<td>42.7%</td>
<td>37.0%</td>
<td>50.6%</td>
</tr>
<tr>
<td>Expensive</td>
<td>44.7%</td>
<td>39.1%</td>
<td>26.3%</td>
</tr>
<tr>
<td>Cheap</td>
<td>42.7%</td>
<td>37.0%</td>
<td>67.4%</td>
</tr>
</tbody>
</table>

Notes: This table reports summary statistics for estimated post-liberalization percentage markups between June 2012 and December 2012. All statistics are unweighted and calculated using data at the week-store-upc level. Products are categorized as “Expensive” (“Cheap”) if their WSLCB price exceeds the median price charged by the WSLCB for that category. All results infer marginal cost from pre-liberalization data. Columns labelled “No Pass-Through” assume that the 10% distributor tax is not passed through to retailers, while columns labelled “Perfect Pass-Through” assume that the entirety of the tax is passed through to retailers. Columns under the header “MST (2015)” report the predicted markups from Table 11 in Miravete, Seim and Thurk (2015).
### Table 3.2: Effect of Store Size on Liquor Licensure

<table>
<thead>
<tr>
<th>Sample</th>
<th>8,000-12,000 Square Feet</th>
<th>5,000-14,000 Square Feet</th>
<th>Full</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
</tr>
<tr>
<td>SQFT ≥ 10</td>
<td>0.283***</td>
<td>0.286***</td>
<td>0.268**</td>
</tr>
<tr>
<td></td>
<td>(0.107)</td>
<td>(0.106)</td>
<td>(0.109)</td>
</tr>
<tr>
<td>(SQFT &lt; 10)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>from cutoff</td>
<td>-0.023***</td>
<td>-0.023***</td>
<td>-0.023***</td>
</tr>
<tr>
<td>(1000s) feet</td>
<td>(0.005)</td>
<td>(0.005)</td>
<td>(0.005)</td>
</tr>
<tr>
<td>(SQFT ≥ 10)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>from cutoff</td>
<td>0.189***</td>
<td>0.106</td>
<td>0.031</td>
</tr>
<tr>
<td>(1000s) feet</td>
<td>(0.065)</td>
<td>(0.093)</td>
<td>(0.131)</td>
</tr>
<tr>
<td>Constant</td>
<td>0.189***</td>
<td>0.106</td>
<td>0.031</td>
</tr>
<tr>
<td></td>
<td>(0.065)</td>
<td>(0.093)</td>
<td>(0.131)</td>
</tr>
</tbody>
</table>

**Controls for Competitors in Zip Code**

- # Groceries & Convenience Stores
- Flexible Controls for Competitor Sizes

| N               | 73 | 73 | 73 | 246 | 246 | 246 | 3969 | 3969 | 3969 |

Notes: This table shows the increase in the likelihood a grocery/convenience store obtains a liquor license at the 10,000 Square foot threshold. Heteroskedasticity-robust standard errors in parentheses. Coefficients are significant at the *10, **5%, and ***1% level. SQFT is the square footage of the store, measured in thousands. Sample is from December, 2012 TDLinx and January, 2013 WSLCB licensure data. Flexible controls for store size include: number of stores in bandwidth (8-12, 5-14, all, respectively), number of stores above bandwidth (12+, 15+, 10+), and number of stores below bandwidth (0-7, 0-4, 0-9), within Zip code.
Table 3.3: Number of Stores that Sell Liquor vs. Stock of Potential Entrants

<table>
<thead>
<tr>
<th># Groceries by Size:</th>
<th>12 &gt; SQFT ≥ 10</th>
<th>8 &gt; SQFT ≥ 8</th>
<th>SQFT ≥ 12</th>
<th>Total</th>
<th># Licensed Groceries</th>
<th>Total # Liquor Licensees</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
<td>(5)</td>
<td>(6)</td>
</tr>
<tr>
<td>12 &gt; SQFT ≥ 10</td>
<td>0.299***</td>
<td>0.299***</td>
<td>-0.142</td>
<td>0.013</td>
<td>0.170</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.114)</td>
<td>(0.114)</td>
<td>(0.135)</td>
<td>(0.055)</td>
<td>(0.181)</td>
<td></td>
</tr>
<tr>
<td>12 &gt; SQFT ≥ 8</td>
<td>0.208***</td>
<td>0.204***</td>
<td>-0.019</td>
<td>-0.014</td>
<td>0.170</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.070)</td>
<td>(0.068)</td>
<td>(0.099)</td>
<td>(0.042)</td>
<td>(0.142)</td>
<td></td>
</tr>
<tr>
<td>8 &gt; SQFT</td>
<td>0.000</td>
<td>-0.000</td>
<td>0.009**</td>
<td>0.009</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.003)</td>
<td>(0.008)</td>
<td>(0.004)</td>
<td>(0.010)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SQFT ≥ 12</td>
<td>0.005</td>
<td>0.827***</td>
<td>-0.024**</td>
<td>0.808***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.008)</td>
<td>(0.027)</td>
<td>(0.010)</td>
<td>(0.030)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SQFT ≥ 10</td>
<td>0.468***</td>
<td></td>
<td></td>
<td>0.387</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.128)</td>
<td></td>
<td></td>
<td>(0.338)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>0.121</td>
<td></td>
<td></td>
<td>0.570*</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.109)</td>
<td></td>
<td></td>
<td>(0.303)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other Variables</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total SQFT Above Cutoff</td>
<td>0.008***</td>
<td></td>
<td></td>
<td>0.011***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
<td></td>
<td></td>
<td>(0.004)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total SQFT Below Cutoff</td>
<td>-0.014</td>
<td></td>
<td></td>
<td>-0.059</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.014)</td>
<td></td>
<td></td>
<td>(0.038)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>512</td>
<td>512</td>
<td>512</td>
<td>512</td>
<td>512</td>
<td>512</td>
</tr>
</tbody>
</table>

Notes: Each observation is a five-digit zip code in Washington state. Data on the number and size of groceries (including convenience stores and Superettes) is from TDLinx. Data on liquor licensure is from the WSLCB. Total SQFT Above Cutoff is the sum, across all groceries above 10,000 square feet in the zip code, of the square footage. Total SQFT Below Cutoff is defined analogously. Heteroskedasticity-robust standard errors reported in parentheses. Coefficients are statistically significant at the *10%, **5%, and ***1% level.
Table 3.4: Effect of Potential Entry on Price, by Liquor Type

<table>
<thead>
<tr>
<th>sample</th>
<th>Full</th>
<th>Top 10% Carried</th>
<th>Bottom 90% Carried</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
</tr>
<tr>
<td></td>
<td>(4)</td>
<td>(5)</td>
<td>(6)</td>
</tr>
<tr>
<td># Grocers</td>
<td>-0.029***</td>
<td>-0.006</td>
<td>-0.002</td>
</tr>
<tr>
<td>12 &gt; SQFT ≥10</td>
<td>(0.015)</td>
<td>(0.005)</td>
<td>(0.005)</td>
</tr>
<tr>
<td></td>
<td>-0.002</td>
<td>-0.006</td>
<td>-0.039*</td>
</tr>
<tr>
<td></td>
<td>(0.005)</td>
<td>(0.004)</td>
<td>(0.023)</td>
</tr>
<tr>
<td></td>
<td>0.032***</td>
<td>0.002</td>
<td>0.042**</td>
</tr>
<tr>
<td>12 &gt; SQFT ≥8</td>
<td>(0.011)</td>
<td>(0.003)</td>
<td>(0.017)</td>
</tr>
<tr>
<td></td>
<td>0.008*</td>
<td>0.004</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td>(0.004)</td>
<td>(0.004)</td>
<td>(0.004)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UPC Fixed Effects</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>N</td>
<td>926013</td>
<td>926013</td>
<td>328322</td>
</tr>
<tr>
<td></td>
<td>328322</td>
<td>597691</td>
<td>597691</td>
</tr>
</tbody>
</table>

Notes: Observations are at the zip code - week - UPC level. Standard errors clustered at the zip code level. Coefficients are statistically significant at the *10%, **5%, ***1% level. Controls include month of the year and product size (liters).
<table>
<thead>
<tr>
<th></th>
<th>Whiskey</th>
<th>Rum</th>
<th>Tequila</th>
<th>Vodka</th>
<th>Gin</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
<td>(5)</td>
</tr>
<tr>
<td># Grocers</td>
<td>-0.022</td>
<td>-0.008</td>
<td>-0.025**</td>
<td>-0.004</td>
<td>-0.019</td>
</tr>
<tr>
<td>12 &gt; SQFT ≥ 10</td>
<td>(0.023)</td>
<td>(0.006)</td>
<td>(0.012)</td>
<td>(0.005)</td>
<td>(0.015)</td>
</tr>
<tr>
<td># Grocers</td>
<td>0.030</td>
<td>0.002</td>
<td>0.013**</td>
<td>0.002</td>
<td>0.018**</td>
</tr>
<tr>
<td>12 &gt; SQFT ≥ 8</td>
<td>(0.019)</td>
<td>(0.004)</td>
<td>(0.008)</td>
<td>(0.003)</td>
<td>(0.009)</td>
</tr>
<tr>
<td>UPC Fixed Effects</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>N</td>
<td>264110</td>
<td>264110</td>
<td>162097</td>
<td>162097</td>
<td>101892</td>
</tr>
</tbody>
</table>

Notes: Observations are at the zip code - week - UPC level. Standard errors clustered at the zip code level. Coefficients are statistically significant at the *10%, **5%, *1% level. Controls include month of the year and product size (liters).
Table 3.6: Effect of Store Sizes on Market Outcomes

<table>
<thead>
<tr>
<th># Grocers</th>
<th>Log Quantity Per Store</th>
<th>Standard Deviation of Price</th>
<th>Number of Products</th>
</tr>
</thead>
<tbody>
<tr>
<td>12 &gt; SQFT ≥ 10</td>
<td>(1) -0.004 (0.048)</td>
<td>(2) 0.003 (0.053)</td>
<td>(3) 0.823 (2.239)</td>
</tr>
<tr>
<td>12 &gt; SQFT ≥ 8</td>
<td>(1) 0.002 (0.039)</td>
<td>(2) 0.001 (0.041)</td>
<td>(3) 1.512 (2.530)</td>
</tr>
<tr>
<td>UPC Fixed Effects</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
</tbody>
</table>

Notes: Observations are at the zip code - week - UPC level in columns (1)-(4). Observations are at the zip code level in column (5). Standard errors clustered at the zip code level. Coefficients are statistically significant at the *10%, **5%, ***1% level. Controls in columns (1)-(5) include month of the year and number of supermarkets. Controls in columns (1)-(5) also include product size (liters). Quantity per store is the average quantity sold in Nielsen Scanner stores matched to the five-digit zip code. Standard deviation of price is calculated at the zip code - week - UPC level. Number of products is the number of unique UPCs sold in the zip code in the first size months of privatization.
### Table 3.7: 2SLS Estimates of Entry on Log Prices

<table>
<thead>
<tr>
<th></th>
<th>OLS</th>
<th>2SLS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Full Sample</td>
<td>Whiskey</td>
</tr>
<tr>
<td># Licensed Grocers</td>
<td>0.026**</td>
<td>-0.149</td>
</tr>
<tr>
<td>12 &gt; SQFT ≥8</td>
<td>(0.011)</td>
<td>(0.179)</td>
</tr>
<tr>
<td># Grocers</td>
<td>0.086</td>
<td>0.071</td>
</tr>
<tr>
<td>12 &gt; SQFT ≥8</td>
<td>(0.077)</td>
<td>(0.072)</td>
</tr>
</tbody>
</table>

| N          | 926013 | 926013 | 264110 | 162097 | 101892 | 343223 | 54691 |

Notes: Observations are at the zip code - week - UPC level. Standard errors clustered at the zip code level. Coefficients are statistically significant at the *10%, **5%, *1% level. Controls include month of the year, number of grocery stores, month, and product size (liters).
Figure 3.12: Törnqvist Price Index Change at Liberalization and Zip 5 Population, State-balanced Panel

Figure 3.13: Törnqvist Price Index Change from Liberalization to End of 2012 and Zip 5 Population, State-balanced Panel
FIGURE 3.14: TÖRNQVIST PRICE INDEX FOR LIQUOR CATEGORIES, UNBALANCED PANEL
Bibliography


123


Haberman, Clyde. 2014 (Oct.). *The Cost of Campaigns*.


Kang, Cecilia, & Gold, Matea. 2014. With political ads expected to hit a record, news stations can hardly keep up.


Miravete, Eugenio J, Seim, Katja, & Thurk, Jeff. 2014. Complexity, Efficiency, and Fairness of Multi-Product Monopoly Pricing.

Mooney, Brian C, & Ailworth, Erin. 2012 (Sept.). Candidates could bump commercial ads on TV.


Silver, Nate. 2012. Tipping Point States.


