As Seen On TV: The Effects of Advertising on Demand for College

Job Market Paper

Elise A. Marifian *

December 2023

Abstract

U.S. colleges spend over $1 billion annually on advertisements, but to unknown effect. This paper investigates the effects of advertising on demand for college. My empirical strategy exploits regulation-induced discontinuities in TV advertising at local media market boundaries and is applied using a novel linkage between the universe of Texas high school graduates and the universe of spot TV advertising in Texas. I find that advertising increases college-going, with relatively larger effects for low-income and Hispanic students and less-selective colleges. I then develop and estimate a discrete choice demand model to quantify the effects of a ban on college advertising. The results indicate that five percent of college enrollees would change their enrollment decisions if college advertising were eliminated, with many students—particularly low-income and minority students—foregoing college altogether.

*Department of Economics, University of Wisconsin-Madison; contact: marifian@wisc.edu. I am indebted to my advisors, Jeff Smith and Chris Taber, and committee members, Naoki Aizawa and Matt Wiswall, for their guidance, support, and dedication. I am grateful to Jesse Gregory, Corina Mommaerts, Martin O'Connell, and Ashley Swanson, whose feedback significantly improved the project. Jonathan Becker, Sharada Dharmasankar, Amrita Kulka, Elan Segarra, Natalia Serna, Sandra Spirovska, and Tatiana Velasco generously offered their time and insight during countless discussions. I also thank Panle Jia Barwick, Chao Fu, John Kennan, Rasmus Lentz, Ananth Seshadri, and seminar participants at the UW-Madison Public Economics Workshop for helpful comments. The conclusions of this research do not necessarily reflect the opinion or official position of the Texas Education Research Center, the Texas Education Agency, the Texas Higher Education Coordinating Board, the Texas Workforce Commission, or the State of Texas. The contents in this paper are the researcher(s)' own analyses calculated (or derived) based in part on data from The Nielsen Company (US), LLC and marketing databases provided through the Nielsen Datasets at the Kilts Center for Marketing Data at The University of Chicago Booth School of Business. The conclusions drawn from the Nielsen data are those of the researcher(s) and do not reflect the views of Nielsen. Nielsen is not responsible for, had no role in, and was not involved in analyzing and preparing the results reported herein. This material is based upon work supported by the National Science Foundation under Grant No. 2149221. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation. In addition, the research reported here was supported by the Institute of Education Sciences, U.S. Department of Education, through Award #R305B150003 to the University of Wisconsin-Madison. The opinions expressed are those of the authors and do not represent views of the Department of Education. All errors are mine.

1 Introduction

College choice in the U.S. is a complex process. Prospective students must navigate a decentralized college system, and with nearly 6,000 diverse institutions, it can be costly to learn about different post-secondary options. To connect with students and lower their search costs, colleges are increasingly turning to paid advertising, spending an estimated $2.2 billion on ads in 2019—nearly $1,000 per new undergraduate student.[1] While advertising may increase a college's visibility and communicate useful information about its offerings, recent consumer protection investigations found that advertising by several large for-profit college chains contained false claims to lure students.[2] These cases have raised concerns that colleges may use advertising to exploit students' naivete, prompting several Congressional bills to regulate college advertising.[3] Despite the magnitude of college advertising spending and the recent policy activity, economic research on college advertising is scant.[4]

This paper studies the effect of college advertising on students' educational choices. I ask two questions: (1) How does advertising affect whether and where students enroll in college? (2) How would a ban on college advertising affect students' enrollment choices? A priori, the effects of advertising on college demand are ambiguous: advertising may increase college-going and/or cause students to switch schools, or neither.[5]

[1]Author's calculations using the advertising statistic from Marcus (2021) and fall first-time undergraduate enrollment from the Jan 2022 Errata "Corrected First-Time Freshman Enrollment," published by the National Student Clearinghouse.

[2]The Federal Trade Commission pursued enforcement action against DeVry University in 2016 and the University of Phoenix in 2019, each leading to a settlement of at least $100 million. DeVry University falsely claimed that one year after graduation, their bachelor's degree graduates had incomes that were 15 percent higher than bachelor's degree graduates from all other colleges. DeVry also falsely claimed that 90 percent of their graduates who were actively seeking employment obtained jobs in their field within six months of graduating. The University of Phoenix falsely claimed to work with high-profile companies to create job opportunities specifically for University of Phoenix students and to develop their curriculum. See https://www.ftc.gov/news-events/news/press-releases/2016/12/devry-university-agrees-100-million-settlement-ftc and https://www.ftc.gov/news-events/news/press-releases/2023/09/ftc-action-leads-us-dept-education-forgive-nearly-37-million-loans-students-deceived-university.

[3]For example, in 2019 Senators Durbin and Hassan introduced legislation that would prevent colleges from using federal funds on advertising, marketing, and recruiting. See https://www.govtrack.us/congress/bills/116/s867/text.

[4]To my knowledge, economists have published one academic article (Cellini and Chaudhary, 2023) and one think-tank report (Cellini and Chaudhary, 2020) focused on paid advertising by U.S. colleges. In a working paper, Armona and Cao (2022) estimate elasticities of for-profit advertising in the sub-baccalaureate market, but the authors ask a different research question than this paper, as they focus on the design of the federal financial aid system.

[5]Colleges may advertise for a variety of reasons, with different implications for the expected effect on

A challenge to estimating the causal effect of advertising on enrollment is that colleges make advertising choices based on unobserved factors that are correlated with demand. I address this endogeneity concern by exploiting an institutional constraint in the ad-buying process that generates a discontinuity in spot TV advertising at the boundaries of Designated Market Areas, or DMAs, which define local media markets. My causal research design is an adaptation of the *border strategy* approach in Shapiro (2018). Specifically, I use a differences-in-differences design applied at the border between two media markets, which allows me to leverage both temporal and spatial variation in the college advertising seen by individuals living in the same local labor markets but different media markets.

Implementing my identification strategy requires data on individuals' locations and college choices and colleges' advertising in those locations. For the former, I use student-level administrative records from the Texas Education Research Center (ERC). My analysis sample covers the universe of Texas public high school graduates, which I link to their initial college enrollment. College advertising data is not collected by government agencies. Instead, I use the Nielsen Company's[6] Ad Intel Database to generate a dataset of all college ads aired on TV in Texas, and perform a novel data linkage between colleges in the ERC data and the college advertisers in the Nielsen data. The resulting analysis dataset offers a near complete portrait of college TV advertising and individual-level college-going in Texas for the 2011-2015 cohorts of high school graduates.

I first examine the aggregate effect of college advertising on the extensive margin. Using the border strategy, I find that advertising increases college-going among recent high school graduates, with an elasticity of demand equal to 0.035. I then investigate whether responses to college advertising vary with demographic characteristics. This analysis reveals that low-income students—and, to a lesser degree, Hispanic students—are highly responsive to college advertising.

These estimates combine ads across all colleges, but it is reasonable to expect the effects of college advertising to vary by college type. To explore this hypothesis, I separately estimate the extensive margin effects of advertising by community colleges, for-profit colleges, public 4-year colleges/universities, and private non-profit colleges. I find large effects for non-selective institutions—particularly community colleges and for-profit colleges; estimated effects for public universities and private colleges are imprecise, likely because these colleges advertise on local TV much less.

After documenting the presence of extensive margin effects, I turn to the question of whether

demand. For instance, non-selective institutions may advertise to attract students on the margin of enrolling. Other colleges may advertise to attract a certain type of student, or because their competitors are advertising. Selective institutions with an eye toward college rankings may advertise to increase applications so that their acceptance rate falls.

[6] Nielsen is a consumer and media data analytics company known for producing TV ratings.

advertising induces students attend a different college than they would in its absence. To answer this question, the ideal empirical framework would accommodate individual choice from a large number of college options, each of which can advertise in different DMAs. To this end, I develop a discrete choice model of college demand with advertising. I leverage the detailed student-level data to estimate heterogeneous preference parameters, similar to the approach in Bayer et al. (2007), and use the insights from Berry (1994) and Berry et al. (1995) to estimate college-specific taste parameters. With the estimated model, I then simulate the effect of the policy counterfactual of interest—banning college advertising—and examine changes to students' enrollment choices.

I find that shutting down advertising would affect both whether and where students attend college. It would reduce overall enrollment and lower the market shares of community colleges and for-profit colleges while increasing the market shares of public universities and private colleges. Additional simulations reveal that more than 5 percent of college enrollees would make a different college choice in the absence of college advertising. Among these, minority and low-income students are more likely to forego college than switch to a different college, while white students are split evenly between these two choices. Reassuringly, results from the counterfactual advertising ban are consistent with the findings from the border strategy.

The primary contribution of this paper is to generate novel empirical facts about college advertising and its effects on demand. The findings shed light on the types of students that respond to college advertising and the types of colleges for which advertising increases college-going.

A secondary contribution is to estimate a structural model of college choice that represents the true breadth of college options that prospective students can consider, as I include *all types of colleges*—selective public 4-year colleges, private non-profit colleges, non-selective public 4-year colleges, 2-year community colleges, and 2- and 4-year for-profit colleges. I am not aware of prior papers using structural models that study college choice in this comprehensive way, as many studies exclude non-selective public colleges and nearly all exclude for-profit colleges. By defining the college set broadly, I am able to allow heterogeneity in preferences for different types of colleges, which generates more realistic substitution patterns.

My paper adds to the literature on information frictions in college choice. Economists studying higher education have long been concerned about the inefficiencies caused by information frictions in the college search and application process (see, e.g. Avery and Kane, 2004; Dynarski and Scott-Clayton, 2006; Hoxby and Avery, 2013; Pallais, 2015; Scott-Clayton, 2012). Gaps in students' knowledge about post-secondary programs—be it their existence, value, price, or other characteristics—are especially common among first-generation college and economically disadvantaged students, and can lead to sub-optimal human capital investment decisions, such as under-matching and under-enrollment (Hoxby and Avery, 2013; Pallais, 2015; Scott-Clayton, 2012; Smith et al., 2013).

Economic research on college information frictions has primarily focused on evaluating policy and information interventions (Avery et al., 2021; Bergman et al., 2019; Bettinger et al., 2012; Dynarski et al., 2021; Hoxby and Turner, 2013, 2015; Page and Scott-Clayton, 2016); little attention has been given to the market's own response to information frictions via advertising. Economics research on colleges' paid advertising is relatively new, beginning with a descriptive report on aggregate trends in college advertising by Cellini and Chaudhary (2020). More recently, Cellini and Chaudhary (2023) investigate the geographic targeting of college ads. My project contributes to this emerging literature by studying how advertising affects students' post-secondary choices.

This paper also builds upon a growing literature focused on credibly identifying the causal effect of advertising on demand (Aizawa and Kim, 2018, 2019; Shapiro, 2018; Sinkinson and Starc, 2019; Dubois et al., 2018; Tuchman, 2019; Spenkuch and Toniatti, 2018). My paper contributes to this literature by studying the effects of advertising on demand in the market for higher education.

The paper is organized as follows. In Section 2, I discuss the institutional setting and describe the data. I then describe the border strategy approach and the corresponding estimation strategy in Section 3, after which I present the results on market expansion (Section 4). In Section 5, I present the discrete choice demand model and its results and perform counterfactual simulations to quantify the effect of a ban on advertising. I offer concluding remarks in Section 6.

2 Setting and Data

2.1 Education in Texas

My empirical analysis uses data from Texas. As an empirical setting to study the effects of college advertising, Texas is attractive for several reasons. Ranking second in terms of state population, GDP, and area, Texas also has the second largest state public school system in the U.S., enrolling over 10 percent (5 million) of the nation's public K-12 students in 2011. In addition, Texas' annual domestic out-migration rate is one of the nation's lowest, at 15 per 1,000 people.

Elementary and secondary education is overseen by the Texas Education Agency (TEA), while the Texas Higher Education Coordinating Board (THECB) regulates nearly 420 post-secondary educational institutions, systems, and independent organizations.[7] The THECB categorizes Texas institutions into one of 5 types or sectors: public universities (4-year), public community/technical/ state colleges (2-year), independent (private non-profit) colleges, career (for-profit) schools and

[7] Of these, approximately 300 institutions are traditional colleges and universities in 2010 (excluding branch campuses and extension centers). See Appendix A, THECB Reporting and Procedures Manuals for Texas Universities, Health-Related Institutions, Community, Technical, and State Colleges, and Career Schools and Colleges 2010, pages A1-A14 for details.

colleges (2- and 4-year), and health-related institutions. The 2010 breakdown of institution types, excluding health-related institutions,[8] is shown in Table 1. Public institutions account for 43 percent of Texas colleges, and more than one-third are for-profit/career colleges.

Table 1: Texas Colleges by Type, 2010

Type/Sector	Level	Control	Count	Percent
Universities	4 Year	Public	44	14.5
Community/Technical	2 Year	Public	88	28.9
Independent	4 Year*	Private Non-Profit	67	22.0
Career	2 and 4 Year	For-Profit	105	34.5

* Includes two independent junior colleges–Jacksonville College and Lon Morris College.
Notes: Percents may not sum to 100 due to rounding. Author's calculations using THECB Reporting and Procedures Manuals for Texas Universities, Health-Related Institutions, Community, Technical, and State Colleges, and Career Schools and Colleges 2010, Appendix A.

2.2 Local TV Advertising

My identification strategy relies on variation in local TV advertising, so in this section I describe the key institutional factors relevant to my research design.

Traditional or "linear" TV is television content transmitted via broadcast, cable, or satellite.[9] In the U.S., television advertising for traditional TV can be purchased both at the national and local levels. The U.S. has 210 local TV media markets, called "Designated Market Areas," or DMAs, which are collections of counties typically centered about a major city. The DMA geography was created by the Nielsen Company in the1960s based on local TV viewing behaviors. The counties in a DMA constitute the major viewing audience for television stations in their metropolitan area.[10] In 1996, the Federal Communications Commission (FCC)[11] began using the DMA geography as its definition of a local market for the purposes of broadcast television station[12] carriage

[8]Health-related institutions primarily serve students at the graduate level in medical, dental, pharmacy, veterinary, and other health-related degree programs.

[9]The term "linear" refers to the fact that the content follows a predetermined programming schedule, in contrast to on-demand content.

[10]Specifically, each county is assigned to the DMA from which the majority of that county's viewed channels originate. The DMA boundaries are reviewed annually but rarely change. DMAs generally correspond to the Office of Management and Budget's metropolitan statistical areas.

[11]The FCC regulates interstate communications by radio, television, wire, satellite, and cable.

[12]The U.S. Code of Federal Regulations (CFR) defines a "local commercial television station" as "any full power television broadcast station...licensed and operating on a channel regularly assigned to its community

rights under cable and satellite television "must-carry" and "retransmission" rules.[13,14] Since then, DMAs have been the primary geography used for television media, advertising sales, and audience measurement.

With few exceptions, all households within a DMA receive the same local television advertising because FCC regulations stipulate that cable and satellite providers may only deliver content to the station's "local market," i.e., the DMA in which it is located.[15] The FCC also monitors broadcasters and broadcast signals to ensure that they serve their "community of license."[16] During the time period studied (2010-2016), DMAs are the smallest geographic area that can be targeted for local TV advertising.[17]

My research design leverages spatial variation in local TV advertising across DMAs. To ensure that this variation is plausibly exogenous, I focus my attention on high school students living near the boundary between two DMAs, as crossing the border generates a discontinuity in advertising levels. I use the term "border county" to denote a county that borders another DMA; the remaining counties are "interior counties." Figure 1b shows the 20 Texas DMAs (outlined in

---

by the Commission that, with respect to a particular cable system, is within the same television market, as defined below in paragraph (e) of this section, as the cable system...." 47 CFR § 76.55(c); 47 U.S.C. § 534(h)(1)(A). Paragraph (e) defines a television market: "a commercial broadcast television station's market ... shall be defined as its Designated Market Area (DMA) as determined by Nielsen Media Research and published in its Nielsen Station Index Directory and Nielsen Station Index US Television Household Estimates or any successor publications." 47 CFR § 76.55(e). The language is similar for satellite carriage.

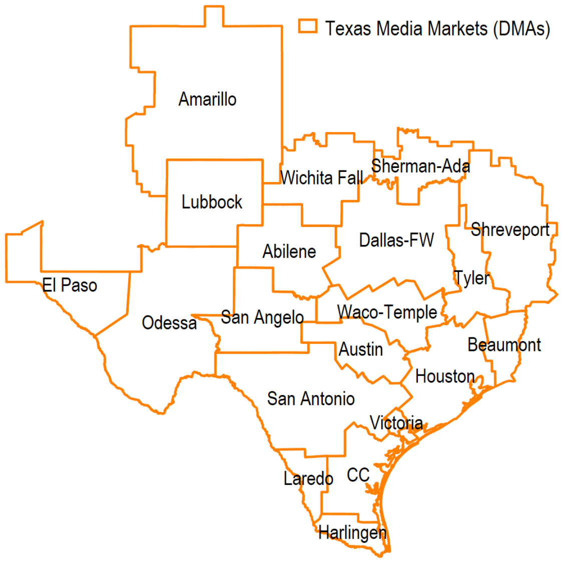
[13]Under the Communications Act, cable systems or other multivideo programming distributors (MVPD) must obtain the consent of a commercial television broadcast station to carry its broadcast signal ("retransmission consent"); this permission may involve compensation from the cable provider to the broadcaster. Alternatively, local broadcast stations may require that cable systems located within their market carry their signal ("must carry"); in this case, the broadcast station cannot demand compensation from the cable operator for carriage of its signal. The broadcast stations must choose once every 3 years (on a system-by-system basis) whether to give retransmission consent or assert mandatory carriage rights. Satellite carriers are subject to the "carry-one, carry all" rule, which requires satellite carriers to provide subscribers all local television broadcast stations' signals in their DMA if the satellite company carries at least one local television broadcast signal. See https://www.fcc.gov/media/engineering/cable-television and https://www.fcc.gov/media/television-broadcast-stations-satellite.

[14]This switch was necessary because The Arbitron Company stopped producing the "Area of Dominant Influence" ("ADI") market list, which the Commission had used as its definition of a local market. See https://www.govinfo.gov/content/pkg/FR-1996-06-10/html/96-14571.htm.

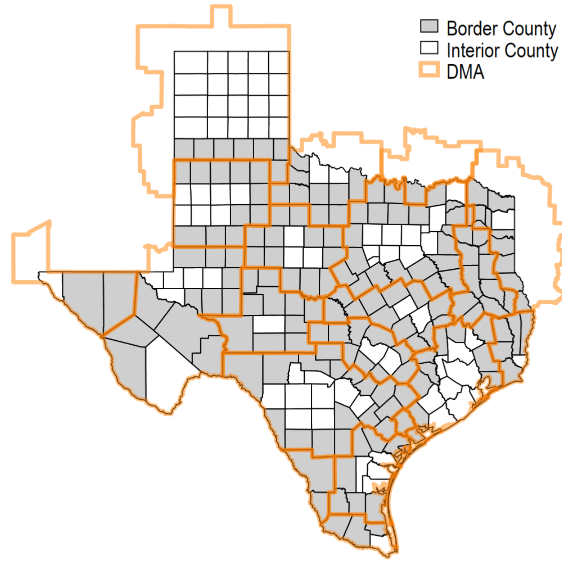
[15]The FCC permits cable and other MVPD providers to petition the FCC to provide a broadcast station's content to an area outside the local market (DMA) if the station is considered "significantly-viewed" in the area (a large portion of the area receives the broadcast signal). These exceptions are published in the FCC's "Significantly Viewed Stations List." See https://www.fcc.gov/sites/default/files/significantly-viewed-stations.pdf.

[16]See `https://transition.fcc.gov/localism/Localism_Fact_Sheet.pdf`.

[17]Recent technological advances increasingly have allowed advertisers to target consumers individually.



(a) DMAs Containing Texas Counties



(b) Border Counties in Texas DMAs

Figure 1: Texas Media Markets

orange) and their counties (outlined in black), with border counties shaded in gray.

2.3 Data

My empirical analysis relies on two key types of data: (1) individual-level data with background information, geographic location, and educational histories, and (2) local TV advertising by colleges. To this end, I construct a novel dataset that combines administrative educational records from the Texas Education Research Center (ERC) and occurrence-level TV advertising data from The Nielsen Company.

2.3.1 Individual-Level Educational Data

Individual-level data come from the longitudinal P-20/Workforce Data Repository housed in the Texas Education Research Center (ERC) at The University of Texas at Austin. The Texas ERC database contains restricted administrative datasets from agencies in Texas and covers individuals' pre-K through post-secondary educational histories. I combine a number of ERC files to create an individual-level dataset that includes demographic, high school academic, and post-secondary choices of Texas high school graduates. Within the repository, the relevant data sources for my study come from the TEA, the THECB, and the National Student Clearinghouse (NSC).

The analysis sample begins with public high school graduates from the TEA. For each student that graduates from a Texas public high school between academic years 2010-2011 and 2014-2015, I obtain demographic, enrollment/graduation, attendance, and standardized test score data, as well as organizational (school, district) and financial data. I then link these students to the THECB and NSC administrative records,[18] which captures all of their enrollment at in-state and out-of-state Texas colleges. Importantly, I do not observe individual students' addresses, so I assign their location as the address of the high school from which they graduate.

The sample of colleges is constructed from two sources: ERC files and surveys from NCES' Integrated Postsecondary Education Data System (IPEDS). I begin with the set of colleges in the Texas ERC files (i.e., those that report to the THECB or the NSC) that enroll a student from my analysis sample. I then match those institutions to their records in the IPEDS surveys to obtain location information.[19] Colleges are then excluded from the sample if they are characterized as a theological seminary or other type of special focus institution (such as a university center or dental school); are a high school or administrative unit; or do not offer undergraduate degrees or certificates. These exclusions yield a sample of 340 Texas colleges from 2011 to 2016.

2.3.2 TV Advertising Data

The U.S. Department of Education does not collect data on colleges' advertising behavior. Therefore, to study college advertising, I use data from The Nielsen Company's Ad Intel database.[20]

Like related studies (Shapiro, 2018; Sinkinson and Starc, 2019), I focus on local TV advertising.[21] Because my objective is to understand how advertising affects college enrollment, it would be useful to observe individual-level exposure to college ads. Unfortunately, to my knowledge, such measures are unavailable in the U.S.[22] Instead, I use ad volumes (i.e., the total number of

[18]The THECB dataset contains records for students who attend Texas post-secondary institutions. The ERC's NSC dataset contains college enrollment data for Texas high school graduates who attend non-Texas institutions (covering 98 percent of students enrolled in U.S public and private higher education).

[19]This matching can be difficult due to differences in reporting to the THECB versus reporting to IPEDS. For instance, sometimes ERC institutions are reported at the branch level whereas IPEDS will combine branches into a single institution. When this is the case, I manually look up the institution's address. Figure 12 shows a map of the Texas colleges in IPEDS in 2014 by institution control.

[20]Ad Intel is accessed through the Kilts Center for Marketing Data at The University of Chicago Booth School of Business.

[21]Even though Ad Intel data also includes occurrence measures for local print, radio, internet and cinema media types, the geographic coverage can vary widely. Many of the Texas DMAs are not covered in the measurement of non-TV media types, with the exception of outdoor media (billboards). Additional details of Nielsen media types are available in Appendix A.1.1.

[22]Even at the DMA-level, exposure measures are unavailable for most media types. Ad Intel contains Nielsen's measure of advertising exposure, referred to as "impressions," for TV, cinema, internet, and radio media, but these data have limited geographic coverage. Specifically, in 18 of the 20 Texas DMAs, local (DMA) demographic-level impressions for spot TV ads are collected for just four weeks per year, during each of the four (week-long) "sweeps" periods, when Nielsen's TV panelists log their TV viewing data in diaries.

local ads) as my measure of the college advertising treatment.[23]

Data coverage begins in 2010, and the data are released annually. I use the television "occurrences" files, which contain records of TV commercials.[24] I restrict the analysis to spot TV ads, for which the measurement coverage is, for all intents and purposes, universal.[25]

Each spot TV observation contains detailed information about the source and characteristics of the ad. For my purposes, the key data elements include the product category, brand, airing date and time, Nielsen's estimates of the spending (in dollars) on the ad, distributor, DMA of airing, and advertiser subsidiary and parent.

2.3.3 Auxiliary Data: High School, College, and County Characteristics

In addition to the ERC and advertising data, I use several auxiliary datasets. I require data on the latitude and longitude of Texas high schools and colleges to construct measures of distance from each high school to each college. The latitudes and longitudes of Texas colleges are not available in a single data source. As mentioned previously, I collect some latitude and longitudes from IPEDS; slightly more than half of the THECB colleges with positive enrollment during my sample period are matched to IPEDS colleges. These are predominately public and private non-profit colleges. For the remaining institutions, which are predominately for-profit colleges, I manually collect the latitude and longitude.[26]

I obtain Texas high school locations from the U.S. Department of Education's Common Core of Data. My empirical strategy requires me to identify high schools near the DMA borders, I use the high school's latitude and longitude to compute the distance from the high school to the nearest DMA border. I also use this information to identify the nearest neighboring county (on the other side of the DMA border), from which I form county pairs (described in Section 3.2).

Finally, because my analysis compares individuals in neighboring counties, it is important to

---

[23]For robustness, I also estimate the main results using spot TV ad spending; in general, the estimates are smaller in magnitude, but the patterns and qualitative conclusions are unchanged. Results are available upon request.

[24]Although Ad Intel includes two types of TV advertising measures—occurrences (airings) and impressions (the views associated with each airing)—the impressions data coverage is limited, so only the occurrences data work in my context. Specifically, year-round impressions data for 18-20 year olds are only available in the 25 largest DMAs in the U.S., so 18 of the 20 Texas DMAs do not have the necessary data. My identification strategy relies on cross-DMA variation, so this limitation prevents me from using the impressions data to construct my advertising measures.

[25]National TV coverage is also universal.

[26]Most of these latitudes and longitudes are obtained by manually matching THECB colleges to firms in the Your Economy Time Series (YTS) database from the Wisconsin Business Dynamics Research Consortium. For the remaining colleges, I collect the institution's address and use the U.S. Census Bureau's Geocoder to convert the address to latitude and longitude. Additional details are available on request.

control for year-to-year variation in factors that may affect local demand for college. For these purposes, I collect annual county-level economic data, including county population, civilian labor force, unemployment rate, and poverty rate data from the U.S. Census Bureau's Small Area Income and Poverty Program.

2.4 Variable Definitions

Cohorts The primary analysis sample comprises Texas public high school graduates from five cohorts, 2011 through 2015. I begin with the 2011 graduating class because Nielsen advertising data are unavailable prior to 2010; thus, the 2011 cohort is the first cohort for whom I observe TV ads during their entire senior year.

Covariates Student-level demographic controls come from the TEA high school enrollment and graduate files and include indicators for sex, race/ethnicity (white, Black, Hispanic/Latino, Asian, American Indian or Alaska Native, Native Hawaiian or Pacific Islander, and two or more races), free or reduced price lunch receipt (FRPL) or other economic disadvantage (social assistance such as WIC, SNAP, or TANF), and English language proficiency. Academic variables include high school graduation date, scaled 8th grade standardized test scores in reading and math (each standardized within the year),[27] and indicators of students identified as gifted or as having special needs, participation in a vocational-technical program, and attendance at an out-of-district high school. I also construct a student's attendance rate in the last year of high school and an indicator of spring high school graduation. School-related variables come from the district and campus files and include indicators of school type (e.g., traditional school, juvenile justice alternative program, etc.), charter status, and metro status.

Advertising Treatments To construct advertising measures, I first restrict the sample to spot TV ads with a product category of "college", "university," or "institute"[28] and exclude ads aired outside the Texas DMAs. I use the ad date to map each ad to an academic year, defined as August 1-July

[27] Prior to the 2011-2012 school year, high school students were tested in reading and math in 9th through 11th grade, and passing the 11th grade test was required for high school graduation. In 2011-2012, the TEA changed the standardized testing system, switching to course-specific tests (e.g. English, Algebra, History, Biology), and students could take these exams at different points in their high school career. As a result, a common and consistent high school test score measure is not available for the majority of the sample. For this reason, I do not include test scores in my regression specifications. I do, however, use the test score data for aggregate descriptive analyses and robustness checks.

[28] I include the institute category because a number of colleges with the word "institute" in their name are categorized by Nielsen as institutes instead of colleges/university (e.g., ITT Technical Institute and MIT).

31. I then collapse the dataset to the brand-DMA-(academic) year level, counting the total number of occurrences and summing the spending.[29]

Once the advertising data is at the brand-DMA-year level, I link the Nielsen brands with the colleges in the THECB data. There is no crosswalk between the advertising brands and U.S. colleges, so I use string and manual matching to identify the brands in Nielsen that are Texas colleges. The unmatched Nielsen brands that advertise in Texas are subsequently grouped into two categories: out-of-state colleges and national online college chains.[30] The matching process is complicated by the fact that a brand in Nielsen may correspond to multiple college campuses; this is often the case with large for-profit college chains (e.g., University of Phoenix, ITT Technical Institute, Kaplan College), which may have several campuses in the same or in different DMAs, but it also can occur with large public university or community college systems (e.g., Texas A&M or Alamo Community College brands). When an advertising brand could match to more than one THECB campus, I also use the DMA of airing to assist with the match.[31]

Outcomes I define college enrollment as enrolling in any college either in the academic year of high school graduation or the subsequent academic year. I do this to capture the first enrollment among individuals that graduate in December or other times of the year outside of the traditional May/June graduation period. Under this definition, the overall college-going rate is 63.8 percent, but it varies by DMA, ranging from 53 percent in Amarillo to 75 percent in Laredo.

2.5 Sample Construction

The primary analysis file is constructed by matching individual-level data from the Texas ERC to data on local TV ads aired in Texas DMAs by Texas colleges.

To construct the Texas ERC sample, I start with individuals who graduate from Texas high schools between 2011 and 2015. For border strategy analyses, I restrict the sample to students whose high school is located in a border county. As previously noted, students' addresses are

[29] I also clean the National TV ad data for use in robustness checks.

[30] There are some brands that correspond to non-regulated (non-THECB) in-state career schools, such as beauty schools; I exclude these institutions from the advertising measures.

[31] I use the following approach: If there is a THECB campus located in the same DMA as the ad is aired, then that campus is matched to the brand-DMA-year observation. If none of the campuses are located in the same DMA as the ad, I match the observation to the closest campus (or campuses if they are located in multiple DMAs and are within a similar distance from the border of the DMA where the ad is aired). When a given brand-DMA-year observation is matched to more than one campus, for each match, I divide the advertising measures (number of ads and spending) by the number of matched campuses. For Texas A&M, The University of Texas at Austin, Texas Tech, and the University of Houston, I assign the ads to the flagship because in many cases, one of the other system campuses appears as its own brand.

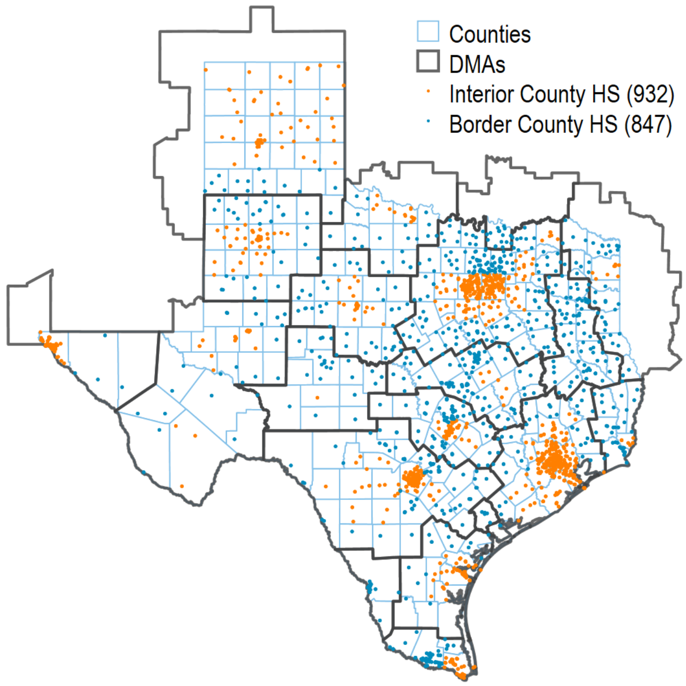


Figure 2: Border County High Schools in Texas

unavailable; the high school is the best proxy for their location of residence. Figure 2 shows a map of Texas DMAs, counties, and high schools. Interior high schools are represented by orange dots, while border county high schools are in blue.

In most specifications I restrict the sample to high schools near a border. For each high school, I compute the distance to the nearest boundary between two DMAs and generate indicators of whether the high school is located within 10 miles and 30 miles of the border.

I link each individual to their post-secondary enrollment data (if any) as follows: For each cohort, I match the individuals to all enrollment spell records. I then restrict the sample to enrollments occurring in the academic year of high school graduation or the subsequent academic year, dropping any observations that are dual-enrollment (attending high school and college simultaneously). Among this sample, some students will appear multiple times if they attend multiple institutions or enroll multiple semesters. I keep each student's first college enrollment occurrence in the two year period.

2.6 Raw Variation in College Advertising

My identification strategy relies on variation in advertising across DMAs over time. In this section I report the raw variation in colleges' spot TV advertising. Figure 3 reports the total number of DMA ads purchased by Texas colleges in 2010 and 2015 across local media. Shading corresponds to quintiles of ad volumes. Panel (a) shows volumes in 2010, the first year of the sample, while panel

(b) shows volumes in 2015. In 2010, the Dallas-Fort Worth DMA had the most college advertising by volume, nearly 63,000 college ads, followed by San Antonio (56,000 ads) and Houston (39,000 ads).

By 2015, Texas college ad volumes had more than doubled in Dallas (210 percent) and San Antonio (219 percent), and nearly doubled in Houston (197 percent). Even though Dallas and Houston are two of the fastest-growing metropolitan areas in the nation, their increase in college advertising volume is an order of magnitude larger than their population growth (+ 20 percent from 2010 to 2020).[32]

3 The Effect of Advertising on Demand for College

Identifying the causal effect of advertising on college demand is challenging because advertising is potentially endogenous. Firms, including colleges, make advertising decisions based on the expected return, so there may be factors that simultaneously affect colleges' advertising choices and demand for college that are unobserved to the researcher. To complicate matters, advertising may correlate positively or negatively with unobserved college demand. For instance, colleges may advertise more in markets with growing tastes for college, and thus higher growth potential. Alternatively, they may target markets where they anticipate a slowdown in demand (from a strengthening labor market) or increased competitive pressures (from increased rival advertising). A naive model, such as a simple regression of enrollment on advertising, would over-estimate the true effect in the former case and under-estimate it in the latter.

As a concrete example, consider how the opening of a new large employer, such as a manufacturing plant, might affect demand for colleges with technical/vocational programs. Some students who otherwise would have enrolled in a vocational program at a nearby community college or for-profit career college may instead seek employment at the new plant. Suppose that local colleges, anticipating this negative shock to demand, respond to the plant opening by increasing advertising, which has a positive effect on their enrollment. The effect obtained from a naive regression model, which does not account for the unobserved negative shock to local college demand, would be downward biased from the true effect. It is possible that the omitted variable bias is large enough to dwarf the true effect and lead the researcher to conclude that college advertising has little to no effect on enrollment, even though it actually increases demand and mitigates the enrollment declines from the plant opening.[33]

[32] See https://www.census.gov/library/stories/2021/08/more-than-half-of-united-states-counties-were-smaller-in-2020-than-in-2010.html

[33] On the other hand, one can tell a slightly different story of the effect of a plant opening. Suppose that

To address this potential endogeneity, I take an approach similar to several recent studies aimed at credibly identifying the causal effect of advertising on demand (Shapiro, 2018; Spenkuch and Toniatti, 2018; Tuchman, 2019; Shapiro et al., 2021; Aizawa and Kim, 2022). My *border strategy* research design uses both spatial and temporal variation in advertising for identification.

One way to leverage the spatial variation in advertising would be to use a *boundary discontinuity design* (BDD) to compare the college-going of individuals who live on either side of the DMA boundaries.[34] The general idea is that, conditional on observables, high school students living near one another have the same latent demand for college (because they share the same labor market opportunities and are exposed to the same local shocks to college demand, like the factory opening described above).[35] However, if these students live in different DMAs, then when they watch the same TV program, they are exposed to different local ads. Therefore, any differences in college-going stem from the difference in advertising at the DMA boundary.

Unfortunately, the BDD is unlikely to be valid in my context. Identification requires that the unobservables that affect demand for college evolve continuously through the DMA border;[36] ideally, I would observe students' home addresses to create an accurate measure of the running variable—the distance to the boundary. Unfortunately, students' locations are not observed in my data. This limitation can jeopardize identification because I do not know which students live closest to the border, for whom the continuity assumption is most likely to hold. The most granular location data I observe for each student is the high school attended. However, using high school location as a proxy for individuals' locations is further complicated by the fact that high school district boundaries can cross DMA boundaries. In such instances, students can live on either side of the boundary—but I do not observe which one—so I cannot assign their advertising treatment.

many of the jobs require some post-secondary training and that students respond to the new opportunities by acquiring the necessary training. In this scenario, the opening of the manufacturing plant generates a positive shock to local college demand. Suppose that colleges, recognizing that more students will want to receive training, respond to the opening by increasing their advertising, hoping to capture new college-goers whose default choice (absent advertising) would be a competitor. Even if college advertising does have a positive effect, the estimated coefficient in the naive model would be upward biased and may cause the researcher to overstate the true effect of college advertising by attributing to it the enrollment gains from increased local demand for post-secondary training.

[34]The BDD is an application of the regression discontinuity design when the running variable—the variable that determines who receives treatment, has a spatial component (e.g., distance). Two important early applications of this logic in education contexts are Black (1999) and Bayer et al. (2007).

[35]See Hahn et al. (2001) for a formal exposition of the identification requirements in a regression discontinuity design.

[36]It must be the case that colleges cannot precisely target their advertising to individuals at the DMA boundary. This assumption is plausible given the institutional constraints of the local TV ad buying process: during the period of study, local TV ads can only be purchased at the DMA level, so colleges cannot micro-target their TV ads to more localized geographies.

Consequently, I need to use individuals whose high schools are far enough from the boundary that the school district is fully contained in one DMA, which makes the continuity assumption less tenable.

Given these limitations, I take an approach that slightly deviates from the BDD, building upon the "border strategy" adopted in Shapiro (2018). My approach takes advantage of the institutional factors that the BDD exploits for identification while also making use of panel variation to address potential violations of the BDD's identifying assumptions in my context.

3.1 Border Strategy

As described above, the problem with using the BDD in my context is that I am unable to restrict the analysis to individuals in a narrow bandwidth around the DMA border.[37] For this reason, I want to allow individuals on either side of the DMA boundary to have different levels of college-going in the absence of advertising. To achieve this, I adopt a difference-in-differences design using students near DMA boundaries.

3.1.1 Econometric Model

Observations are at the individual level, and each individual is observed once. Denote individuals by i, high schools by s, counties by c, and DMAs by m. Furthermore, define a *border pair* as two adjacent counties in different DMAs (i.e., two counties that share an edge which is also a DMA boundary), and denote border pairs by b. Last, let t denote high school graduation cohorts (August 1 to July 31).

To identify the effect of advertising on college-going, I specify a differences-in-differences model where the sample is restricted to individuals in border pairs. The baseline specification takes the form:

$$y_i = a_{m(i)t(i)}\beta + z_i\gamma_1 + x_{c(i)t(i)}\gamma_2 + \theta_{b(i)t(i)} + \tau_{s(i)} + \varepsilon_i. \tag{1}$$

The variable y_i is a binary indicator of whether individual i enrolls in any college by the end of the next academic year. On the right-hand side, the treatment of interest is $a_{m(i)t(i)}$, the number of spot TV advertisements by Texas colleges aired in i's DMA in the year they graduate high school and the subsequent year. There are two sets of covariates: z_i (individual covariates) and $x_{c(i)t(i)}$ (county-level economic covariates in the year i graduates from high school). Finally, ε_i is an error.

The model includes two sets of fixed effects, both of which are important for identification. The first is a set of border pair-by-cohort fixed effects $\theta_{b(i)t(i)}$, which control for unobserved factors

[37] Bayer et al. (2007) use bandwidths of 0.2 miles or less. In my setting, the high schools located 0.2 miles from DMA border are most likely to enroll students who live in both DMAs.

affecting college-going that are common to the individuals in a given border pair-cohort. Implicitly, by including these fixed effects and restricting the sample to individuals near the DMA borders, my model partially mimics a boundary discontinuity design, which would compare individuals that live near one another but receive different advertising. Importantly, the border pair-by-cohort fixed effects net out local labor and college market shocks that are shared by students on either side of the DMA boundary within a cohort. They also partial out other time-specific effects within the border-pair, which would include other sources of ads that are the same in both counties (e.g., national TV ads).

On their own, the border pair-by-cohort fixed effects would assume that individuals near the border between two DMAs are as good as randomly assigned to either side. Because I exclude the high schools closest to the border, it is important to control for underlying differences in college demand across the DMA boundary within a border pair. To this end, I include a set of high school fixed effects, $\tau_{s(i)}$. These fixed effects control for unobserved, time-invariant heterogeneity at the high school level. They allow the college-going behaviors to differ on either side of the DMA within a border pair and help absorb unobserved factors that are high school-specific and time-invariant, such as a highly skilled college counselor.[38]

Given these fixed effects, the remaining unobserved heterogeneity can come from two sources: time-varying unobserved heterogeneity at the DMA, county, or high school level, and individual-level factors that are correlated with both advertising exposure and college going. To control for time-varying heterogeneity in college-going between the two counties in a border pair, I include controls for county population and median household income in the vector $x_{c(i)t(i)}$.[39]

To account for unobserved heterogeneity at the individual level that is not absorbed by the high school fixed effects, I include a set of student-level demographic and academic covariates z_i that may be correlated with TV viewing behavior and/or colleges' ad targeting and that affect college attendance. These variables include the student's high school attendance rate in their senior year, indicators capturing socioeconomic disadvantage (receipt of free lunch or other means-tested social assistance, and reduced price lunch eligibility), race/ethnicity, participation in a vocational-technical curricular program in high school, limited English proficiency, academic giftedness (as defined by the district), and receipt of special education services. I include a flag for students who attend an out-of-district high school, which may reveal unique academic preferences and greater willingness to travel for education. In addition, I include an indicator for spring graduation because

[38]Note that it is not possible to include county-by-cohort or high school-by-cohort fixed effects because they would fully absorb the variation in advertising (which is at the DMA-year level).

[39]For robustness, I also tried specifications adding unemployment rate, civilian labor force, and percent of the population in poverty, but they had no effect on the treatment coefficients.

students who graduate in the fall or winter may have more time to watch TV in the spring semester. Finally, I include a flag for graduation during the high school's modal graduation month to control for possible differences in college-going by students on a non-traditional academic schedule.

In summary, my border strategy is implemented as a differences-in-differences model where the parallel trends assumption is enforced separately within each border pair. The identifying assumption is that any conditional differences in college-going trends between the two counties in a border pair arise from differences in advertising. Note that the border strategy allows differences in advertising levels between two neighboring DMAs to arise from strategic choices by colleges; i.e., a college's optimal DMA-level advertising choices may be correlated with aggregate DMA-level demand shocks. The idea is that demand shocks in heavily populated areas (i.e., the interior of the DMA) are likely to have a greater influence on colleges' (DMA-level) advertising choices; whereas demand shocks in remote areas, which make up a relatively small portion of the DMA population, are likely negligible for colleges' advertising decisions. The border comparison is valid as long as college-going behaviors within a border pair would trend in parallel on either side of the border in the absence of advertising.

To gain intuition for the strategy, recall that institutional factors force advertisers to purchase local TV ad spots at the DMA level, rather than at a more granular geography, which would allow them to target their intended audience more effectively. If colleges could micro-target their ads, one would expect that their optimal advertising choices would differ between the border and interior counties but would be similar for neighboring border counties. The discontinuity in advertising at the DMA boundary exists precisely because colleges cannot target ads to more granular geographies.

3.2 Estimation

In this section, I describe the estimation details for the baseline model. The baseline analysis sample includes students who graduated from a public high school that is located in a border pair between 10 and 30 miles from the shared boundary.

3.2.1 Border Pairs

To implement the border strategy, I assign each high school to a unique border pair.[40] To do this, I keep only high schools that are located in a county that borders another DMA. Then for each high school, I determine the closest adjacent county in a different DMA. These two counties become a *candidate* border pair. I retain only those candidate pairs where at least one high school in each county belongs to the pair. This procedure yields approximately 80 border pairs.

3.2.2 Additional Specifications

To facilitate comparison of the estimates with effects in the literature, I also specify an elasticity model:

$$\log(y_{st}) = \log(a_{m(s)t})\alpha + \bar{z}_{st}\delta + \theta_{b(s)t} + \tau_s + \eta_{st}. \tag{2}$$

The unit of observation is a high school-cohort. On the left hand side, y_{st} indicates the share of high school graduates in high school s and cohort t that enroll in college within a year. Both the outcome and the advertising treatment are transformed using the natural log. The vector $\bar{z}_{st}$ contains the school-cohort means of individual covariates, and the same identifying fixed effects $\theta_{b(s)t}$ and τ_s are included. I include DMA-cohort advertising by out-of-state and online colleges as separate covariates.[41] With the log transformations of the outcome and treatment, the model has a constant elasticity interpretation.

In addition to the baseline model in Equation (1), I consider how advertising effects differ by college type. Letting w denote college type, $w \in \{$Public 4-Year, Community College, Private Non-Profit, For-Profit$\}$, I first specify a set of market expansion models where each model has two treatments: the advertising for the focal college type, $a^{w}_{m(i)t(i)}$, and the aggregate advertising of all other college types, $a^{-w}_{m(i)t(i)} \equiv \sum_{w' \neq w} a^{w'}_{m(i)t(i)}$; the latter includes advertising by out-of-state and online colleges. Specifically, for each w, I estimate

$$y_i = a^{w}_{m(i)t(i)}\beta^{w} + a^{-w}_{m(i)t(i)}\beta^{-w} + z_i\sigma_1 + x_{c(i)t(i)}\sigma_2 + \theta_{b(i)t(i)} + \tau_{s(i)} + \mu_i. \tag{3}$$

All border strategy specifications include fixed effects for the border pair-by-cohort $\theta_{b(i)t(i)}$ and high school τ_s and cluster standard errors at the DMA-by-cohort level.

[40] An alternative approach would be to define the set of border pairs as every pair of adjacent counties separated by a DMA boundary, and assign each high school to every border pair that contains its county. With this approach, however, some counties would appear in multiple border pairs. As described below, I use the high school distance to the boundary to define border pairs in order to avoid duplicating counties (and every observation within those counties).

[41] Although I do not focus on these effects, they are identified under the same assumptions as the effects for Texas colleges' advertising.

3.3 Variation in Advertising

The use of cross-DMA variation in the border strategy relies on advertisers not buying advertising uniformly across neighboring DMAs and over time.

To examine the temporal variation in spot TV advertising, I plot DMA-level spot TV ad volumes for each year in the sample. Because the amount of college spot TV advertising can vary substantially across the Texas DMAs, I categorize DMAs into one of three groups (high, medium, and low) based on the ad volumes and plot the groups separately. Figures 5, 6, and 7 report this time variation in DMA-level advertising. There does not appear to be a common time trend in advertising across DMAs. For example, in Figure 5, advertising volumes increase steadily in the Dallas-Fort Worth DMA from 2011 to 2015, while in the San Antonio DMA, advertising volumes are generally declining from 2012 to 2015.

I also examine how often colleges advertise in a single DMA versus multiple DMAs. Figure 8 plots the distribution of the number of DMAs in which a college advertises over the sample period. The analysis includes any brand that advertises at least once in any DMA during the period. The large spike at zero shows that many colleges do not advertise at all in one of the years, suggesting time variation in brand-level advertising. That behavior is more common among community colleges and regional institutions, while for-profit chains often have a presence in multiple DMAs and are more likely to advertise across Texas.

Last, I investigate the breakdown in enrollment and advertising by college type. Table 2 shows each college type's share out of the total for in-state enrollment, the number of TX colleges' spot TV ads, and the total number of college brands that advertise. While community colleges enroll 57 percent of in-state students, they account for 21 percent of the college brands and just 6 percent of the spot TV ads aired by Texas colleges in Texas DMAs. Similarly, public 4-year colleges capture one-third of in-state enrollment but account for a mere 2 percent of the spot TV ads. In contrast, for-profit colleges enroll the smallest share of students, at 3 percent, but they make up 38 percent of the advertisers and over 90 percent of the ads aired on spot TV.

4 Border Strategy Results

This section reports estimates of the causal effect of advertising on college-going for the models in Equations (1) - (3). I first discuss main results, where the treatment of interest is the total number of spot TV ads (in tens of thousands) aired by all Texas colleges in the student's DMA, and the outcome is an indicator of enrollment in (any) college within one academic year of high school graduation. As described in Section 3.2.2, I also estimate the treatment effect as an elasticity. I view the elasticity specifications as complementary to the baseline specification (in which the advertising treatment enters in levels), as the log-log model allows the effect of advertising to

Table 2: Enrollment and Advertising Shares by College Type

College Type	In-State Enrollment	College Brands	Spot TV Ads
Public 4-Year	0.34	0.24	0.02
Community College	0.57	0.21	0.06
Private Non-Profit	0.06	0.17	0.01
For-Profit	0.03	0.38	0.91
Total	1.00	1.00	1.00

Notes: Table 2 presents each college type's share of the total in-state enrollment, college brands (advertisers), and spot TV ads for the sample period. A college brand is any of the brands in the Ad Intel database that can be mapped to at least one of the Texas colleges in the ERC.

exhibit decreasing returns to scale.

For the full sample period (2011-2015 high school graduation cohorts), the average number of spot TV ads aired in a DMA for each cohort (which covers two academic years) is approximately 20,000. The median is substantially lower, at approximately 5,000. The college-going rate is 0.599 state-wide and 0.614 for the border-pair sample. The border pair sample covers 103 counties.

4.1 Extensive Margin Effects of College Advertising

This section reports results of the effect of advertising on college-going. The preferred specification, which uses the border strategy, is shown in the rightmost column (column 3) of Table 3. For comparison, I also report results from two alternative specifications. In column 1, I report estimates from a regression with two-way fixed effects for high school and cohort.[42] In contrast to the preferred specification, this model is estimated on the full sample of high school graduates from all Texas counties. In column 2, I report estimates from a model that uses a modified boundary discontinuity design; this model includes border pair-by-cohort fixed effects and a quadratic polynomial for the high school's distance to the border and is estimated using the sample of students in border pairs.[43] To provide a benchmark for the magnitude of the effect, it is helpful to note that 10,000 ads is the DMA average per year,[44] which is approximately 27 college spot TV ads per day.

[42]The estimating equation is $y_i = a_{m(i)t(i)}\beta^{FE} + z_i\lambda_1 + x_{c(i)t(i)}\lambda_2 + \tau_{s(i)} + \sigma_{t(i)} + \mu_i$. The identifying assumption is that, conditional on the covariates and fixed effects, any differences in college-going arise from differences in advertising.

[43]The estimating equation in this model is $y_i = a_{m(i)t(i)}\beta^{BDD} + f(d_{sk(b)}) + z_i\psi + \theta_{b(i)t(i)} + \nu_i$, where $d_{sk(b)}$ is the (shortest) distance from high school s to the DMA boundary $k(b)$ in its border pair. This model leverages within-year variation in advertising between two counties that share a DMA border. The identifying assumption is that $E[\nu_i|f(d_s), z_i\theta_{b(i)t(i)}]$ is continuous at the boundary $k(b)$.

[44]The DMA-year average for each cohort (a 2-year period) is shown in the bottom of Table 3.

Table 3: Effects of College Advertising Under Alternative Research Designs

Outcome: Enrollment at Any College	Two-Way Fixed Effects	Boundary Discontinuity	Border Strategy
Spot TV Ads by TX Colleges (10,000s)	0.0050	0.0033^{***}	0.0178^{***}
	(0.0034)	(0.0012)	(0.0043)
Percent Change	0.8	0.5^{***}	2.9^{***}
Average Dependent Variable	0.599	0.614	0.614
Sample	All counties	BP counties	BP counties
Fixed Effects	HS, Cohort	BP-by-Cohort	BP-by-Cohort, HS
DMA Average # Ads in t and $t+1$	$20,874$	$20,209$	$20,209$
Adjusted R-Squared	0.157	0.164	0.172
Observations	$1,345,885$	$208,410$	$208,405$

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Notes: Table 3 displays regression results of the effect of local TV advertising on the probability of enrolling in college. Observations are at the individual level. The sample includes Texas public high school graduates from academic years 2010-11 through 2014-2015. The outcome is a binary indicator of enrollment in any college within one year of the high school graduation year. The treatment of interest is advertising by Texas colleges, measured as the total number of local (DMA-specific) spot TV ads aired from August 1-July 31 in the academic year of high school graduation (t) and the subsequent year ($t+1$). Estimates in the table are scaled by 10,000 to be interpreted as the causal effect of an increase in 10,000 college ads on the probability of enrolling in college (in percentage points). Columns correspond to different model specifications. The preferred specification is in column 3, while columns 1 and 2 show the effects when key elements of the border strategy are removed. Column 1 reports estimates from a two-way fixed effects design with fixed effects for high school and cohort and is estimated using high school graduates in all counties. Column 2 reports estimates from a boundary discontinuity design which includes fixed effects for border pair-by-cohort and is estimated using the sample of high school graduates in border pairs. Column 3 reports estimates from the border strategy, which includes fixed effects for border county pair-by-cohort and high school and is estimated using high school graduates in border pairs. All models include individuals' demographic covariates. Standard errors are in parentheses and are clustered at the DMA-by-year level. *BP*= Border Pair, *HS*= High School, *DMA*= Designated Market Area (media market).

The results in Table 3 show that advertising increases the probability of enrolling in college. Coefficient estimates are positive across model specifications. The first column shows results from the two-way fixed effects specification. The parameter estimate is 0.0050 ($SE = 0.0034$), which corresponds to a 0.8 percent increase in college-going, but it is not statistically different from zero. Column 2 shows results from the border discontinuity design, which discards data from the interior counties and leverages within-year variation in advertising between two counties that share a DMA border.[45] The 0.0033 percentage point effect, equivalent to a 0.5 percent increase in enrollment, is smaller than in the two-way fixed effects specification, however it is estimated more precisely ($SE = 0.0012$) and is statistically significant at the one percent level.

I now turn to the results from the border strategy, which is my preferred causal research design. As with the boundary discontinuity model reported in column 2, the border strategy model in column 3 is estimated only using high school graduates in border pairs. Unlike the boundary discontinuity design, the border strategy does not assume that individuals in a border pair are randomly assigned to either side of the DMA boundary (which, as described in Section 3, is unlikely to be a valid assumption in my case because I cannot limit the analysis to a narrow bandwidth around the boundary). The border strategy is able to relax this assumption by adding high school fixed effects, which allow for time-invariant differences in the level of college-going on either side of the border.As shown in column 3, the border strategy yields a substantially larger estimated effect of advertising at 0.0178 percentage points ($SE = 0.0043$), equivalent to a 2.9 percent increase in college-going. Compared to the boundary discontinuity design, the addition of high school fixed effects in the border strategy increases the estimated coefficient five-fold, indicating that there are unobserved differences in college-going between individuals in border pairs that are correlated with advertising. Specifically, students with lower college-going rates appear to live in DMAs that receive more college advertising, so when high school fixed effects account for time-invariant differences in college-going, the effect of advertising is larger. The estimate is statistically significant at the one percent level.

To summarize, results from the border strategy indicate that increasing colleges' DMA-year spot TV ad volumes by 27 ads per day (10,000 ads per year) increases the probability of enrolling in college by 1.8 percentage points, or approximately 3 percent from the baseline college-going rate of 0.614. Colleges spend between \$175 and \$200 per spot on average, which is approximately \$1.75 to \$2 million dollars in spot TV ad spending per DMA-year, on average.

[45] The panel variation is not used in this model, so identification relies on the assumption that, conditional on the covariates, individuals in each border pair are as good as randomly assigned to either side of the border.

4.1.1 Differential Responses by Demographic Group

Which students are more responsive to college ads? In this section I investigate whether advertising generates differential effects on college-going for different demographic groups.

In these analyses, I estimate the baseline (border strategy) model separately for six different groups. To facilitate comparisons across groups, I focus on the elasticity estimates.[46]

Anecdotal evidence suggests that many colleges, especially for-profit colleges, target ads to students of color, so I investigate how the effect of advertising on college-going differs across racial groups (white, underrepresented minority, Hispanic, and Black).[47] In addition, advertising has the potential to reduce gaps in awareness or information between advantaged and disadvantaged students, so I also examine effects for low-income students (those eligible for free school lunches or whose family receives social assistance such as TANF or SNAP).

The estimated elasticities are shown in Table 4. Each column reports results from the elasticity regression in (2) restricted to the demographic group indicated in the column. The estimates, which are statistically significant for nearly all subgroups,[48] reveal meaningful differences in responsiveness to college ads across demographic groups. The leftmost column reports the overall elasticity (all students) as 0.0357. White students are somewhat more responsive than the average student, with an elasticity of 0.0383, but Hispanic students are the most responsive among the racial subgroups, with an estimated elasticity of 0.0513. The rightmost column reveals that females are somewhat less responsive than the average student, with an elasticity of 0.0279. Among all groups, low-income students have the highest advertising elasticity of demand, at 0.1305. Low-income students are less likely to have a parent with a college degree, so their post-secondary choices may be more influenced by college advertising.

Having discussed the effect of advertising on the probability of enrolling in a Texas college, I now turn to examining heterogeneous effects by college type.

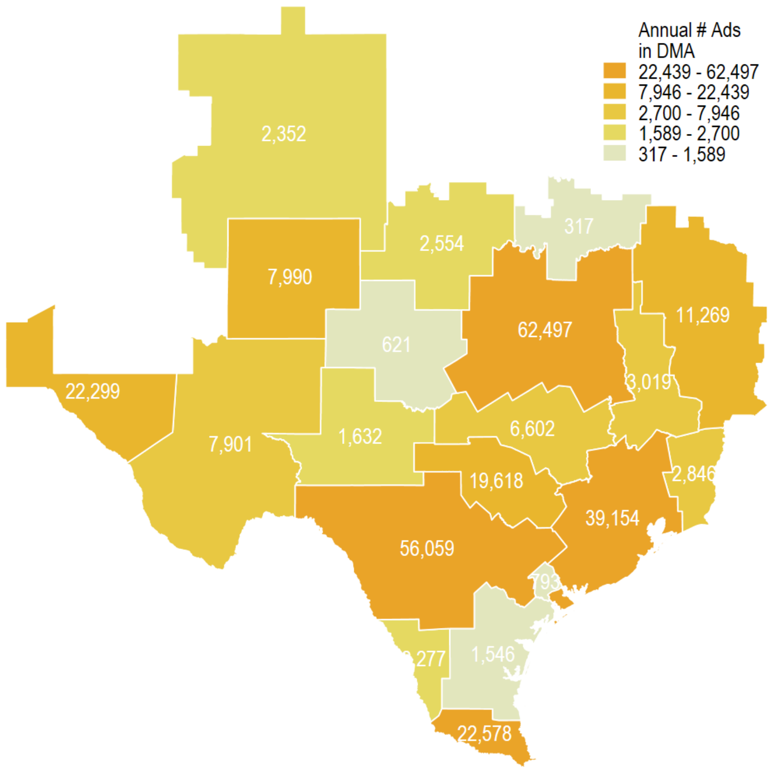
[46] Estimates in levels corresponding to Equation (1) are reported in Table 11 in the Appendix.

[47] Underrepresented minority includes students who identify as Hispanic, Black, or American Indian/Alaska Native.

[48] The exception is the model estimated on the sample of Black students. The lack of precision is likely due to the small sample size (22,987).

Figure 3: Advertising Volumes in Texas DMAs

(a) 2010



(b) 2015

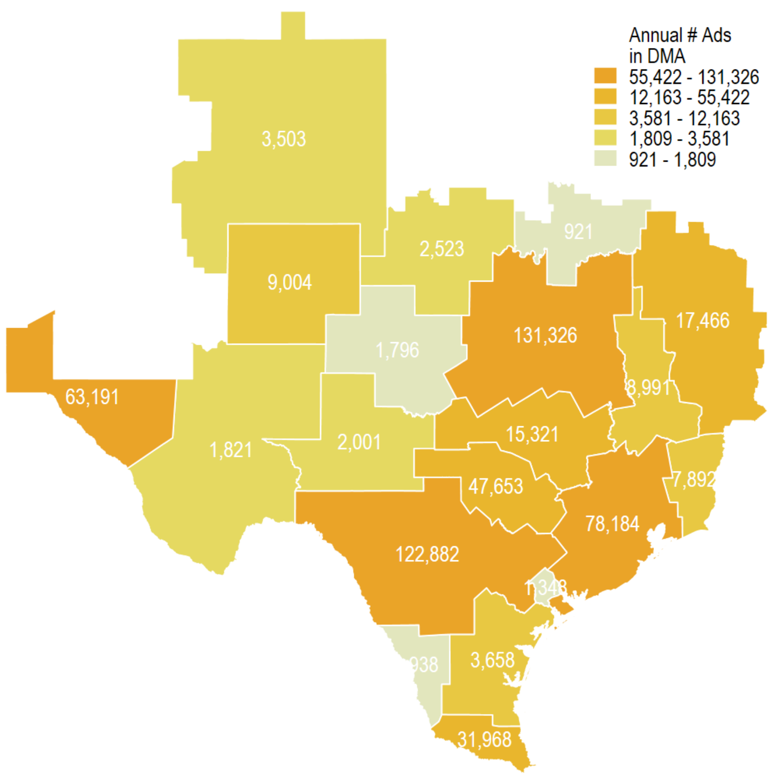


Figure 4: Interior and Border County High Schools in Texas

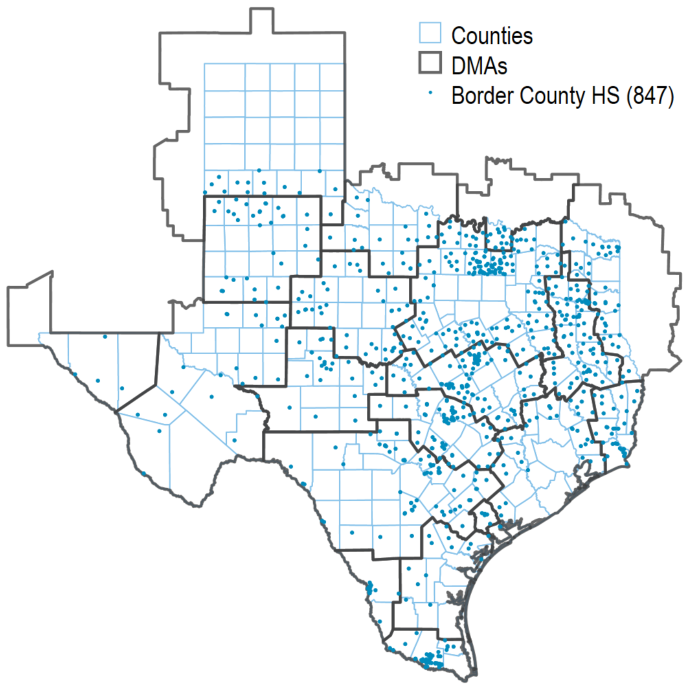


Figure 5: Advertising Variation Over Time: High Advertising DMAs

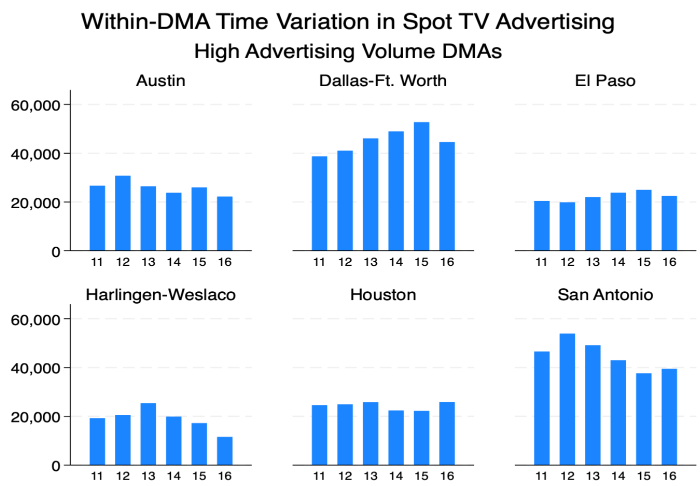


Figure 6: Advertising Variation Over Time: Medium Advertising DMAs

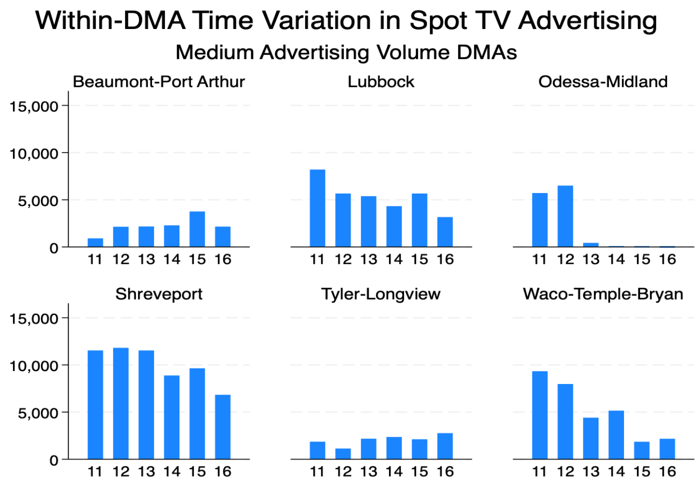
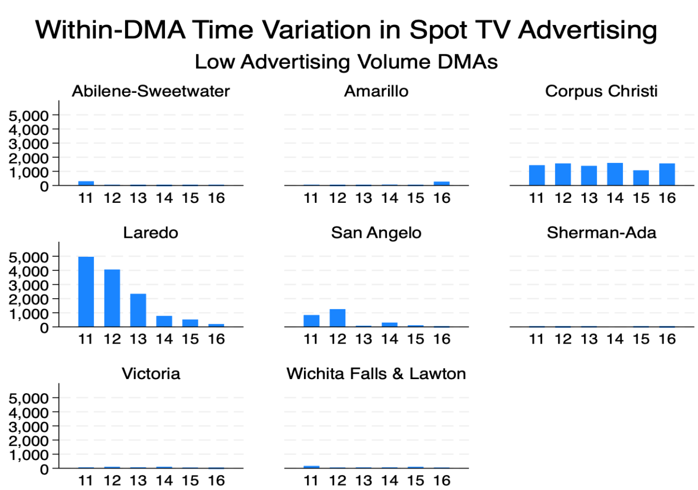


Figure 7: Advertising Variation Over Time: Low Advertising DMAs



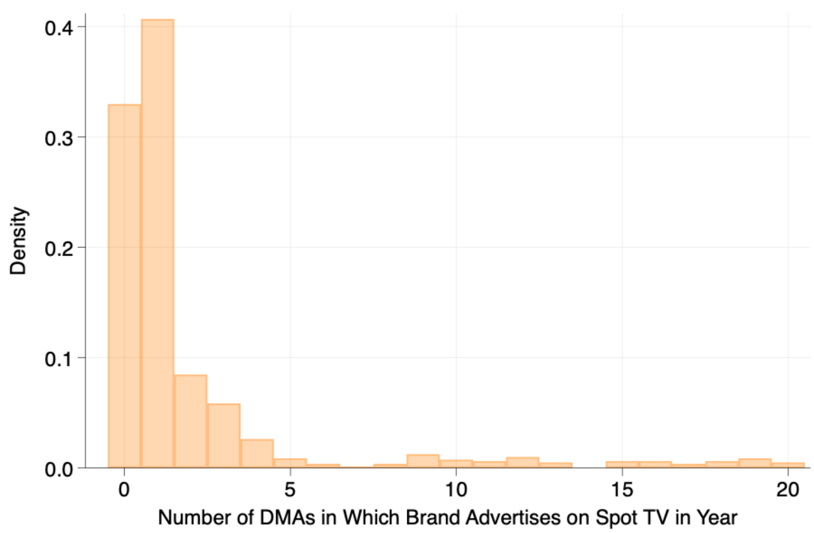


Figure 8: Texas Colleges (Brands) that Ever Advertise in Texas 2010-2015

Table 4: Advertising Elasticities of Demand by Demographic Group

Outcome: Log(Share Enrolled in College)	All	White	URM	Hispanic	Black	Low-Income	Female
Log(TX College Ads)	0.0357***	0.0383***	0.0375**	0.0513***	0.0165	0.1305***	0.0279***
	(0.0082)	(0.0115)	(0.0155)	(0.0162)	(0.0500)	(0.0270)	(0.0095)
Average Dependent Variable	0.556	0.606	0.526	0.515	0.625	0.455	0.625
Mean # Ads (DMA-Year)	19,727	19,044	19,727	19,727	20,667	19,727	19,727
Median # Ads (DMA-Year)	4,875	4,465	4,875	4,875	4,875	4,875	4,875
Adjusted R-Squared	0.897	0.865	0.862	0.871	0.761	0.821	0.848
Adjusted R-Squared (within)	0.095	0.108	0.075	0.052	0.122	0.129	0.078
Border-County Pairs	79	78	79	79	59	79	79
Counties	103	102	103	102	71	102	103
Years	5	5	5	5	5	5	5
Observations	206,103	106,636	89,228	61,620	22,987	61,375	101,987

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Notes: Table 4 displays regression results from the elasticity models presented in Section 3. Observations are at the high school-by-year level. For each column, the sample is restricted to the students belonging to the subgroup indicated in the column. Each column is estimated using Equation (1); all specifications include fixed effects for the border pair-by-cohort and high school. The outcome is the log of the share of each high school's graduates that enroll in college within one year after the year of high school graduation. The treatment is the log of advertising by Texas colleges, measured as the total number of local (DMA-specific) spot TV ads aired from August 1-July 31 in the academic year of high school graduation (t) and the subsequent year ($t+1$). Estimates in the table are elasticities, interpreted as the percent increase in demand for college from a 1 percent increase in college ads. All models include high school-by-year averages of a subset of the individual-level demographic covariates. Standard errors are in parentheses and are clustered at the DMA-by-year level.

Table 5: Effects on College-Going by Advertiser Type

Outcome: *College Enrollment*	*Extensive Margin Effect of 1,000 Spot TV Ads by* Public 4-Years	Community Colleges	Private Non-Profits	For-Profits
Percentage Points	–0.0080	0.0089***	–0.0047	0.0020***
	(0.0061)	(0.0025)	(0.0034)	(0.0004)
Percent	–1.31	1.46***	–0.779	0.329***
Average Dependent Variable	0.612	0.612	0.612	0.612
Observations	207,934	207,934	207,934	207,934

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Notes: Table 5 displays regression results of the effect of local TV advertising on the probability of enrolling in college. Observations are at the individual level. Each column is a separate regression model that decomposes the aggregate advertising treatment into two variables: advertising by focal college type (indicated in the column) and advertising by all other college types. The outcome is a binary indicator of enrollment in any Texas college within one year of the high school graduation year. All specifications include fixed effects for the border pair-by-cohort and high school. Standard errors (in parentheses) are clustered at the DMA-by-year level.

4.1.2 Differential Effects by Advertiser's College Type

In this section, I ask: Does the effect of advertising on overall college-going vary by the type of college that advertises? I report estimates from the models given by Equation (3), where I disaggregate the advertising treatment by college type. Results are presented in Table 5. Note that, in contrast to the main effects, these coefficients are scaled by 1,000 ads. The estimated effects for public 4-year colleges and private non-provide colleges are negative but are not statistically different from zero. Community college advertising has a relatively large effect on overall college-going. The estimated coefficient of 0.0089 translates to a 1.46 percent increase in college enrollment when community college advertising is increased by 1,000 local spots, holding constant advertising by all other college types. This effect is statistically significant at the one percent level.

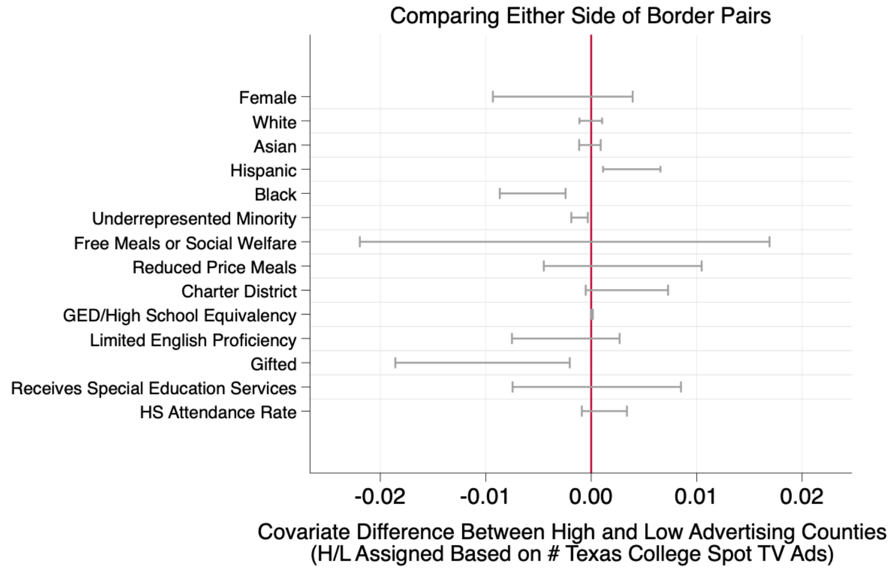
Advertising by for-profit colleges also contributes to market expansion. For the same increase in ad spots (1,000), the effect of for-profit college advertising (0.0020) is substantially smaller in magnitude than the effect of community college advertising. However, it is worth noting that for-profit colleges purchase fifteen times more spot TV ads than community colleges (see Table 2), so if advertising operates through an awareness channel, marginal increases in for-profit ad spots would be expected to generate a smaller effect.

4.2 Threats to Identification

In this section, I consider the primary threats to the internal validity of the border strategy results.

A key assumption in my strategy is that exposure to local TV advertising is uncorrelated

Figure 9: Students in Border Pair Counties Have Similar Characteristics



with the unobserved factors that affect individuals' college-going behaviors, conditional on the covariates and fixed effects. One way to shed light on the validity of this assumption is to examine whether observed student characteristics appear to be similar between the two groups being compared.

I conduct such an exercise using the baseline analytic sample at the student level. For a set of binary student covariates, I estimate a regression of the covariate on (i) an indicator for whether the student's county has higher numbers of spot TV ads within the border pair, (ii) the other covariates, and (iii) fixed effects for border pair-by-cohort and county. The model is

$$z_i^k = High_{b(i)t(i)}\alpha + z_i^{-k}\phi + \theta_{bt} + \tau_c + \xi_i, \text{ where } -k = \{k' \in K \text{ s.t. } k' \neq k\}, \tag{4}$$

where i indexes students, z denotes student demographic covariates (indexed by k), and $High$ is a dummy variable indicating that more spot TV ads were aired in i's county than in the neighboring county in the border pair for cohort t. Figure 9 plots the coefficients and 95 percent confidence intervals from each regression, which are labeled by the outcome covariate. The coefficients capture the mean differences (in percentage points) of binary student demographic covariates between the two counties in a border-pair. As seen in the figure, most of the confidence intervals include zero, and all of the point estimates are small in magnitude (less than 3 percentage points), indicating that within a pair of border counties, being in the high ad volume county has little to no effect on student demographic characteristics.

Another concern is that individuals in border counties view TV ads aired in the neighboring DMA. In this case, the advertising measure for their DMA does not accurately capture their potential exposure to college TV ads. I refer to this scenario as "cross-DMA-viewing." Cross-DMA-viewing could affect my results if it is common or systematic.

There is little reason to expect cross-DMA viewing to be common in my context. The most probable scenario for cross-DMA-viewing is when households receive broadcast signals from local TV stations in neighboring counties outside their DMA. Since broadcast stations are more likely to be located in urban areas, border county households are more likely to be able to pick up signals from multiple DMAs. The existence and signal strength of such out-of-DMA channels would vary across households based on the household's distance to the local TV station. The ideal reception for digital broadcast TV is within approximately 35 miles of the TV station's broadcast tower. Although it is possible to receive a TV signal further away, the curvature of the Earth weakens over-the-air reception, such that beyond approximately 70 miles, the signal becomes unavailable.

Although such cross-DMA viewing is possible, it is likely to be limited in my setting, as only 10 percent of TV homes are "broadcast only" during my period of study.[49] The remaining 90 percent of households obtain their TV content through a paid subscription from a cable, satellite, telecommunications, or other multichannel video programming distributor (MVPD), which is regulated by the FCC to deliver DMA-specific content based on the customer's location.[50] Households in border counties are more likely to use a paid subscription service because those locations are the least likely to catch a strong enough broadcast signal to exclusively rely on over-the-air television delivery. Given the overwhelming penetration of pay TV, the geological constraints on tuning into a TV station outside of one's DMA, and the small representation of those channels among all potential channels to be watched, there seems to be limited scope for cross-DMA-viewing using broadcast signals from neighboring DMAs.

Of course, other situations could generate cross-DMA-viewing. Even if cross-DMA-viewing is common, it is likely to occur randomly.[51] In this case, it would generate classical measurement error, causing attenuation bias in the estimated advertising parameter. Recall that I estimate the regressions restricting the sample to students whose high schools are between 10 and 30 miles from the DMA boundary. Excluding households 10 miles from either side of the DMA border reduces the likelihood that neighboring DMA broadcast signals can be received with sufficient strength; it also reduces the likelihood that a student's social network extends to the neighboring

[49] Specifically, in December 2013, 10 percent of TV homes in the U.S. were "broadcast only;" (The Nielsen Company, 2013). The share of broadcast only households differed slightly by race: white 9%, Black/African American 12%, Hispanic 15%, Asian American 13%.

[50] Although in theory it is possible for these homes to use a digital antenna to tune into available local broadcast stations, such behavior seems unlikely.

[51] It is difficult imagine how or why students would systematically shift their TV viewing to a neighboring DMA.

DMA, leading her to spend time watching TV in the neighboring DMA.

4.3 Limitations of the Border Strategy

A discussion of the external validity of the effect estimates under the border design is warranted. As with any causal identification strategy that leverages exogeneity generated by discontinuities, there is an internal-external validity trade off: estimated effects are "local" to areas near the discontinuity–in this case, the border counties. The primary reason why exogeneity is a reasonable assumption in the border strategy is that the border counties are likely quite different from the interior counties. It is precisely because college demand and unobserved demand shocks are likely to differ between border and interior counties that we would expect the cross-border differences in advertising over time to be uncorrelated with cross-border differences in college demand shocks in the border counties. That such an assumption is unlikely to hold for the interior counties (because college demand shocks in the urban areas are likely to drive the DMA-level college advertising decisions) demonstrates the appeal of the "local" analysis. Nevertheless, for many policy-relevant questions related to advertising, it is desirable to generalize the estimated effects to the broader population. Although it is difficult to know the extent to which such extrapolation is reasonable, a useful exercise is to examine differences in demographics between border and interior counties.

In Figure 16, I use the full population of high school graduates in Texas to examine raw differences in student and school characteristics. Figure 16a plots coefficients and 95 percent confidence intervals, each of which comes from a regression of the demographic attribute (vertical axis) on indicators for border county and cohort year, with standard errors clustered by school of high school graduation. The coefficients represent percentage point differences in student attributes between border and interior counties.

Figure 16a shows that certain characteristics can be quite different between border and interior counties. Students in border counties are about 6 percentage points less likely to have a family income that qualifies them for free meals through the National School Lunch Program, 6 percentage points less likely to be Hispanic or Latino, 5 percentage points less likely to have been considered at risk of dropping out of high school, and 4 percentage points less likely to be black. On the other hand, they are nearly 13 percentage points more likely to be white. The remaining differences between border and interior county students (receipt of special education services or having a parent that is a migratory agricultural worker) are small ($<$ 1 percentage point). I conduct a similar comparison using school characteristics in Figure 16b. As expected, the differences between border and interior counties involve the location of the schools (namely, the metropolitan area categorization of the county in which the school is located). Attributes such as charter type do not differ between border and interior counties. In summary, one should exercise caution in generalizing the results to the interior counties.

Notably, individuals living in rural areas have fewer college options nearby. Hillman (2016) shows that living in an "education desert"—a place that is geographically isolated from colleges—is an important barrier to college-going. With the border design, I am measuring the effect of advertising on college-going among a more geographically isolated sub-population, so to the extent that policymakers are interested in increasing college-going among this sub-population, results from my sub-population (rather than the broader population) may be an effect of interest.

5 A Model of College Choice with Advertising

Results from the design-based analyses show that college advertising increases college-going among recent high school graduates. I estimate an advertising elasticity of demand equal to 0.038, which is more than double the median elasticity estimated in Shapiro et al. (2021), suggesting that students are relatively responsive to college ads.[52] I also find evidence that enrollment effects of college advertising vary across demographic groups. Low-income students are particularly responsive to advertising, with elasticity estimates exceeding the 90th percentile in Shapiro et al. (2021).

In this section, I develop a structural model to simulate the effects of a ban on college advertising, which recently was proposed in Congress. I model the college enrollment decision of individuals as a discrete choice among differentiated products (McFadden, 1973) and draw upon recent demand models in advertising (Dubois et al., 2018; Shapiro, 2018; Sinkinson and Starc, 2019; Tuchman, 2019) and higher education (Howell, 2010; Kapor, 2020; Cook, 2021; Bleemer, 2021) to estimate a relatively flexible multinomial choice model of college demand with advertising.

After describing the model, I briefly discuss the estimated parameters. I then use the model to perform the policy counterfactual of interest: examining how college enrollment would change if college advertising were banned. To this end, I simulate college choices in the absence of advertising and analyze the implied changes in college enrollment. The counterfactual exercise complements the border strategy analyses: while the design-based approach enabled me to quantify the aggregate/net effects of college advertising, the discrete choice model makes it possible to analyze the relative importance of different response margins.

[52] Shapiro et al. (2021) estimate the causal effect of TV advertising on sales for a variety of consumer products and provide summaries of the distribution of elasticities across the 288 brands. For the specification most comparable to mine, which defines advertising as occurrence stock and uses the border strategy, the median elasticity is 0.0160, the mean is 0.0327, and the 75th and 90th percentiles are 0.0516 and 0.1270, respectively.

5.1 Setup

5.1.1 Environment

As before, individuals are indexed by $i = 1, ..., N$, and their county of residence is denoted $c(i)$. After graduating from high school, individuals choose one of $j = 1, ..., J$ mutually exclusive and exhaustive college alternatives, or the outside option of not attending college, which is denoted $j = 0$. The college choice set is the same for all individuals and includes every Texas college and two other alternatives: online college or out-of-state college.[53] Colleges may purchase local TV ads to air in any of Texas' 20 DMA, which as before are denoted by m.[54]

5.1.2 Colleges

Because my aim is to study the demand responses to college advertising rather than colleges' behavior, colleges are passive in the model. I abstract from the admissions process and assume that colleges have neither capacity constraints nor preferences to enroll students with certain attributes (e.g., higher ability). These assumptions are reasonable for the majority of Texas colleges, which are open-access or have limited admissions requirements. Moreover, Texas high school students who graduate in the top 10 percent of their class qualify for automatic admission to most public universities in Texas (for The University of Texas at Austin, the rule is top 6 percent). Given this legislative rule, the admissions process to Texas' more selective public colleges is trivial for high academic achievers, as is the admissions decision from the college's perspective.

5.1.3 Preferences

Agent i obtains indirect utility U_{ij} from enrolling in college j, which is a function of college advertising, distance to j, other non-advertising characteristics of j (both observed and unobserved by the researcher), observed demographic attributes of i, and an error. Observed demographic attributes are denoted by the K-dimensional row vector z_i. The distance to college j for individual i is denoted d_{ij}.[55] Individual i is (potentially) exposed to TV advertising by alternative j in the media market where she lives; denote the number of ads by j in her media market $a_{jm(i)}$. The other observed characteristics of alternative j are denoted by the R-dimensional vector x_j.

[53] I include in the sample the high school graduates that choose to attend a college outside of Texas. Approximately 4.5 percent of my sample (7.25 percent of college enrollees) attends an out-of-state college within the first year after high school graduation. Aside from a handful of institutions (Rice University, Baylor University, Southern Methodist University, Texas Christian University, and The University of Texas at Austin), Texas colleges do not compete for students in the national market.

[54] Recall that DMAs are collections of counties, so all counties in m receive the same local TV advertising.

[55] I construct d_{ij} as the distance to college j from i's high school.

I assume that utility takes the general form

$$U_{ij} = V(a_{jm}, d_{ij}, x_j, z_i, \beta_i) + \delta_j + \varepsilon_{ij}, \tag{5}$$

where $V(a_{jm}, d_{ij}, z_i, x_j, \beta_i) := V_{ij}$ is a function of the tastes of i, represented by β_i, and the observed variables. Utility includes two unobserved terms. The δ_j term captures unobserved tastes for college j that are common across individuals, while ε_{ij} is an i.i.d. idiosyncratic error term with joint density $f(\varepsilon_i)$, where $\varepsilon_i' = \langle \varepsilon_{i0}, \ldots, \varepsilon_{iJ} \rangle$.

5.1.4 Choice

Agent i chooses the option j that gives her the highest utility; that is, she solves

$$\max_{j \in J} U_{ij}.$$

Let the vector $y_i = \langle y_{i0}, \ldots, y_{iJ} \rangle$ denote i's choice for each j, where

$$y_{ij} = \begin{cases} 1 & \text{if } U_{ij} > U_{ij'} \quad \forall\, j' \neq j \\ 0 & \text{otherwise.} \end{cases} \tag{6}$$

Then, given the observed data and parameters, the conditional probability that individual i chooses alternative j, $Pr(y_{ij} = 1 | a, x, z, d, \beta_i) = P_{ij}$, is

$$P_{ij} = Pr(U_{ij} > U_{ij'} \quad \forall j' \neq j). \tag{7}$$

5.2 Discussion: Modeling Advertising

Before describing the parametric specification of the model, I provide some motivation for my modeling decisions. In particular, I consider three key questions that seem most relevant to characterize how advertising affects college choice: (i) Does advertising affect awareness? (ii) Does advertising affect utility directly? (iii) Does advertising by other colleges $(-j)$ enter i's utility for college j?

5.2.1 Advertising and Awareness

The first question is whether to model college advertising by j as affecting individuals' *awareness* of j.[56] Goeree (2008) considers the awareness effects of advertising in the context of personal computers, a market in which the full information assumption of traditional discrete-choice models

[56]It also is possible that advertising by j increases awareness of other colleges, a point which I consider in the discussion for question (iii).

is unrealistic because new products are introduced frequently. In the college choice context, advertising also may affect awareness, and, consequently, the set of colleges a student considers. Some colleges may advertise because they have low brand awareness or want to increase their brand recognition. Colleges in thick markets may advertise to make their existence *salient*.[57] The awareness channel of college advertising may be particularly relevant for individuals from low-income backgrounds or who would be the first in their families to attend college ("first-gen" students).

In a context where advertising affects awareness of colleges, the choice probability would take the general form

$$\begin{aligned} Pr(i \text{ chooses } j) &= Pr(i \text{ chooses } j | i \text{ is aware of } j) \times Pr(i \text{ is aware of } j) \\ &= Pr(U_{ij} > U_{ij'} \quad \forall j' \neq j | j \in C_i) \times Pr(j \in C_i), \end{aligned} \tag{8}$$

where C_i indicates i's consideration set. Comparing (8) to my specification of the choice probability in Equation 11 shows that I assume $Pr(i \text{ is aware of } j) = 1$, so individuals are aware of all J alternatives. I omit an explicit treatment of advertising's effects on consideration because measures of individuals' baseline awareness of each college are unavailable; absent such data, identification of these effects would rely exclusively on functional form.

5.2.2 Advertising and Utility

The second question is whether advertising affects utility directly. The advertising literature offers different views of advertising, broadly categorized as the complementary, informative, and persuasive views, each of which offers a different characterization of the relationship between advertising and utility (see Bagwell, 2007, for a comprehensive review). Within the complementary view, Becker and Murphy (1993) model advertising as entering the fixed preferences of consumers in a way that complements consumption of the advertised product. For products that confer social status, such as luxury goods, it seems reasonable that advertising could directly increase the utility of consuming the good. In the college choice context, one might anticipate that advertising by elite or highly selective colleges could generate a similar "social prestige" effect. In Texas, however, such institutions might account for a handful of the state's colleges; for the typical broad-access

[57] An increase in awareness (due to advertising) can be interpreted in several ways. As described, individuals could be completely unaware of the existence of a product j (e.g., because it is new, or because she has not been exposed to it), and an advertisement could make her aware of its existence. Yet awareness effects also could reflect a situation wherein an individual had learned of the existence of j in a prior time period, but at the time of decision-making, j was not *salient* to her, and she could not recall it without an external reminder such as an ad. Thus, an ad for j could have a salience effect. Distinguishing between an awareness effect and a salience effect would be difficult without additional data on the individual's information set over time.

college, it seems less likely that individuals obtain utility from the ads themselves.

Under the informative view, advertising may affect (expected) utility directly if it contains information about the product. Anecdotal evidence suggests that at least some college advertising provides information about verifiable product attributes (such as offering scholarships or weekend classes). For consumers that do not know the value of an observable attribute for college j because they have not exerted the cost to learn it (for instance, by searching j's website or taking a campus tour), informative advertising can improve their knowledge of the attribute's true distribution (e.g., reveal the mean), reduce their uncertainty about the distribution, or reveal the exact value of the attribute.

Persuasive advertising, on the other hand, *changes* the utility function by "distort[ing] the consumer's decisions as compared to those that reflect his 'true' preferences" (Bagwell, 2007). Institutional aspects of the college market provide an environment for persuasive advertising to arise. Given that some college attributes are costly for consumers to observe or remain private information to the college, the information environment may incentivize persuasive advertising. Consider, for example, college quality, which is difficult to define and estimate (Dillon and Smith, 2017) but also is likely valued by a majority of prospective college students. Such information asymmetries may incentivize colleges to use ads that contain hard-to-verify claims or that distort the perceived importance or effect of an attribute, either of which can shift an individuals' beliefs away from the full-information benchmark.

Although both informative and persuasive advertising likely arise in the college choice context, to model each channel in consumers' preferences and estimate the corresponding parameters would require data that do not exist in my context.[58] As such, I abstract from an explicit treatment of advertising's informative and persuasive channels and model advertising in a reduced-form way, allowing advertising to enter utility directly. My baseline model therefore combines the awareness, informative and persuasive effects of advertising into a single parameter.

5.2.3 Rival Advertising

The third question is whether to include advertising by other colleges in i's utility function for college j. To understand the implications of this choice, it is helpful to consider a model that only includes a college's own advertising. A reasonable hypothesis is that advertising by j has a positive effect on its own enrollment (otherwise, it would be sub-optimal to continue to advertise). In a traditional discrete choice framework, if the effect of own advertising on own demand is positive, then advertising by j necessarily has a negative effect on enrollment at college j'. In this sense, a model

[58] For example, data on individuals' beliefs about product attributes.

that only includes own advertising imposes that advertising is exclusively *business-stealing*.[59] Is this a reasonable assumption? In the context of college choice, advertising could generate positive spillovers if it increases the salience or attractiveness of higher education. For example, viewing a college ad by j may cause an individual to consider the general value of higher education and prompt her to conduct a search of other colleges. If she enrolls in a different college j', then the advertising by j will have generated a positive enrollment spillover to j'. Understanding the strategic incentives colleges face when deciding whether and how much to advertise requires uncovering the relative magnitudes of business-stealing and positive spillover effects. The potential importance of positive spillover effects—which would disincentivize advertising due to free-riding—suggests that the ideal model would be flexible enough to allow i's choice of j to be affected by other colleges' $(-j)$ advertising. The current version of the model, however, does not include a separate term for rival advertising.[60]

5.3 Parameterization

I specify conditional indirect utility U_{ij} as

$$U_{ij} = \beta_i^o a_{jm}^o + \beta_i^d d_{ij} + \beta_i^{x\prime} x_j + \delta_j + \varepsilon_{ij}. \tag{9}$$

Both advertising and distance vary across colleges and individuals: Advertising (a_{jm}^o) varies across colleges and DMAs and distance varies across colleges and high schools. The other non-advertising college characteristics vary across colleges. The i subscripts on the βs indicate that the marginal utilities of each covariate are consumer-specific. Specifically, I take advantage of the detailed information on high school students' demographic and academic characteristics and allow the effect

[59] To understand why business-stealing effects of advertising are imposed by the structure of a traditional random utility discrete choice modeling framework, consider an increase in the probability of choosing j—arising, for example, from improvements in a utility-relevant attribute such as college quality. Because market shares must sum to one, this increase necessarily lowers the probability that one of the other alternatives is chosen. As such, including advertising by j in the utility function automatically generates a business-stealing effect: increases in j's advertising reduces enrollment at other alternatives (assuming that advertising has a positive effect on utility).

[60] I am in the process of adding rival advertising to the estimation procedure. In this case, college advertising will comprise j's own advertising (denoted with "o" superscripts) and rival advertising by other colleges (denoted with "r" superscripts): $a_{jm} = \langle a_{jm}^o, a_{jm}^r \rangle'$; a_{jm}^o represents the number of TV ads by college j in the media market where i resides (i.e., the ads to which agent i is potentially exposed), while a_{jm}^r represents advertising by a subset of other colleges G_j. I will include a college in j's rival group if it is located in the same county as j.

of advertising, distance, and college characteristics to vary with observed student attributes z_{ik}:

$$\begin{aligned}
\beta_i^o &= \bar{\beta}^o + \sum_{k \in K} z_{ik} \alpha_k^o \\
\beta_i^d &= \bar{\beta}^d + \sum_{k \in K} z_{ik} \alpha_k^d \qquad (10) \\
\beta_i^x &= \sum_{k \in K} z_{ik} \alpha_k^x .
\end{aligned}$$

The coefficients on advertising a_{jm} and d_{ij} are each the sum of a common component and a set of heterogeneous components. The $\bar{\beta} = \langle \bar{\beta}^o, \bar{\beta}^d \rangle'$ parameters are constant over consumers and are interpreted as the average effect in the population, while the heterogeneous components ($\alpha = \langle \alpha^o, \alpha^x, \alpha^d \rangle$) capture individual variation around those means. The x_j are college characteristics.

To understand the interpretation of the heterogeneous terms, consider the first term in (9): $\beta_i^o a_{jm}^o = \bar{\beta}^o a_{jm}^o + (\sum_{k \in K} z_{ik}^1 \alpha_k^o) a_{jm}^o$. The summation term of $\beta_i^o a_{jm}^o$, $(\sum_k z_{ik} \alpha_k^o) a_{jm}^o$, captures the portion of i's utility (above and beyond $\bar{\beta}^o a_{jm}^o$) from alternative j's own advertising in m; put differently, the marginal utility of advertising varies with observed consumer demographic attributes z_{ik}. The flexible specification captures the extent to which high school graduates' exposure or responsiveness to ads by j varies with their observed attributes.

The conditional probability that individual i chooses alternative j, $Pr(y_{ij} = 1 | a, x, z, d, \beta_i) = P_{ij}$, is

$$\begin{aligned}
P_{ij} &= Pr(U_{ij} > U_{ij'} \quad \forall j' \neq j) \qquad (11) \\
&= \int_{\varepsilon} I\left[\varepsilon_{ij'} - \varepsilon_{ij} < (V_{ij} + \delta_j) - (V_{ij'} + \delta_{j'}) \quad \forall j' \neq j\right] f(\varepsilon_i) d\varepsilon_i,
\end{aligned}$$

where $\varepsilon_i' = \langle \varepsilon_{i0}, \ldots, \varepsilon_{iJ} \rangle$ and I is an indicator function which equals one if the term in brackets is true and zero otherwise.

To close the choice model, I assume that the ε_{ijt} are independently and identically distributed Type 1 Extreme Value,[61] which simplifies to the well-known closed-form choice probability

$$P_{ij} = \frac{\exp(V_{ij} + \delta_j)}{\sum_{j'=0}^{J} \exp(V_{ij'} + \delta_{j'})}. \qquad (12)$$

Before discussing identification, I highlight several key differences between the border strategy model and the structural model.

[61] The main content of this assumption is that the ε_{ijt} are uncorrelated over alternatives and have the same variance.

Because my objective is to assess the enrollment effects of a ban on college advertising, it makes less sense to use only individuals in border pairs. For one, I would not be able to estimate taste parameters for all Texas colleges, which requires the broader sample of students. A related issue is that using only border county students would bias the heterogeneous preference parameters, as there are important demographic differences between border and interior students. The benefits of using students statewide are not without cost. The main trade off is that there is not a natural (and computationally feasible) way to implement the border strategy when students on the interior are included in the estimation sample.[62] For this reason, the structural model relies on different identifying variation than the border strategy, which I discuss in further detail below. In addition, the structural model is not indexed by time, which implicitly assumes away cohort effects. This is in contrast to the models estimated using the border strategy, which explicitly allow each cohort to have a different college-going rate.[63]

5.4 Identification

In this section I discuss how the model parameters are identified. The aim of the model is to study how advertising affects college choices and conduct policy counterfactuals in which college advertising is removed. Given this, the key parameters are the preferences for college advertising.

The first challenge in identifying the marginal utility of advertising is that colleges with higher advertising may also possess other attributes that increase utility, such as unobserved quality. The college-specific terms δ_j help address this source of endogeneity. Even controlling for common tastes, the cross-DMA differences in advertising for a college are unlikely to be random, as colleges likely target their advertising to the populations that will respond the most to the ads; these populations are likely to be geographically closer to the college and located in the same DMA as the college. To address this concern, I again exploit the discontinuity in spot TV advertising at media market borders, however my approach differs from the border strategy employed above.

The identifying assumption is that the number of spot TV ads by college j that student i is (potentially) exposed to is independent of the unobserved utility error term ε_{ij}, conditional on δ_j, distance to j (d_{ij}), and the interaction between student characteristics z_{ik} and d_{ij}. Under this assumption, the differences in utility (and choices) between two students who live equidistant from

[62]Enforcing comparisons in the border would involve adding a fixed effect for each county and college, i.e., replace δ_j with δ_{jc}. For 2013 alone, this would translate to 254 (counties)$\times$ 270 (colleges) = 68,580 δ_{jc} fixed effects.

[63]I made the decision to use a single year of data because of limited computational resources available on the shared server used for the analysis. I am in the process of adjusting the model in order to use all years in the sample period.

college j and who share the same demographic characteristics z_k arise exclusively from the exogenous jump in spot TV advertising that occurs because of FCC regulations and college advertising choices.

The identification of mean tastes for colleges (δ_j) follows standard arguments in the literature (Berry, 1994; Berry et al., 1995, 2004; Bayer et al., 2007). Specifically, these parameters are identified from the shares of high school graduates that choose each college, such that the model's predicted enrollment shares for college j equal the enrollment shares observed in the data.

Finally, because only differences in the alternative-specific constants δ_j are identified,[64] I set the overall level of utility by normalizing the constant for the outside option, δ_0, to zero.[65]

With this notation and normalization, the choice probability is

$$P_{ij} = \frac{\exp(V_{ij} + \delta_j)}{1 + \sum_{j'=1}^{J} \exp(V_{ij'} + \delta_{j'})}. \tag{13}$$

5.5 Estimation

I estimate the utility parameters using maximum likelihood. My estimation sample includes 2013 high school graduates for the entire state.[66] Let $\theta = \langle \delta, \bar{\beta}, \alpha \rangle$ denote the parameters of the model, where $\delta = \langle \delta_1, \ldots, \delta_J \rangle'$ and $\alpha = \langle \alpha_k^a, \alpha_k^x, \alpha_k^d \rangle'$. To estimate the δ_j parameters, I use a procedure similar to that in Bayer et al. (2007), which entails recursively solving for the optimal mean utility parameters using the contraction mapping algorithm of Berry et al. (1995).

The probability of agent i choosing the alternative that she was observed to choose is

$$\prod_j (P_{ij})^{y_{ij}}. \tag{14}$$

Under the assumption that individuals' choices are independent of one another, the probability of each person choosing the alternative that she was observed to choose is

$$\mathscr{L}(\theta)) = \prod_i \prod_j (P_{ij})^{y_{ij}}, \tag{15}$$

[64] To see why, note that the choice probability is $P_{ij} = Pr(U_{ij} > U_{ij'} \quad \forall j' \neq j) = Pr(U_{ij} - U_{ij'} > 0 \quad \forall j' \neq j)$, so P_{ij} depends only on the difference in utility, not the absolute level. Because only differences in utility are relevant to individuals' choices, it follows that only differences in the alternative-specific constants matter.

[65] With the normalization, utility for the outside option simplifies to $U_{i0} = \varepsilon_{i0} \; \forall \, i$.

[66] The decision to use a single year of data was driven by computational limitations on the shared server, but I am in the process adjusting the model to use multiple years of data. This change also will allow me to leverage additional sources of identifying variation.

and the log-likelihood function is

$$\log(\mathscr{L}(\theta)) = \sum_{i=1}^{N} \sum_{j=0}^{J} y_{ij} \log(P_{ij}). \tag{16}$$

The maximum likelihood estimator is the value of θ which maximizes the log-likelihood given the observed data.

The individual attributes interacted with advertising, distance and college characteristics include indicators for female, Black, Hispanic, and low-income (as proxied by free or reduced lunch price eligibility).[67] The college characteristics x_j are indicators for college j's sector.

5.5.1 Estimation of δ Using a Contraction Mapping

For 2013, the full analysis sample contains over 304,000 high school graduates attending approximately 270 colleges, so estimating the parameters in the choice model entails estimation of 270 δ_j parameters in addition to the heterogeneous parameters α. Instead of estimating the δ_j using maximum likelihood, I follow the approach in Bayer et al. (2007) and use the contraction mapping algorithm provided by Berry et al. (1995) to estimate the optimal college taste parameters given the heterogeneous parameters, $\delta^*(\alpha)$, while the α vector is estimated by maximum likelihood.

The logic of this approach is as follows: Berry (1994) observes that in the correctly specified model, the model predicted market shares $\hat{S}_j(\delta) = \sum_i P_{ij}(\delta)/N$ should equal the actual market shares $S_j = \sum_i y_{ij}/N$,

$$S_j = \hat{S}_j(\delta), \tag{17}$$

and shows that observed market shares can be explained by a unique vector of utility means. This relationship can be exploited to estimate the δ_j parameters,[68] and Berry et al. (1995) show that the fixed point relation $h(\delta_j) = \delta_j$ is a contraction mapping, so δ_j may be solved recursively using

$$\delta_j^{h+1} = \delta_j^h + \ln(S_j) - \ln\left(\hat{S}_j(\delta^h; \hat{\alpha})\right) \tag{18}$$

where $\hat{\alpha}$ is a guess of α.

---

[67] The heterogeneous parameters consist of (i) K parameters (α_k^o) on the interactions between the observed individual attributes z_{ik} and advertising a_{ij}^o, (ii) $R \cdot K$ parameters α_{rk}^x on the interactions between the K attributes z_{ik} and the R non-advertising product characteristics x_j, and (iii) K scalar parameters α_k^d on the interactions between z_{ik} and distance d_{ij}. There are a total of $K + RK + K$ heterogeneous parameters.

[68] Berry (1994) observes that the model's predicted market shares $\hat{S}_j(\delta) = \sum_i P_{ij}/N$ should equal the realized share of consumers in the market that choose college j, $S_j = N^{-1} \sum_i y_{ij}$, when the choice model is correctly specified. He shows that the equation $S_j = \hat{S}_j(\delta)$ is invertible, i.e., there is a unique δ^* that satisfies $S_j = \hat{S}_j(\delta)$, so the mean utility levels δ_j can be calculated from the market shares S_j.

Table 6: Estimates of Utility Parameters

Parameter	Estimate	Standard Error
Advertising ×		
Omitted	0.1102	0.0015
Female	0.0095	0.0016
Hispanic	0.0060	0.0020
Low-income	0.0189	0.0019
Black	−0.0074	0.0029
Distance ×		
Omitted	−2.4674	0.0046
Female	0.1262	0.0073
Hispanic	−0.3922	0.0097
Low-income	−0.5784	0.0098
Black	0.1538	0.0127

Notes: Table 6 reports select estimates from maximum likelihood estimation of the model presented in Sections 5.3 and 5.5. Advertising and distance are measured in standard deviation units. The magnitudes of the parameter estimates, which are measured in utility units, are not directly interpretable, but the signs indicate the direction of the effect, and the relative magnitudes reflect the relative strength of effects. Coefficients for Female, Hispanic, Low-income, and Black are relative to the omitted group, which is white, not low-income males.

5.6 Estimation Results

The full set of estimated parameters from the demand model are reported in Figure 10. I focus on a few key takeaways.

Table 6 shows maximum likelihood estimates of the structural demand model using data from 2013. Recall that the parameters capture differences in utility relative to the outside option, which is no college. When the model is estimated using a single year of data and includes mean utility parameters for each college (δ_j), it is not possible to identify the mean utility for each college type (public 4-year university, community college, private non-profit, and for-profit). To capture preferences for different types of colleges, I interact each college type with student demographic variables. These parameters capture demographic-specific tastes for a given college type relative to the mean utility tastes δ_j. In contrast to the college type variables, advertising and distance vary within college and thus may enter the model without interactions (see Equation 10). My model specification includes these mean utility parameters for advertising and distance as well as heterogeneous parameters on the interactions with students' demographic characteristics. In practice, however, the demographic characteristics I use are binary, so the mean effects for advertising and distance capture the marginal utility for the reference group.

5.6.1 The Marginal Utility of Advertising

A key takeaway from the model is that college advertising increases utility relative to the outside option, so advertising increases college enrollment. The estimated parameter on advertising (measured in standard deviation units) is 0.1102 and corresponds to the utility effect for the omitted/reference group (non-Black, non-Hispanic, and non-disadvantaged males) of a one standard deviation increase in advertising. The magnitudes of the interaction parameters are smaller than the main effect, indicating that advertising has a positive effect on utility for all demographic groups included in the model. In particular, advertising has a relatively large effect for low-income students (0.0189) and a smaller effect for females (0.0095). Hispanics also experience greater utility from advertising than non-Hispanics (0.0059). In contrast, the effect of advertising is lower for Black students relative to non-Black students (-0.0074).

5.7 Tastes for Distance and College Type

As expected, students prefer closer colleges. The parameter on distance (which also is measured in standard deviation units) is -2.467 for the reference group. The negative signs on the heterogeneous parameters for disadvantaged (-0.5784) and Hispanic (-0.3922) students indicate that individuals in these groups experience additional disutility from distance. In contrast, the coefficients for Black and female students are positive, so these students experience less disutility from distance.

The heterogeneous parameters on college type reveal variation in preferences across demographic groups. The college type parameters (public 4-year university, community college, private non-profit, and for-profit) are positive for female students and Black students and negative for disadvantaged and Hispanic students.

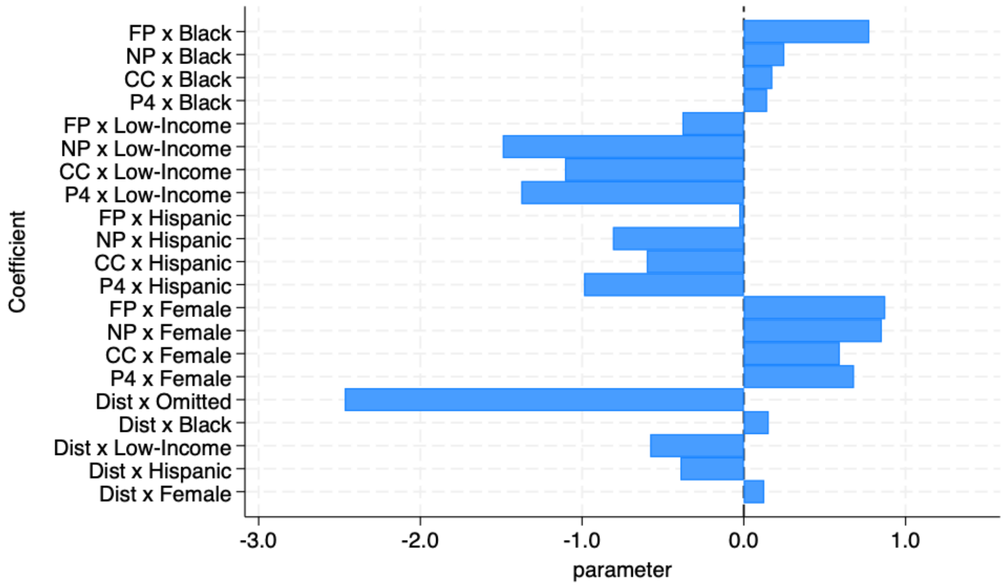
5.8 The Enrollment Responses to a Ban on College Advertising

In 2019, U.S. deputy undersecretary of education Diane Auer Jones described aggressive marketing and advertising as "the biggest consumer protection issue" in higher education.[69] In the years leading up to her remarks, the Federal Trade Commission (FTC) had pursued enforcement actions against DeVry University and the University of Phoenix—two of the largest for-profit college chains in the U.S., alleging that their TV and radio ads contained misleading content and false claims to lure students into enrolling. By 2020, lawmakers had introduced several bills in Congress to regulate college advertising.[70]

[69] See https://tcf.org/content/report/much-education-students-getting-tuition-dollar/.

[70] The 2020 College Affordability Act proposes to ban colleges from using federal funds on advertising and marketing if colleges do not spend a minimum amount on instruction. See

Figure 10: Demand Model: Estimated Utility Parameters



Notes: Figure 10 reports heterogeneous parameter estimates from maximum likelihood estimation of the model presented in Sections 5.3 and 5.5.

Motivated by these proposals, I use the estimated demand model to simulate student responses to advertising regulation. My primary objective with this analysis is to quantify how college advertising affects the allocation of students across different colleges and the outside option. As such, the counterfactual of interest is a total ban on college advertising.[71] I conduct two types of policy counterfactuals to examine how high school graduates' immediate post-secondary enrollment choices would change in absence of advertising.

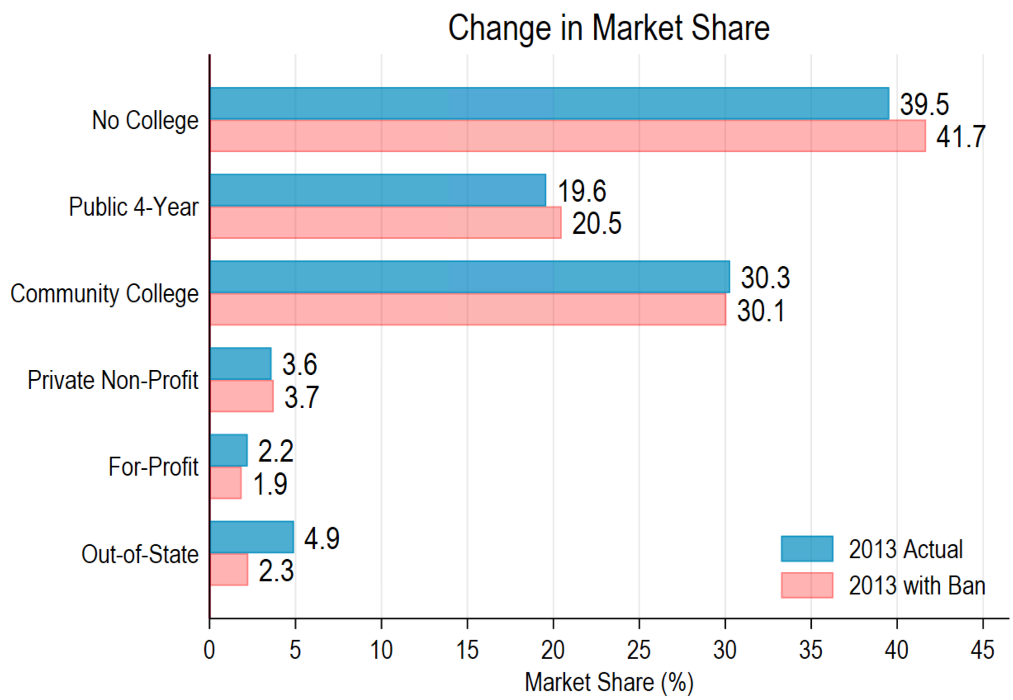
First, I consider how market shares of different college types change when all advertising is banned. The other components of the model (estimated parameters and data) remain the same. I compute the predicted probabilities for each individual at each college under the zero advertising regime. I then aggregate these probabilities by college type.

The results of this exercise are shown in Figure 11. The first thing to observe is that banning advertising causes the outside option (no college) market share to increase by over 2 percentage

https://hechingerreport.org/with-competition-up-enrollment-down-colleges-are-spending-billions-on-marketing-and-advertising/.

[71] Even though the consumer protection concerns of college advertising are largely a response to advertising abuses by for-profit colleges, the proposals in Congress are written as applying to *all* colleges that receive federal funds, including most public universities and private non-profit institutions. In light of this, I focus on the effects of regulation that would ban all colleges from advertising.

Figure 11: Policy Counterfactual: Effect of a Ban on College Advertising



Notes: Figure 11 shows each college type's market share before and after the counterfactual ban, which uses the estimates from the model defined in Equation (5.3) and sets advertising by all colleges equal to the minimum (in standard deviation units) while keeping everything else the same.

points, from 0.395 to 0.417. In contrast, community colleges, for-profit colleges, and out-of-state colleges lose market share when advertising is banned.[72] These results are consistent with the results reported in the design-based section, which indicated that advertising by for-profit colleges and community colleges increases overall college-going. Finally, the figure shows that public 4-year and private non-profit market shares increase after the ban.

Second, I examine variation in responses across demographic groups. To do so, I generate a simulated population of individuals by drawing Type 1 Extreme Value errors for each of the observed students in my dataset. I then use the estimated model parameters, the students' observed covariates, and the error to compute each simulated individual's utility and choice at the 2013 advertising levels and when advertising is banned. The results are shown in Table 7.

The results indicate that 4.9 percent of students who enrolled in college in 2013 would make a different choice if advertising were banned. For a large portion of these students, advertising affects the extensive margin decision: 3.1 percent of individuals would stop attending college if advertising were shut down. There are interesting differences across demographic groups. Among white students who make a different choice when advertising is banned, they are as likely to switch from one college to another as they are to exit the market. In contrast, racial minorities and low-income students are more likely to forego college altogether after the ban. For example: 3.4 percent of Hispanic students choose not to enroll, compared to only 1.2 percent opting to switch schools. The numbers for low-income students are similar. In addition, compared to Hispanic students, a larger share of Black students make a different choice.

6 Conclusions

In this paper, I use rich data on college enrollment and TV advertising to estimate the effect of college advertising on demand for college. To identify the causal effect, I leverage the discontinuity in advertising at the boundary between neighboring media markets in combination with panel variation. Regression results from the border design indicate that advertising has a positive effect on overall college enrollment. Extensive margin effects are particularly large for low-income and Hispanic students. I also find that market expansion effects are largest for community colleges.

After estimating effects using the border strategy, I develop a simple discrete choice demand model to simulate the effects of counterfactual advertising regulation. I find that banning advertising would cause a substantial share of students to stop attending college, and consistent with the design-based results, low-income and Hispanic students are more likely to forego college altogether.

[72] The decline in out-of-state college share is an unexpected result which I am investigating further.

Table 7: Heterogeneous Responses to a College Advertising Ban

	Any Different Choice (%)	Switch College (%)	Not Enroll (%)
All Students	4.9	1.8	3.1
White	5.7	2.8	2.8
Black	5.0	2.1	2.9
Hispanic	4.6	1.2	3.4
Low-income	4.4	1.0	3.3
Not low-income	5.3	2.5	2.8
Female	5.5	2.4	3.1
Male	4.3	1.3	3.0

Notes: Table 7 reports simulation results of changes in college choices after an advertising ban. The first column shows the percent of students within each row (unconditional on enrollment) that made a different college choice after the advertising ban. The second column shows the percent that attend a different college after the ban, while the rightmost column shows the percent that stop attending college.

The analyses in this paper provide an important foundation for understanding how advertising affects individuals' enrollment choices. In light of my findings, an important question that emerges is whether students are better off after being induced to enroll in college. Fruitful extensions of this work would study the effects of college advertising on college persistence/degree completion and earnings. The findings would be informative about the channels through which advertising affects enrollment choices: increases in degree completion and earnings gains would be consistent with advertising generating awareness or information effects. In contrast, negative outcomes—such as faster dropout or diminished earnings—would be more likely to result if persuasive advertising causes students to over-value the advertised product and make enrollment mistakes. A related and promising direction for future work is to provide direct evidence of college advertising's informative and persuasive effects.

References

Aizawa, N. and Kim, Y. S. (2018). Advertising and risk selection in health insurance markets. *American Economic Review*, 108(3):828–867.

Aizawa, N. and Kim, Y. S. (2019). Government Advertising in Market-Based Public Programs: Evidence from Health Insurance Marketplace. *SSRN Working Papers [Unpublished manuscript]*.

Aizawa, N. and Kim, Y. S. (2022). Public and Private Provision of Information in Market-Based Public Programs: Evidence from Advertising in Health Insurance Marketplaces. *SSRN Working Papers [Unpublished manuscript]*.

Armona, L. and Cao, S. (2022). Redesigning Federal Student Aid in Higher Education. *SSRN Working Papers [Unpublished manuscript]*.

Avery, C., Castleman, B. L., Hurwitz, M., Long, B. T., and Page, L. C. (2021). Digital messaging to improve college enrollment and success. *Economics of Education Review*, 84(September):102170.

Avery, C. and Kane, T. (2004). Student Perceptions of College Opportunities. The Boston COACH Program. In Hoxby, C. M., editor, *College Choices: The Economics of Where to Go, When to Go, and How to Pay for It*, pages 355–394. University of Chicago Press.

Bagwell, K. (2007). The Economic Analysis of Advertising. In Armstrong, M. and Porter, R. H., editors, *Handbook of Industrial Organization*, volume 3, pages 1701–1844. North-Holland, 1 edition.

Bayer, P., Ferreira, F., and McMillan, R. (2007). A Unified Framework for Measuring Preferences for Schools and Neighborhoods. *Journal of Political Economy*, 115(4):588–638.

Becker, G. S. and Murphy, K. M. (1993). A simple theory of advertising as a good or bad. *Quarterly Journal of Economics*, 108(4):941–964.

Bergman, P., Denning, J. T., and Manoli, D. (2019). Is Information Enough? The Effect of Information about Education Tax Benefits on Student Outcomes. *Journal of Policy Analysis and Management*, 38(3):706–731.

Berry, S., Levinsohn, J., and Pakes, A. (1995). Automobile Prices in Market Equilibrium. *Econometrica*, 63(4):841–890.

Berry, S., Levinsohn, J., and Pakes, A. (2004). Differentiated products demand systems from a combination of micro and macro data: The new car market. *Journal of Political Economy*, 112(1):68–105.

Berry, S. T. (1994). Estimating Discrete-Choice Models of Product Differentiation. *The RAND Journal of Economics*, 25(2):242–262.

Bettinger, E. P., Terry Long, B., Oreopoulos, P., and Sanbonmatsu, L. (2012). The Role of Application Assistance and Information in College Decisions: Results from the H&R Block FAFSA Experiment. *The Quarterly Journal of Economics*, 127(3):1205–1242.

Black, S. E. (1999). Do Better Schools Matter? Parental Valuation of Elementary Education. *Quarterly Journal of Economics*, 114(2):577–599.

Bleemer, Z. (2021). Top Percent Policies and the Return to Postsecondary Selectivity. *Center for Studies in Higher Education Research and Occasional Papers Series [Unpublished manuscript]*.

Cellini, S. R. and Chaudhary, L. (2020). Commercials for college? Advertising in higher education. Technical report, Brookings.

Cellini, S. R. and Chaudhary, L. (2023). Where do colleges advertise? Demographic targeting by U.S. colleges. *Economics of Education Review*, 94:102402.

Cook, E. E. (2021). Competing Campuses: Equilibrium Prices, Admissions, and Undergraduate Programs in US Higher Education. *[Unpublished manuscript]*, page 47.

Dillon, E. W. and Smith, J. A. (2017). Determinants of the Match Between Student Ability and College Quality. *Journal of Labor Economics*, 35(1):45–66.

Dubois, P., Griffith, R., and O'Connell, M. (2018). The Effects of Banning Advertising in Junk Food Markets. *Review of Economic Studies*, 85(1):396–436.

Dynarski, S., Libassi, C., Michelmore, K., and Owen, S. (2021). Closing the Gap: The Effect of Reducing Complexity and Uncertainty in College Pricing on the Choices of Low-Income Students. *American Economic Review*, 111(36):1721–56.

Dynarski, S. M. and Scott-Clayton, J. E. (2006). The cost of complexity in federal student aid: Lessons from optimal tax theory and behavioral economics. *National Tax Journal*, 59(2):319–356.

Goeree, M. S. (2008). Limited Information and Advertising in the U.S. Personal Computer Industry. *Econometrica*, 76(5):1017–1074.

Hahn, J., Todd, P., Klaauw, W. V. D., Econometrica, S., and Jan, N. (2001). Identification and Estimation of Treatment Effects with a Regression-Discontinuity Design. *Econometrica*, 69(1):201–209.

Hillman, N. W. (2016). Geography of College Opportunity: The Case of Education Deserts. *American Educational Research Journal*, 53(4):987–1021.

Howell, J. S. (2010). Assessing the impact of eliminating affirmative action in higher education. *Journal of Labor Economics*, 28(1):113–166.

Hoxby, C. and Avery, C. (2013). The Missing "One-Offs": The Hidden Supply of High-Achieving, Low-Income Students. *Brookings Papers on Economic Activity*, (Spring):1–61.

Hoxby, C. and Turner, S. (2013). Expanding College Opportunities for High-Achieving, Low Income Students. *SIEPR Discussion Paper*, (12-014):1–57.

Hoxby, C. M. and Turner, S. (2015). What High-Achieving, Low-Income Students Know About College. *American Economic Review: Papers & Proceedings*, 105(5):514–517.

Kapor, A. (2020). Distributional Effects of Race-Blind Affirmative Action. *[Unpublished manuscript]*.

Marcus, J. (2021). From Google ads to NFL sponsorships: Colleges throw billions at marketing themselves to attract students. *The Hechinger Report*.

McFadden, D. (1973). Conditional Logit Analysis of Qualitative Choice Behavior. In Zarembka, P., editor, *Frontiers in Econometrics*, pages 105–142. Academic Press, New York.

Page, L. C. and Scott-Clayton, J. (2016). Improving college access in the United States: Barriers and policy responses. *Economics of Education Review*, 51:4–22.

Pallais, A. (2015). Small Differences that Matter: Mistakes in Applying to College. *Journal of Labor Economics*, 33(2):493–520.

Scott-Clayton, J. (2012). Information constraints and financial aid policy. *Student Financing of Higher Education: A Comparative Perspective*, pages 75–97.

Shapiro, B. T. (2018). Positive Spillovers and Free Riding in Advertising of Prescription Pharmaceuticals: The Case of Antidepressants. *Journal of Political Economy*, 126(1):381–437.

Shapiro, B. T., Hitsch, G. J., and Tuchman, A. E. (2021). TV Advertising Effectiveness and Profitability: Generalizable Results From 288 Brands. *Econometrica*, 89(4):1855–1879.

Sinkinson, M. and Starc, A. (2019). Ask Your Doctor? Direct-to-Consumer Advertising of Pharmaceuticals. *Review of Economic Studies*, 86(2):836–881.

Smith, J., Pender, M., and Howell, J. (2013). The full extent of student-college academic undermatch. *Economics of Education Review*, 32:247–261.

Spenkuch, J. and Toniatti, D. (2018). Political advertising and election results. *Quarterly Journal of Economics*, 133(4):1981–2036.

The Nielsen Company (2013). A Look Across Media. Technical Report December.

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

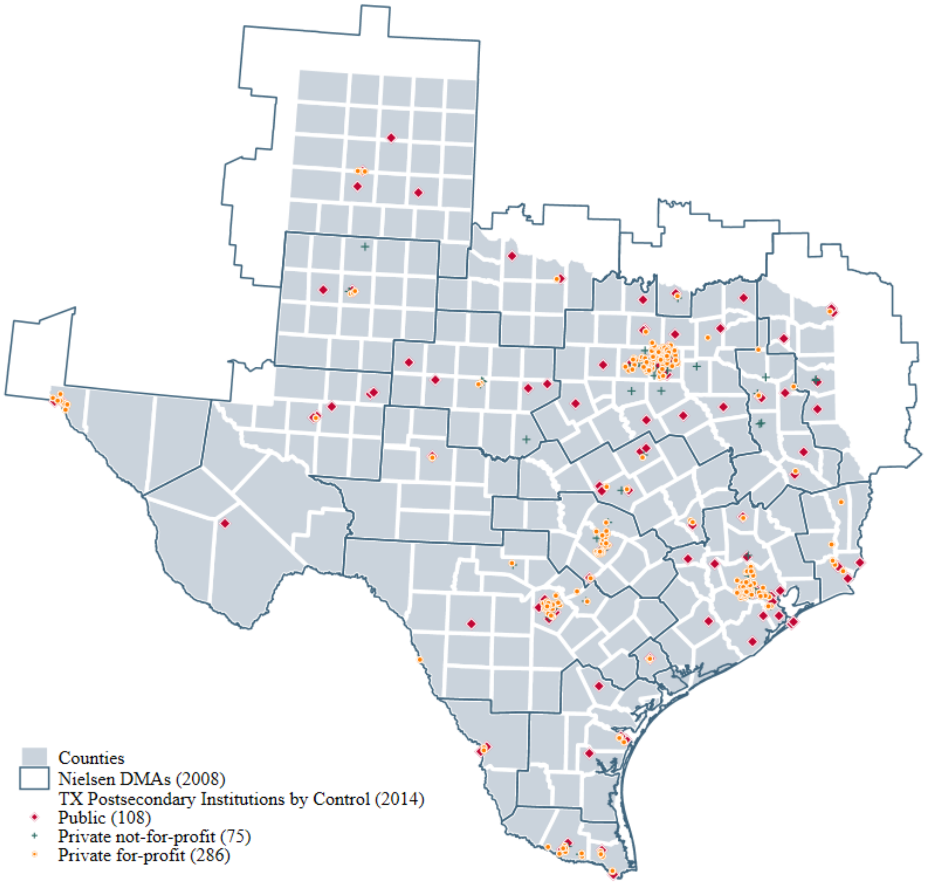


Figure 12: Geographic Distribution of Texas Colleges

A Data and Institutional Details

A.1 Data: Nielsen Ad Intel

In this section, I provide additional details about the coverage of the advertising data obtained from Nielsen's Ad Intel database.

A.1.1 Media Types and Coverage

The Nielsen Ad Intel data cover the media types shown in Table 9 below.

Data coverage varies by media type for local media. For instance, according to Nielsen documentation, as of December 2013, local magazines covered the top 42 DMAs, local radio covered 43 DMAs, local supplements covered the top 50 DMAs, local/regional cable TV covered 51 DMAs, local newspapers covered the top 75 DMAs, local internet covered 79 DMAs, and local billboards covered over 180 DMAs,

Outdoor (billboard) occurrences are assigned a monthly date. Individual spot radio occurrences are available, but the date provided for spot radio ads is for the Monday of the week of the

Table 8: Texas DMAs

DMA Rank (Nationally)	DMA Name	# TX HHs in DMA	% of DMA HHs Located in TX
5	DALLAS-FT. WORTH	2,616,600	100.00
10	HOUSTON	2,200,500	100.00
37	SAN ANTONIO	856,600	100.00
44	AUSTIN	717,000	100.00
83	SHREVEPORT	112,000	28.62
87	HARLINGEN-WESLACO-BROWNSVILLE-MCALLEN	361,500	100.00
89	WACO-TEMPLE-BRYAN	347,400	100.00
97	EL PASO (LAS CRUCES)	244,700	76.93
109	TYLER-LONGVIEW	272,700	100.00
129	CORPUS CHRISTI	201,800	100.00
131	AMARILLO	157,200	79.92
141	BEAUMONT-PORT ARTHUR	171,900	100.00
142	LUBBOCK	162,900	100.00
146	WICHITA FALLS & LAWTON	82,400	52.09
151	ODESSA-MIDLAND	145,900	98.58
161	SHERMAN-ADA	47,000	35.88
165	ABILENE-SWEETWATER	116,200	100.00
186	LAREDO	71,000	100.00
196	SAN ANGELO	55,900	100.00
204	VICTORIA	32,100	100.00

Notes: Data are for 2010.
Source: In-State Broadcast Programming: Report to Congress Pursuant to Section 304 of the Satellite Television Extension and Localism Act of 2010 (Report DA 11-1454), Media Bureau, U.S. Federal Communications Commission, 2011.

occurrence.

National Cinema data are received by Nielsen from national cinema companies on a monthly basis and are reported at the ad campaign flight level detail. Cinema media cover on-screen commercial and slide campaigns appearing on 100% of the markets available in a Cinema company's inventory. Nielsen estimates that each National Cinema commercial runs approximately 4 times per day per screen for the length of the campaign flight. In 2013, Nielsen's estimates of total screen potential for any given National campaign (given the contributing companies) was approximately 34,000+ screens. The universe of coverage for Regional Cinema data consists of on-screen commercial and slide campaigns appearing in at least two but in less than 100% of the markets available in a cinema company's inventory. Total screen potential for any given Regional Cinema campaign varies depending on the cities which they aired.

Table 9: Media Types Covered in Ad Intel

Level	Media Group	Media Types
Local	TV	Spot, Network and Syndicated Clearance, Local/Regional Cable
Local	Radio	Spot Radio
Local	Print	Local Newspaper, Outdoor, Local Magazine, Local Supplements
Local	Cinema	Regional Cinema
Local	Internet	Local Internet,
National	TV	Network, Cable, Syndicated, Spanish Language Network and Cable
National	Radio	Network Radio
National	Print	National Newspaper, National Magazine, Sunday Supplements
National	Cinema	National Cinema
National	Internet	National Internet
National	Digital	National Digital
National	FSI	Free-Standing Insert Coupons*

*During the analysis period, there are no college ads for Free Standing Insert Coupons.
Source: 2020 Ad Intel Dataset Manual, Kilts Center for Marketing

The internet media types consist of all reportable sites and sub-sites captured via Nielsen's probing technology (ad supported websites only). Micro-sites are not tracked. These data are reported weekly and cover the two weeks prior. Occurrence-level detail is not available.

National digital media type becomes available in 2018, replacing national and local internet. National Digital captures ads on desktop web (all operating systems), mobile web (iPhone and Android), tablet Web (iPad and Android), displays on all desktop and mobile+tablet sites, native content (BuzzFeed, Yahoo, Outbrain, Taboola, etc.), IAB standard formats, skins, animated HTML5, common non-standard formats, pre-video ads on all desktop and mobile+tablet web on YouTube only. It does not capture ads in social media (Facebook, Twitter, etc.), Mobile+Tablet in-app, search, non-standard executions, highly targeted campaigns (i.e., very specific demos, individual targeting, etc.), re-targeting, or anything behind a login or pay-wall.

A.1.2 Ad Occurrences

Each Ad Intel advertisement occurrence record contains rich information about an ad's source and characteristics. Table 10 provides examples of the source variables used to describe records in the Nielsen data.

Nielsen categorizes TV media programming into four broad groups: Network TV, Cable TV, Syndicated TV, and Spot TV. Although Network and Syndicated TV ads are purchased nationally, they are broadcast locally at TV stations. Nielsen provides two types of occurrence records for these ads: the national-level record (e.g. Network TV Media Type) indicates when the ad should

Table 10: Examples of Ad Source Information

Brand	Advertising Subsidiary	Advertising Parent	Product Type
University of Phoenix	University of Phoenix Inc.	Apollo Group Inc.	University
Kaplan University	Kaplan Higher Education Corp.	Graham Holdings Co.	University
Everest College	Zenith Education Group Inc.	ECMC Group Inc.	College
Sinclair Comm. Coll.	City of Dayton	State of OH	College

have been aired at each local station, while the local-level record (e.g. Network Clearance Spot TV Media Type) indicates the realized date and time when each ad is aired by local TV stations (which implicitly also indicates which markets aired the ad). An analogous coding occurs for Syndicated and Syndicated Clearance Spot TV. Local TV stations may preempt national ad purchases if another advertiser offers a higher price for the spot; furthermore, ads can get moved around due to breaking news or other changes to programming schedules. For these reasons, an ad purchased nationally may not air in all DMAs. The percentage of markets that air the ad determine the ad's "clearance" rate.

B Additional Tables and Figures

Figure 13: On Average, Young Adults Watch Over Two Hours of TV Daily

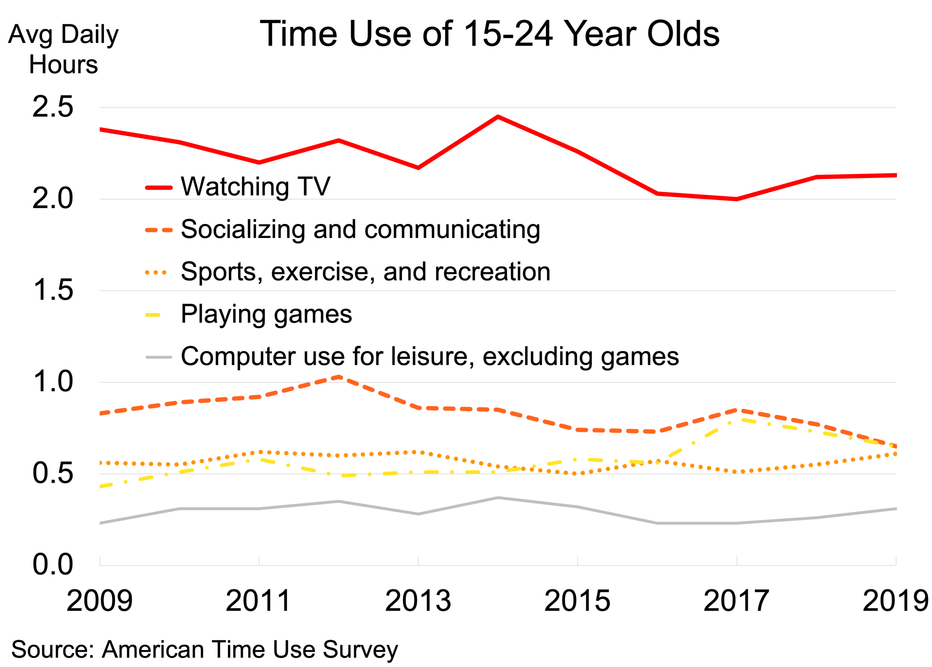


Figure 14: Most Video Viewing is on Traditional TV

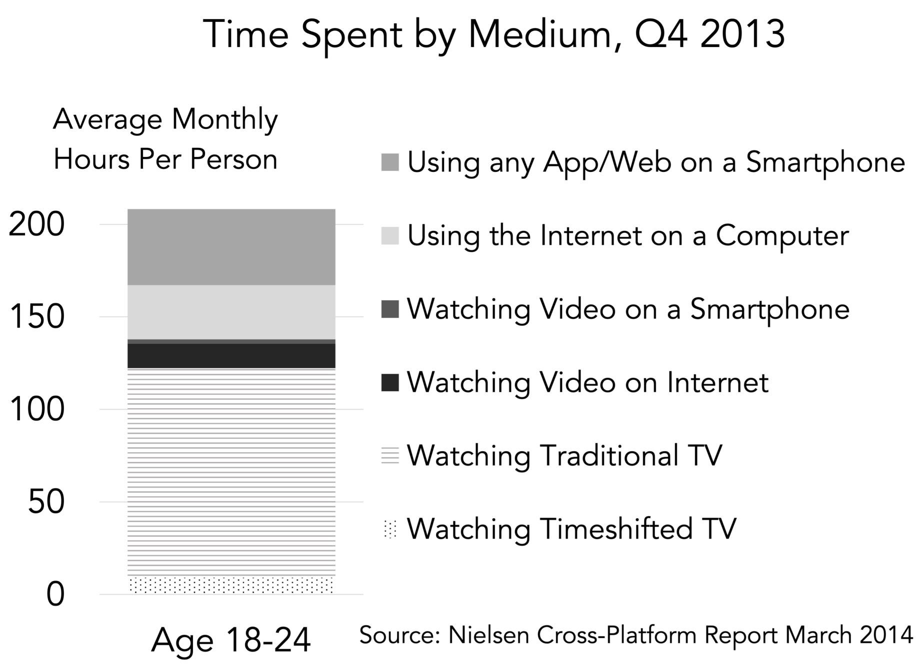
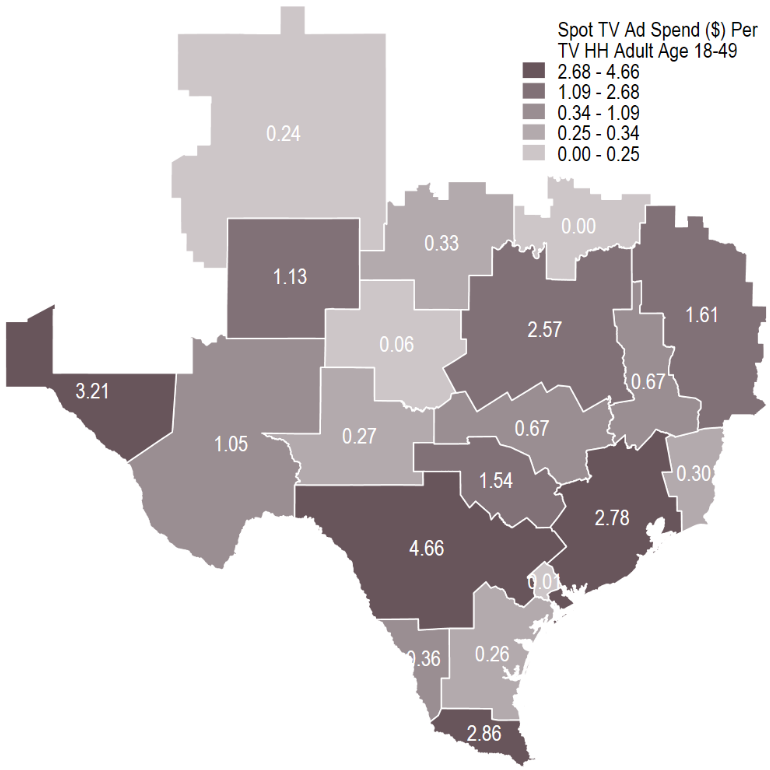


Figure 15: Per-Capita Spot TV Advertising Spending in Texas DMAs

(a) 2010



(b) 2015

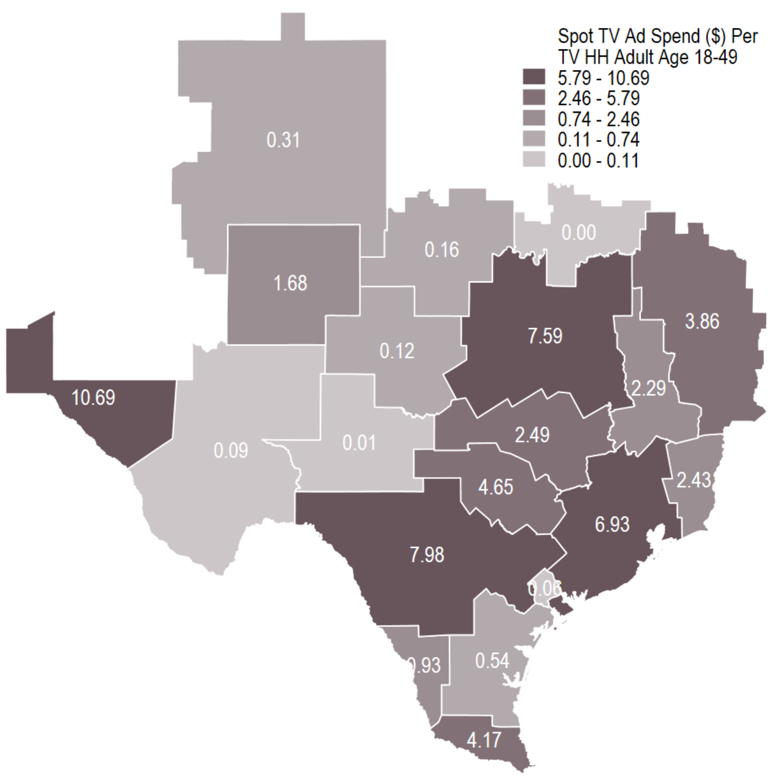


Table 11: Market Expansion Effects of College Advertising by Demographic Group

Outcome: College Enrollment	White	URM	Hispanic	Black	Low-Income	Female
TX College Spot TV Ads (10,000s)	0.0015	0.0155***	0.0238***	–0.0016	0.0062	0.0039
	(0.0040)	(0.0049)	(0.0060)	(0.0137)	(0.0052)	(0.0038)
Percent Change	0.002	0.029***	0.046***	–0.003	0.013	0.006
Avg Dep Var	0.647	0.541	0.518	0.596	0.463	0.657
Mean # Ads (DMA-Year)	19,068	19,328	19,328	18,890	19,328	19,328
Median # Ads (DMA-Year)	4,465	4,520	4,520	4,465	4,520	4,520
Adj R-Sq	0.175	0.171	0.182	0.142	0.154	0.163
Adj R-Sq (within)	0.122	0.115	0.109	0.096	0.098	0.113
Border-County Pairs	88	89	89	69	89	89
Counties	119	122	122	86	122	122
Years	5	5	5	5	5	5
Observations	153,385	127,885	90,881	31,034	92,809	144,784

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Notes: Table 11 displays regression results from the model presented in Equation (1) with the sample restricted to the students belonging to the subgroup indicated in the columns. All specifications include fixed effects for the border pair-by-cohort and high school. Standard errors (in parentheses) are clustered at the DMA-by-year level. For additional details, see the note for Table 3.

C Computational Details

C.0.1 Analytical Gradient

The speed in which the optimization routine computes the numerical likelihood can be improved drastically by providing the analytical gradient of the log likelihood (with respect to the heterogeneous parameters α) as an input to the optimizer. In this section I derive the analytical gradient of ℓ with respect to the vector of heterogeneous parameters α.

The observed component of indirect utility, V_{ij}, is

$$V_{ij} = \bar{\beta}^a a_{ij} + \left(\sum_{k \in K} z_{ik}\alpha_k^a\right) a_{ij} + \bar{\beta}^d d_{ij} + \left(\sum_{k \in K} z_{ik}\alpha_k^d\right) d_{ij} + \left(\sum_{k \in K} z_{ik}\alpha_k^x\right) x_j. \tag{19}$$

The parameters are estimated via maximum likelihood. The $(K+K+RK)$-dimensional vector of heterogeneous parameters $\alpha = \langle \alpha^a, \alpha^x, \alpha^d \rangle$ can be partitioned into three components: (i) the K-dimensional vector $\alpha^a = \langle \alpha_1^a, \ldots, \alpha_K^a \rangle$, which contains the parameters on the interaction between z_{ik} and $a_{jm(i)}$ (college j's advertising in i's DMA) (ii) the K-dimensional vector $\alpha^d = \langle \alpha_1^d, \ldots, \alpha_K^d \rangle$, which contains the parameters on the interaction between z_{ik} and the individual's distance to college j, d_{ij}, and (iii) the RK-dimensional vector $\alpha^x = \langle \alpha_{11}^x, \ldots \alpha_{R1}^x, \ldots, \alpha_{1k}^x, \ldots, \alpha_{RK}^x \rangle$, which contains the parameters on the interactions between the K observed individual attributes z_{ik} and the R indicators of college type x_{rj}.

The log likelihood function is

$$\ell = \log(\mathscr{L}) = \sum_{i=1}^{N} \sum_{j=0}^{J} y_{ij} \cdot \log(P_{ij}), \tag{20}$$

where P_{ij} is the probability that individual i chooses product j, given by

$$P_{ij} = \frac{\exp(V_{ij} + \delta_j)}{\sum_{j'} \exp(V_{ij'} + \delta_{j'})}. \tag{21}$$

Substituting (21) into (20) and simplifying gives

$$\ell = \sum_i \sum_j y_{ij} \cdot \left[(\delta_j + V_{ij}) - \log\left(\sum_{j'} \exp(\delta_{j'} + V_{ij'})\right)\right]$$

$$\ell = \underbrace{\sum_i \sum_j y_{ij}\left(\delta_j + V_{ij}\right)}_{\ell_A} - \underbrace{\sum_i \sum_j y_{ij}\left[\log\left(\sum_{j'} \exp(\delta_{j'} + V_{ij'})\right)\right]}_{\ell_B}. \tag{22}$$

To derive the analytical gradient of the log-likelihood with respect to the heterogeneous parameters, I take the partial derivative of ℓ with respect to each individual parameter in α. For interactions between individual characteristics z_i and college attributes x_k, the parameters are α^x_{rk} in α^x. The partial derivatives of ℓ_A and ℓ_B with respect to each α^x_{rk} are

$$\frac{\partial \ell_A}{\partial \alpha^x_{rk}} = \sum_i \sum_j y_{ij} z_{ik} x_{rj}$$

$$\frac{\partial \ell_B}{\partial \alpha^x_{rk}} = \sum_i \sum_j y_{ij} \left(\sum_{j'} z_{ik} x_{rj'} P_{ij'} \right)$$

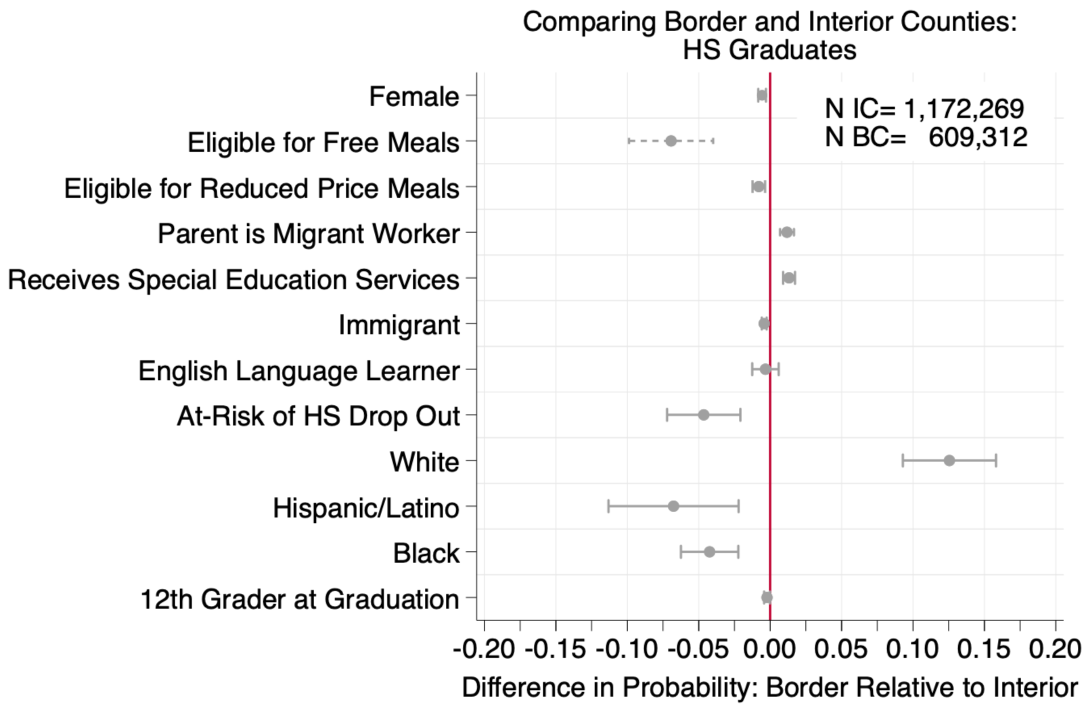
Similarly, the partial derivatives of ℓ_A and ℓ_B with respect to each parameter α^d_k in α^d are

$$\frac{\partial \ell_A}{\partial \alpha^d_k} = \sum_i \sum_j y_{ij} z_{ik} d_{ij}$$

$$\frac{\partial \ell_B}{\partial \alpha^d_k} = \sum_i \sum_j y_{ij} \left(\sum_{j'} z_{ik} d_{ij'} P_{ij'} \right).$$

Figure 16: Comparing Border and Interior Counties

(a) Student Characteristics



(b) School Attributes