Advertising and Demand for Addictive Goods: The Effects of E-Cigarette Advertising

Anna E. Tuchman*†

August 7, 2019

Abstract

Although TV advertising for traditional cigarettes has been banned since 1971, advertising for electronic cigarettes remains unregulated. The effects of e-cigarette ads have been heavily debated, but empirical analysis of the market has been limited. Analyzing both individual and aggregate data, I present descriptive evidence showing that e-cigarette advertising reduces demand for traditional cigarettes and individuals treat e-cigarettes and traditional cigarettes as substitutes. I then specify a structural model of demand for cigarettes that incorporates addiction and allows for heterogeneity across households. The model enables me to leverage the information content of both datasets to identify variation in tastes across markets and the state dependence induced on choice by addiction. Using the demand model estimates, I evaluate the impact of a proposed ban on e-cigarette television advertising. I find that in the absence of e-cigarette advertising, demand for traditional cigarettes would increase, suggesting that a ban on e-cigarette advertising may have unintended consequences.

*Assistant Professor of Marketing, Northwestern University Kellogg School of Management, Email: anna.tuchman@kellogg.northwestern.edu. I wish to thank Lanier Benkard, Kristina Brecko, Liran Einav, Pedro Gardete, Brett Gordon, Wes Hartmann, Sridhar Narayanan, Navdeep Sahni, Stephan Seiler, and Brad Shapiro for their insightful comments. I am especially grateful to my advisor, Harikesh Nair, for his guidance and support. Thanks also to the participants at the 2016 FTC Microeconomics, IIOC, Marketing Science, and 9th Workshop on Economics of Advertising and Marketing conferences, as well as seminar participants at Columbia, Duke, Emory, Harvard, HKUST, INSEAD, Northwestern, Rochester, Stanford, UC Berkeley, UCLA, UCSD, UChicago, University of Colorado, USC, UPenn, UTDallas, Wash U in St. Louis, and Yale for many helpful comments and suggestions.

†Calculated (or Derived) based on data from The Nielsen Company (US), LLC and marketing databases provided by the Kilts Center for Marketing Data Center at The University of Chicago Booth School of Business. The conclusions drawn from the Nielsen data are those of the researchers and do not reflect the views of Nielsen. Nielsen is not responsible for, had no role in, and was not involved in analyzing and preparing the results reported herein.

1 Introduction

Smoking cigarettes is still the leading cause of preventable death in the United States, killing more than 480,000 people a year. As a result, cigarette advertising remains a public health issue that is intensely debated by cigarette companies, policy makers, and academic researchers. Although all TV and radio advertising for traditional cigarettes has been banned since 1971, attention to the advertising ban has been renewed by the entry of e-cigarettes into the market. E-cigarettes first entered the U.S. market in 2007 and quickly grew to become a $2 billion industry by 2014 (Crowley (2015)). E-cigarette advertising does not fall under the tobacco advertising ban and thus remains unregulated. Advertising for e-cigarettes has proliferated in recent years on television, online, and in print media outlets. By 2013, e-cigarette marketing spending exceeded $79 million with the majority of spending going towards TV and magazine advertising (Kantar Media (2014), Kim, Arnold, & Makarenko (2014)).[1] Advocates for a ban on e-cigarette advertising argue that e-cigarette ads glamorize smoking and that e-cigarettes may act as a gateway into smoking traditional cigarettes and marijuana. Physicians have also expressed concern that e-cigarette advertising may divert smokers who want to quit away from clinically-proven cessation products. Proponents of e-cigarettes argue that e-cigarettes may be used as a tool to effectively help quit smoking, and thus e-cigarette advertising may reduce demand for tobacco cigarettes. To date, there exists little empirical evidence in support of either of these positions.

In this paper, I use data from 2010 to 2015 to empirically test whether e-cigarette advertising increases or decreases demand for traditional cigarettes and smoking cessation products. Using both descriptive and structural methods, I find that e-cigarette advertising reduces demand for both traditional cigarettes and smoking cessation products. Next, I quantify the impact of e-cigarette advertising on demand by considering a counterfactual in which e-cigarette TV advertising were banned between 2012-2015. The model predicts that cigarette sales would have been 1.0% higher than observed sales if there had been no e-cigarette ads during this period. This is an economically significant increase when compared to the fact that sales of cigarettes in the U.S. fell by 2.2% between 2011 and 2012 (FTC (2015)). The counterfactual also predicts that sales of smoking cessation products would have been 1.0% higher if e-cigarette advertising had been banned. This result provides some support

[1] Kim et al. (2014) use data from Kantar Media and Nielsen to estimate that in 2012, $18.3 million in e-cigarette ad spending was split between TV (27%), magazines (59%), internet (1%), radio (9%), and newspapers (4%). The authors predict that going forward “TV expenditures will likely outpace other channels given the recent national cable network campaigns for Blu eCigs and NJOY."

for physicians' concerns that e-cigarette advertising may lead smokers who want to quit to substitute away from clinically-proven cessation products.

Although the market for e-cigarettes is still small relative to tobacco cigarettes, awareness and use of e-cigarettes has been growing steadily in recent years. Despite being a quickly growing new category, much is still unknown about e-cigarettes to date. Existing research relating to e-cigarettes has generally been focused on addressing three types of questions: i) what are the health effects of e-cigarettes to users and non-users, ii) are e-cigarettes an effective tool to help smokers quit smoking, and iii) do e-cigarettes hamper existing tobacco control efforts. This paper primarily relates to the third category.

In general, whether e-cigarettes have a positive or negative impact on public health and tobacco control depends on the interplay between the potential benefits to current smokers and the undesired adoption of nicotine products by non-smokers. The World Health Organization's 2014 report on electronic nicotine delivery systems discusses the two primary arguments made by advocates for a ban on e-cigarette advertising: the gateway and renormalization effects. The *gateway* effect refers to the possibility that e-cigarettes will lead more non-smokers to initiate nicotine use and that once addicted to nicotine, non-smokers will be more likely to switch to smoking cigarettes than they would if they were not e-cigarette users. The *renormalization* effect refers to the possibility that marketing that portrays e-cigarettes as an attractive product will increase the attractiveness of cigarettes as well. The WHO (2014) report acknowledges that the existence and magnitude of the gateway and renormalization effects is an empirical question that is still understudied due to the limited availability of data.[2]

To my knowledge, this paper is among the first to empirically analyze the effects of e-cigarette advertising on demand for traditional cigarettes, e-cigarettes, and smoking cessation products.[3] First, I use store sales data and local advertising data to determine whether e-

[2] Advocates for a ban on e-cigarette advertising often bring up the gateway and renormalization effects in the context of teen consumption. The 2014 National Youth Tobacco Survey found that for the first time, middle and high school students used e-cigarettes more than any other tobacco product, including conventional cigarettes. However, middle and high school students did not increase their overall tobacco use between 2011 and 2014; the increase in e-cigarette use was offset by a decline in traditional cigarette and cigar use. Still, researchers are concerned about the long-term consequences of teenagers adopting e-cigarettes since surveys indicate that about 90% of current smokers first tried cigarettes as teens and that about 75% of teen smokers continue to smoke as adults (2012 Surgeon General's Report). My ability to study the important question of youth adoption of e-cigarettes is unfortunately limited by the availability of data on the nascent industry.

[3] Zheng, Zhen, Nonnemaker, & Dench (2016) estimate an AIDS demand model for cigarettes, e-cigarettes, and other tobacco products using monthly, market-level convenience store sales data and TV advertising data for 20 Nielsen markets in the U.S. They estimate a short-run e-cigarette TV ad elasticity of 0.008 and a long-run own-ad elasticity of 0.11. They also report a small positive effect of e-cigarette advertising on demand for tobacco cigarettes (long-run elasticity of 0.001). The paper does not discuss the potential endogeneity of advertising at all.

cigarette advertising increases or decreases demand for cigarettes. Identifying advertising effects can be challenging and is the focus of a large body of academic research. Randomization and instrumental variables are tools frequently used by researchers to identify causal effects of advertising. My strategy for identifying advertising effects is a hybrid regression discontinuity difference-in-differences approach based on the recent work of Shapiro (2018), and similar to the identification approaches taken by Card & Krueger (1994) and Black (1999). The idea is to take advantage of discontinuities in television market borders that lead similar individuals to be exposed to different levels of advertising. In this way, each border discontinuity can be thought of as a natural experiment through which we can learn about the causal effect of advertising.

I present difference-in-differences regressions which indicate that e-cigarette advertising increases demand for e-cigarettes and decreases demand for traditional cigarettes and smoking cessation products. After identifying advertising effects in the aggregate data, I use household purchase panel data to document the substitution patterns between e-cigarettes and traditional cigarettes. Household purchase patterns indicate that e-cigarettes are a substitute to traditional cigarettes. The household data also reveals a pattern consistent with addiction; current period demand for cigarettes is increasing in past consumption.

Finally, to quantify the effects of a proposed ban on e-cigarette advertising, I construct and estimate a model of demand for cigarettes that allows me to leverage the strengths of both aggregate and household data. The demand model aggregates in an internally consistent way, such that equations governing household and aggregate demand are functions of the same underlying structural parameters. The model enables me to utilize the information content of the two datasets in a unified way. I identify advertising effects off the aggregate data, while accounting for both heterogeneity in tastes and the persistence in choices generated by addiction. I estimate the model by adapting an integrated procedure proposed by Chintagunta & Dubé (2005) that recovers mean utility levels and unobserved demand shocks from aggregate data and identifies parameters governing heterogeneity off of household purchase data. I extend this procedure to a model with state dependence which allows me to identify addiction using the household data. Finally, I show how the discontinuities I exploit in the descriptive linear model port to the nonlinear structural model in an intuitive way, thus showing how to leverage the same identification in all model specifications. I then use the estimated model parameters to predict the impact of a ban on e-cigarette advertising.

In comparison, in this paper, I address the potential endogeneity of advertising using detailed weekly, market-level data on advertising intensity and an identification strategy that exploits across-market variation in advertising over time. Duke et al. (2014) document the increase in youth exposure to e-cigarette advertising, but they do not link this advertising exposure to purchase outcomes.

My research contributes to the ongoing policy debate as to whether e-cigarette TV advertising should be banned and suggests that a ban on e-cigarette advertising may have unintended consequences. More generally, my approach contributes to the study of advertising in categories with state dependence and to the analysis of substitution and complementarities in demand across categories. The methodology I develop to study this question is useful beyond just the study of addictive goods and can be used to estimate aggregate demand for any type of good that exhibits state dependent demand.

This paper contributes to a small but growing literature that seeks to better-understand the burgeoning e-cigarette market and how it relates to the more established market for tobacco cigarettes. Huang, Tauras, & Chaloupka (2014) measure own-price elasticities of e-cigarettes ranging between -1.2 and -1.9 and a positive cross-price elasticity of e-cigarette demand with respect to traditional cigarette prices. Friedman (2015) and Pesko & Currie (2016) study the effects of legislation that restricts the sale of e-cigarettes to minors. Both papers find that cigarette sales to minors increase when e-cigarette purchase age restrictions are put in place, providing further evidence of substitution between e-cigarettes and tobacco cigarettes. My results also indicate the substitutability of these products, providing convergent evidence on this point.

My work is also closely related to the stream of research that seeks to measure the effects of the 1971 ban on cigarette TV advertising (Ippolito, Murphy, & Sant (1979), Schneider, Klein, & Murphy (1981), Porter (1986), Baltagi & Levin (1986), Seldon & Doroodian (1989)). Despite extensive work in the area, researchers have come to mixed conclusions. Many studies conclude that the ban did not significantly reduce cigarette consumption, while others have found evidence that the marginal productivity of cigarette advertising fell after the ban (Tremblay & Tremblay (1995)). Researchers have pointed to but not resolved the potential endogeneity of advertising and advertising regulation, as well as firms' ability to substitute advertising to other media as factors that have complicated empirical analyses of the effects of the advertising ban. The majority of papers analyzing the 1971 ad ban were limited to using data on aggregate advertising expenditures. In this paper, I am able to address the endogeneity of advertising using detailed weekly, market-level data on advertising intensity and an identification strategy which exploits across-market variation in advertising over time.

Finally, my work relates to a large body of literature in economics that analyzes markets for addictive goods. In classic models of addiction, a good is considered to be addictive if past consumption of the good raises the marginal utility of present consumption. Researchers have empirically tested for addiction in the cigarette market and found strong evidence that current consumption is increasing in past consumption (Houthakker & Taylor (1970), Mullahy (1985)). In addition, many empiricists have applied myopic and forward-looking models of addiction to data in order to measure the responsiveness of demand for addictive goods to changes in price. Researchers have found that temporary price changes for addictive goods have little impact on demand. However, long-run responses to permanent price increases are substantially larger than short-run reductions in demand (Chaloupka & Warner (1999)). These results suggest that ignoring the addictive nature of demand for tobacco will lead to biased predictions of long-run responses to price changes. In this paper, I focus on the measurement of advertising effects, which may also be biased if the presence of addiction is not accounted for. I address this feature of the market by presenting a myopic model of cigarette addiction in which past consumption is complementary to current consumption.

In the sections that follow, I first describe the industry context and data in more detail. Then I discuss my identification strategy and present descriptive analyses of aggregate and household-level purchase data. Motivated by these results, the second half of the paper introduces a demand model for nicotine products and describes an integrated estimation procedure that utilizes both the aggregate and household data. I then use the demand estimates in a counterfactual analysis to predict the impact that a ban on e-cigarette TV advertising would have on cigarette demand. I conclude by summarizing the key findings and outlining directions for future research.

2 Empirical Setting

2.1 Tobacco Advertising Ban

In the mid to late 1960s, cigarettes were one of the most heavily advertised products on TV. Under pressure to reduce youth exposure to cigarette ads, in 1969 Congress approved the Public Health Cigarette Smoking Act, which effectively banned *all* advertising for cigarettes on TV and radio. The ban went into effect on January 1, 1971, and is still in effect today.

2.2 E-Cigarettes

In 2004, the Chinese company Ruyan introduced the world's first e-cigarette. The product entered the U.S. market soon after in 2007. An e-cigarette is an electronic device that contains a nicotine-based liquid. When heated, the liquid becomes a vapor which the user inhales. E-cigarettes do not contain tobacco and do not produce smoke because they do not use combustion. There are two main variants of e-cigarettes – a durable, re-usable product that can be recharged with included batteries and refilled with replacement cartridges, and a disposable product. Many e-cigarette companies sell both a refillable and a disposable device. Although e-cigarettes vary greatly in appearance, the most popular brands bear a close physical resemblance to traditional cigarettes. E-cigarettes are available in many flavor varieties including tobacco, mint, and cotton candy. Opponents to e-cigarettes argue that these flavors increase the product's attractiveness to youth.

Until early 2012, the e-cigarette market was composed of many small independent brands. In April 2012, Lorillard (the 3rd largest U.S. tobacco company) acquired Blu Ecigs for $135 million. They became the first of the Big Tobacco companies to enter the e-cigarette market. Reynolds (the 2nd largest U.S. tobacco company, now merged with Lorillard) launched its own brand Vuse in July 2013. Altria (the largest U.S. tobacco company) launched its own brand, MarkTen, in August 2013.

Compared to tobacco cigarettes, e-cigarettes are more loosely regulated. E-cigarettes are sold in retail stores and online and are not federally taxed, as are traditional tobacco cigarettes. Until early 2016, e-cigarette minimum purchase age restrictions were determined by state governments, with no federal age restrictions like those imposed on cigarette purchases. In May 2016, the FDA finalized a rule that extends its regulatory authority to include e-cigarettes. Among other changes, this regulation set 18 as a national minimum age to purchase e-cigarettes.

With the increasing popularity of e-cigarettes, a growing body of literature has developed around studying the health effects of e-cigarette use and second-hand exposure. The long-term health effects of e-cigarettes are still being investigated by clinical researchers, but initial studies seem to indicate that e-cigarettes appear to be less harmful than traditional cigarettes, but more harmful than abstaining from nicotine products altogether.[4] Most e-cigarettes contain nicotine, the highly addictive stimulant found in tobacco cigarettes that raises the heart rate, increases blood pressure, and constricts blood vessels (Benowitz & Gourlay (1997)). Long-term exposure to nicotine has been linked to hypertension and heart disease, including congestive heart failure and arrhythmias. Nicotine has also been shown to negatively affect the neurological development of adolescents and developing fetuses. E-cigarettes, however, do not contain tar and other cigarette residues that are the ingredients in traditional combustion cigarettes that

[4]For example, a recent report by the Royal College of Physicians asserts that e-cigarettes are only 5% as harmful as traditional cigarettes (Royal College of Physicians (2016)). The CDC website states "Are e-cigarettes less harmful than regular cigarettes? Yes – but that doesn't mean e-cigarettes are safe" (CDC (2018)).

have been shown to cause lung cancer.[5]

A second stream of research has explored whether e-cigarettes are an effective smoking cessation tool. Proponents of e-cigarettes argue that they deliver nicotine to the user without many of the harmful byproducts contained in tobacco smoke and that e-cigarettes may be a more effective smoking cessation aid than other existing products because they mimic the tactile and sensory process of smoking. Although e-cigarettes have not yet been approved as a smoking cessation device by any government agency, various public health organizations have acknowledged that e-cigarettes may prove to be an effective smoking cessation tool for some users.[6]

2.3 E-Cigarette Advertising

The primary goal of this paper is to determine the effect of e-cigarette advertising on demand for cigarettes. It is thus important to understand the messages that e-cigarette ads communicate to viewers. On one hand, e-cigarette advertising may reduce aggregate consumption of cigarettes by encouraging smokers to switch from traditional cigarettes to e-cigarettes and by highlighting the unappealing attributes of cigarettes.[7] Alternatively, e-cigarette ads could generate positive spillovers if they increase demand for the category of cigarettes as a whole or if they portray e-cigarettes as a complement to traditional cigarettes.

Matthew Myers, president of the Campaign for Tobacco-Free Kids, has expressed concern that “e-cigarettes are using the exact same marketing tactics we saw the tobacco industry use in

[5] Researchers are also interested in the effects of second-hand exposure to e-cigarette aerosol, which can help inform whether e-cigarette use should be regulated indoors as is the smoking of traditional cigarettes. E-cigarette aerosol is not simply water vapor. It contains chemicals including formaldehyde and acetaldehyde, though these chemicals are present at rates 9 to 450 times lower than in smoke from combustible cigarettes (Crowley (2015)).

[6] A 2015 report released by Public Health England concludes that electronic cigarettes “can help people to quit smoking and reduce their cigarette consumption" (McNeill et al. (2015)). As of 2016, the C.D.C. in the U.S. had taken an opposing stance, maintaining the position that “There is currently no conclusive scientific evidence supporting the use of e-cigarettes as a safe and effective cessation tool at the population level. The science thus far indicates most e-cigarette users continue to smoke conventional cigarettes" (Tavernise (2016)). However, in 2018, the CDC’s website was updated in a way that suggests that the organization is revising their stance on e-cigarettes. The website reads "e-cigarettes have the potential to benefit adult smokers who are not pregnant if used as a complete substitute for regular cigarettes and other smoked tobacco products... While e-cigarettes have the potential to benefit some people and harm others, scientists still have a lot to learn about whether e-cigarettes are effective for quitting smoking" (CDC (2018)). Based on the marginally positive but limited existing studies that explore the efficacy of e-cigarettes as a smoking cessation tool, The World Health Organization concludes that “the use of ENDS [electronic nicotine delivery systems] is likely to help some smokers to switch completely from cigarettes to ENDS” and that e-cigarettes may “have a role to play in supporting attempts to quit” for smokers who have previously attempted and failed to quit using other cessation aids.

[7] For example, a Blu e-cigarettes ad emphasizes "no odor, no ash, no tobacco smoke, only vapor."

Figure 1: E-Cigarette Ads Use the Same Marketing Tactics Used by Traditional Cigarette Ads

the 50s, 60s and 70s [...] The real threat is whether, with this marketing, e-cigarette makers will undo 40 years of efforts to deglamorize smoking." The Lucky Strike cigarette and Blu e-cigarette ads in Figure 1 illustrate the similarities in advertising tactics that have generated concern that e-cigarette advertising will hinder existing tobacco control efforts and renormalize cigarettes in society. Characteristics of these ads include asserting an independent identity and associating nicotine use with celebrities, fashion, and youth.

Ad spillovers may also arise if consumers either consciously or subconsciously confuse the product that is being advertised. For example, in the FIN advertisement on the left of Figure 2, the physical appearance of the product is virtually indistinguishable from that of a traditional cigarette. On the company website, FIN describes its product as an "electronic cigarette that looks and feels like a traditional cigarette." This physical similarity is important because it raises the possibility that viewers could misinterpret ads for e-cigarettes to be ads for traditional cigarettes. In an experimental study, Maloney & Cappella (2015) found that e-cigarette advertisements with visual depictions of people using e-cigarettes increased daily smokers' self-reported urge to smoke a tobacco cigarette relative to daily smokers who saw e-cigarette ads without visual cues. These results suggest that e-cigarette advertisements may generate positive spillovers and increase demand for traditional cigarettes.

Other e-cigarette ads, such as the Blu ad in Figure 2, inform consumers about the fact that e-cigarettes do not fall under most indoor smoking bans that apply to traditional cigarettes. The underlying message communicated by these ads is that you do not need to quit smoking,

Figure 2: E-Cigarette Ads May Generate Positive Ad Spillovers

you may continue to smoke cigarettes when permitted, and you can supplement your nicotine consumption with e-cigarettes when you are prohibited from smoking indoors or in public places. The additional nicotine consumption coming from supplemental vaping indoors may reinforce addiction and increase demand for cigarettes in the future. In short, these ads may increase demand for traditional cigarettes by suggesting that e-cigarettes are complementary to traditional cigarettes.

To summarize, to the extent that e-cigarettes act as a substitute to traditional cigarettes, e-cigarette advertising can decrease demand for cigarettes. To the extent that e-cigarette ads and usage generate positive spillover effects for traditional cigarettes either through renormalization or complementarities, e-cigarette advertising can increase demand for cigarettes. In the sections that follow, I explore both the net effect of advertising on cigarette demand as well as heterogeneity in this effect across markets.

3 Data

Ultimately, whether e-cigarette advertising increases or decreases demand for cigarettes is an empirical question. Data on both purchase volume and advertising intensity is necessary in order to tease out which effect of e-cigarette advertising dominates. I analyze retail sales data, household purchase panel data, and market-level TV advertising data collected by Nielsen. Each of these datasets is described in more detail below. In addition, I use yearly county population data from the U.S. Census Bureau and data on yearly changes to state cigarette excise taxes collected by the Campaign for Tobacco-Free Kids.

3.1 Retail Sales Data

The Nielsen database includes weekly store sales data reporting prices and quantity sold at the UPC-level. The data records sales of e-cigarettes, traditional cigarettes, and smoking cessation products. Store location is specified at the county level. The data is available from 2010–2015 and the sample is partially refreshed annually.[8]

There are 64 brands and 540 unique e-cigarette UPCs recorded in the retail sales data. These UPCs are a mixture of rechargeable starter kits, refill cartridges, and disposable e-cigarettes. Starter kits cost between $30–50, refills (sold in 3–5 cartridge packs where each cartridge is roughly equivalent to 1–2 packs of cigarettes) cost between $10–15, and disposable e-cigarettes (equivalent to 1.5–2 packs of cigarettes) cost about $10. In all subsequent analyses, I focus on sales of refill cartridges and disposable e-cigarettes because i) these products account for 96% of unit sales and 89% of e-cigarette dollar sales, ii) quantities are more clearly indicated in the data and iii) these products have similar prices from which I can construct an aggregate price series.

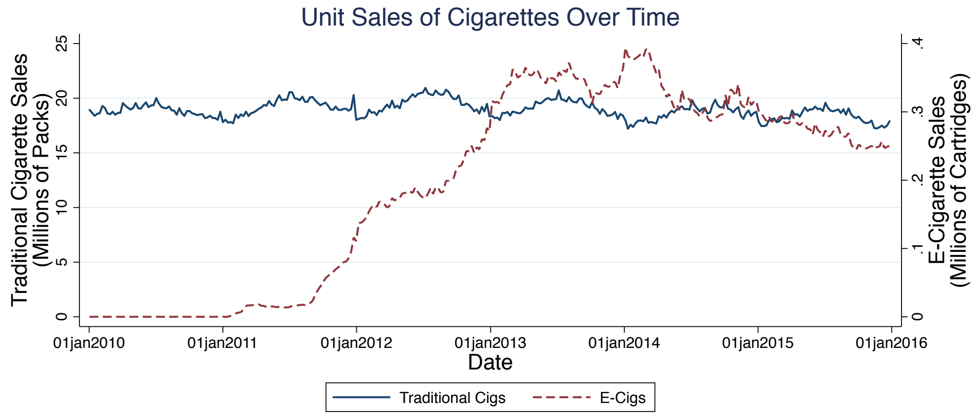
The Nielsen data records sales of 455 tobacco cigarette brands and 9,257 unique UPCs. Cigarettes are sold primarily as packs (20 cigarettes in a pack) and cartons (10 packs in a carton). Federal, state, and local cigarette excise taxes are reflected in the prices in the data. Differences in state and local excise taxes create extensive variation across markets in the average price of a pack of cigarettes. The quantity-weighted average price of a pack of cigarettes across all stores in the panel is $5.61, but this price varies across counties from a low of $3.16 in Barton County, MO to a high of $10.66 in Bronx County, NY.

Figure 3 plots the trend in aggregate cigarette and e-cigarette sales over time for the 31,634 stores who are active in the panel each year between 2010–2015. E-cigarette sales were low until mid 2011, after which the quantity of units sold began to grow rapidly. The plot shows that there is seasonality in the quantity of cigarette packs sold with lower sales during the winter and higher sales during summer months.

Finally, turning to cessation products, Nicorette gum and Nicoderm CQ patches are the dominant products in the category, accounting for 97% of store sales. Nicorette gum is most frequently purchased in packs of 20 and 100 pieces of gum, while Nicoderm CQ patches are

[8] Each year the retail data tracks sales from approximately 35,000 individual stores pertaining to roughly 90 retail chains. As of 2011, estimated coverage as a percent of all commodity volume by channel was: Food (53%), Drug (55%), Mass Merchandise (32%), and Convenience Store (2%).

Figure 3: Trend in Weekly Sales of Cigarettes and E-Cigarettes



most commonly purchased in packs of 7 and 14 patches. Nicorette is intended to be used as needed to satisfy nicotine cravings, while each Nicoderm CQ patch is intended to last a whole day. Ten pieces of gum and a single nicotine patch have roughly an equivalent amount of nicotine to a pack of cigarettes. The quantity-weighted average price per unit is $0.46 per piece of gum and $3.34 per patch.

3.2 Household Purchase Data

Nielsen also collects daily UPC-level purchase data for a sample of approximately 50,000 U.S. households.[9] Purchases of e-cigarettes, traditional cigarettes, and smoking cessation products are recorded. The data reports price paid, number of units purchased, and, when available, identifying information for the store at which the purchase was made. Like the store sample, the household sample is partially refreshed annually.

Between 2010–2015, 2,288 households made a total of 10,962 purchases of any type of e-cigarette. Of the 895 of these households who are tracked in the panel for all six years, 84% of households are observed to buy traditional cigarettes before buying e-cigarettes for the first time, 3% of households report purchasing e-cigarettes before later making a purchase of traditional cigarettes for the first time, and the remaining 13% of households never report any purchases of traditional cigarettes. It is these latter two groups that policy makers are especially worried about.

[9]Nielsen strives to ensure that the panel provides the best possible representation of all demographic groups. However, Nielsen documentation notes that young people are hard to recruit and maintain. Further, purchases cannot be linked to specific individuals within multi-person households. These data limitations inhibit my ability to study youth and young-adult consumption of e-cigarettes.

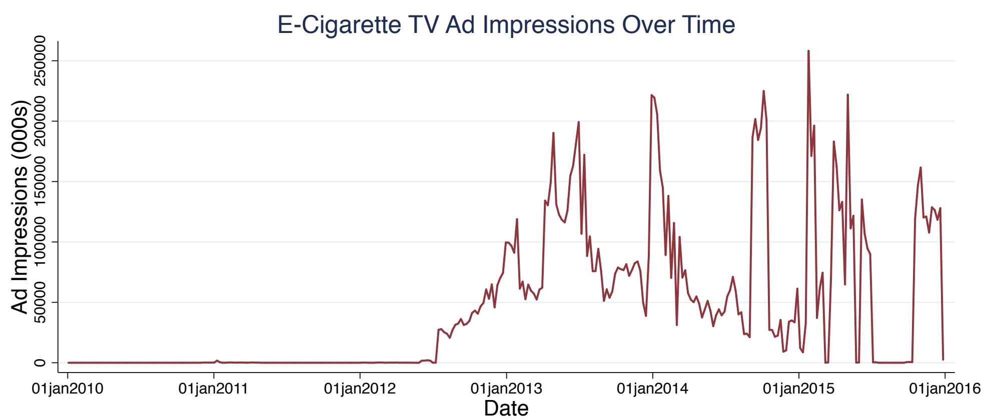
3.3 Advertising Data

Weekly, product-level television advertising data from 2009–2015 comes from Nielsen. The data records ad impressions, units, expenditures, and gross rating points (GRPs). GRPs are a measure of advertising intensity, calculated as exposures per capita times 100.

Figure 4 plots the trend in total e-cigarette ad impressions over time. There was very little advertising until mid 2012, at which point the number of ad impressions began to grow quickly. Firms buy advertising at both the national and local DMA level.[10] Although the majority of advertising is bought nationally, about 20% of ad-spending is on local advertising.

The data records advertising for e-cigarette brands as well as smoking cessation products. Table 1 reports market shares for the top e-cigarette and smoking cessation brands. From 2010 to 2015, Blu was the market leader amongst e-cigarette brands with 55% of e-cigarette store sales and 59% of all e-cigarette ad impressions. Lorillard acquired Blu in April 2012, shortly before the observed spike in advertising in mid 2012. Nicorette and Nicoderm CQ account for 94% of the advertising for smoking cessation products.

Figure 4: Trend in E-Cigarette TV Ad Impressions



4 Descriptive Analysis

In this section I explore the purchase and advertising data further in order to better understand the role of advertising in the market and to identify the substitution patterns between e-cigarettes and traditional cigarettes. First, using market-level data I show that e-cigarette

[10]Cable, Network, and Syndicated advertising is purchased at the national level while Spot advertising is purchased at the local level.

Table 1: E-Cigarette and Smoking Cessation Brands by Market Share (2010–2015)

	Market Share	Ad Impression Share
Blu (Lorillard)	54.5%	58.5%
Vuse (RJ Reynolds)	2.1%	30.8%
NJOY	8.6%	6.2%
Fin	11.4%	3.0%
Other	23.4%	1.5%
Total	$366,429,000	13,899,376,000

Nicorette	78.8%	52.5%
Nicoderm CQ	18.6%	41.0%
Other	2.6%	6.5%
Total	$948,970,000	62,287,168,000

advertising increases demand for e-cigarettes and decreases demand for traditional cigarettes and smoking cessation products. Next, I illustrate the substitution patterns between traditional and e-cigarettes and show patterns that are consistent with addiction using the household purchase data.

4.1 Identifying Advertising Effects with Aggregate Data

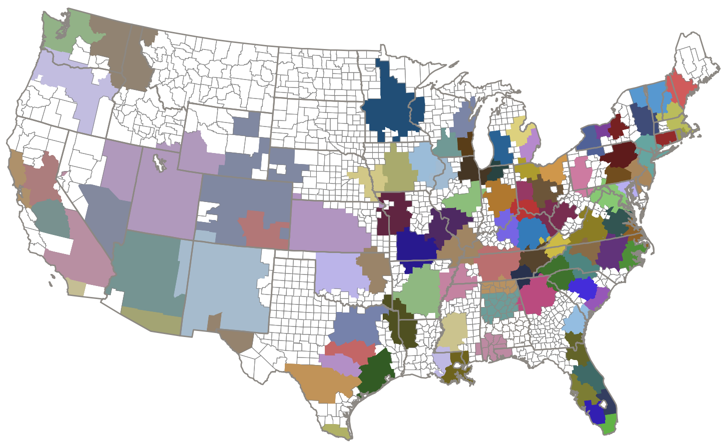
4.1.1 Identification Strategy for Ad Effects

I am ultimately interested in measuring the causal effect of e-cigarette advertising on demand. Identifying the causal effect of advertising is complicated by the fact that local advertising is not assigned randomly. The concern is that firms might target higher levels of advertising to markets and time periods with high demand. If not accounted for, this endogeneity would lead to biased estimates of the effects of e-cigarette advertising.[11]

I address this endogeneity concern by exploiting a discontinuity in local advertising markets that was first pointed out by Shapiro (2018). Nielsen delineates local television markets or Designated Market Areas (DMAs) by grouping counties based on their predicted interest in TV program content and quality of over-the-air TV signal. Firms buy local advertising at the DMA level, so all households residing in a given DMA see the same television programming and

[11] Appendix A presents county-level regressions with common week fixed effects as a comparison to the border strategy results. The ad elasticities in the county-level regressions are slightly biased in the positive direction relative to the following border analysis.

Figure 5: Top 100 DMAs



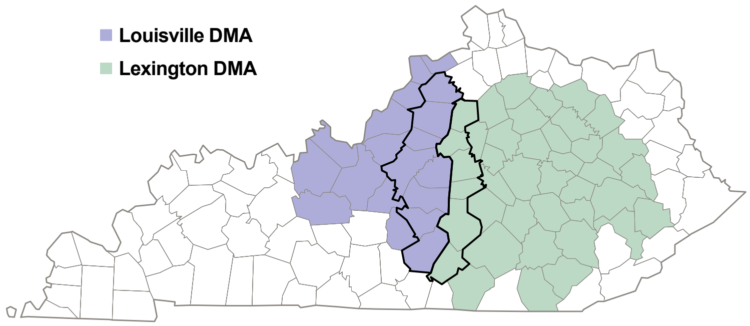
ad content.[12] Thus, if advertisers do not uniformly buy advertising across DMAs, households on opposite sides of a DMA border can be exposed to different levels of advertising. I refer the reader to Shapiro (2018) for a thorough discussion of television advertising markets.

Identification comes from comparing sales in counties just to the left of a border to sales in counties just to the right of the border over time. I aggregate store sales to the county level because county is the finest level of geographic variation I observe in the store sales data. The identifying assumption is that these border counties experience the same unobserved demand shocks, and thus, in the absence of an advertising intervention, sales in these bordering markets would follow the same trend. This strategy is analogous to the approaches used in important early studies on program evaluation including Card & Krueger (1994)'s study of minimum wage effects and Black (1999)'s analysis of the economic value of education. However, while Card and Krueger use state boundaries and Black looks across school district attendance boundaries, DMA boundaries do not necessarily coincide with state or other geo-political boundaries that we worry would likely be correlated with advertising and demand for cigarettes. A map of the top 100 DMAs ranked by viewership is shown in Figure 5.

DMAs tend to be centered around cities, while the borders between DMAs typically fall in more rural areas. Firms tend to set advertising for a given DMA based on the urban center of the DMA, where the majority of the population resides. This suggests that we might see different levels of advertising at the border between two DMAs, but that these differences are not being driven by differences in the characteristics of households in these rural border areas. The intuitive way to think about identification here is that the individuals living on either side

[12]Although nearly all households now watch TV using cable or satellite dish as opposed to watching over-the-air, it is still the case that television providers show households within a given DMA the same TV content and ads.

Figure 6: Louisville and Lexington DMA Border Counties



of a border are similar on unobservables, but they are exposed to different levels of advertising because of differences in the major cities located at the centers of their respective DMAs. If this is true, then we can think of each border as a natural experiment with two treatment groups. In the absence of differences in advertising, we would expect demand to follow the same trend on either side of the border. Thus, ad effects will be identified off of the covariance between differences in advertising and deviations from the common trend in sales.

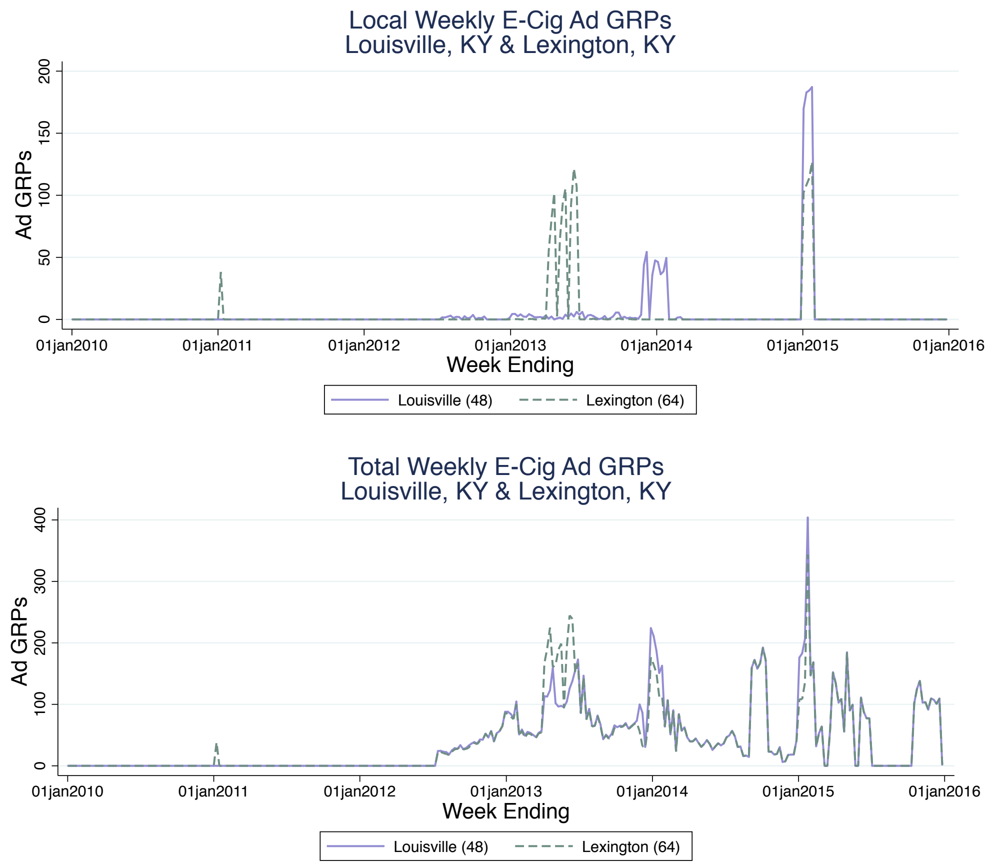
Take, for example, the border between the Louisville, KY and Lexington, KY DMAs shown in Figure 6. There are 8 counties in the Louisville DMA that share a border with a county in the Lexington DMA and 6 counties in the Lexington DMA that share a border with a county in the Louisville DMA. The population of these border counties makes up a small share of the total population of the corresponding DMAs; the border county population share of the Louisville and Lexington DMAs are 10% and 12% respectively. I focus on borders between the top 100 DMAs, resulting in 149 borders. After restricting to the markets that contain at least one store selling the focal products, I am left with 141 borders and 282 border-markets. The median and mean border county population shares across these border-markets are 9% and 17% respectively.

The identification strategy relies on the extent to which there is variation in advertising intensity both across borders and over time. For example, Figure 7 plots the local and total (local plus national) weekly e-cigarette ad GRPs in the Louisville and Lexington DMAs and shows that there is variation in both the intensity and time in which the two DMAs are exposed to advertising. Table 2 reports statistics summarizing the variation in advertising for the entire border sample.[13] The average difference in average weekly advertising across each pair of border markets[14] is 3.9 GRPs, confirming that there is a discontinuity in advertising across

[13]Statistics reported for the period May 2012 – Dec 2015, the period in which the vast majority of e-cigarette advertising occurs (see Figure 4).

[14]$\Delta a_b = |\bar{a}_{bm_1} - \bar{a}_{bm_2}|$ where $\bar{a}_{bm_1} = \frac{1}{T}\sum_{t=1}^{T} a_{bm_1 t}$

Figure 7: Local and Overall Variation in E-Cigarette Ad GRPs in the Louisville and Lexington DMAs



neighboring DMAs. The coefficient of variation calculated for each market as the standard deviation in weekly ad GRPs divided by the mean weekly GRPs is large and shows that there is within-market variation in advertising over time. Note that in this analysis and all analyses going forward, advertising refers to total advertising (local plus national). Figure 7 shows that in some weeks, one DMA may be exposed to more advertising than its neighbor, and in other weeks, the opposite may occur. In order to quantify this variation, I calculate the absolute value of the difference in weekly ad GRPs for each pair of bordering DMAs.[15] In 63% of the 26,881 week-border observations, both sides of the border are exposed to the same intensity of e-cigarette advertising. In the remaining 37% of observations there is significant variation in the magnitude of the ad differential. This variation is summarized in the last row of the top panel in Table 2. In more than 10% of observations, the difference in treatment is greater than 10%

[15] $|\Delta a_{bt}| = |a_{bm_1t} - a_{bm_2t}|$

Table 2: Variation in Advertising for the Border Market Sample

	N	Min	Median	Mean	Max
Ave Weekly E-Cigarette GRPs	282	59.6	61.5	63.0	100.7
Difference in Ave Weekly E-Cigarette GRPs	141	0.0	2.4	3.9	40.8
Coeff Var E-Cigarette GRPs	282	0.76	0.83	0.85	1.04
Abs Difference in Weekly E-Cigarette GRPs	9,823	0.0	2.0	14.6	344.2
Ave Weekly Smoking Cessation GRPs	282	158.6	163.3	164.4	204.5
Difference in Ave Weekly Smoking Cessation GRPs	141	0.0	2.2	3.5	40.7
Coeff Var Smoking Cessation GRPs	282	0.80	0.89	0.89	0.92
Abs Difference in Weekly Smoking Cessation GRPs	17,463	0.0	4.4	9.3	285.3

of the average treatment.[16] The bottom panel of the table reports the analogous variation in smoking cessation advertising. Notably, smoking cessation products are advertised at a higher intensity, but there is slightly less variation in this advertising across borders. Together, these statistics confirm that the data contains significant variation in advertising that can be used to identify the effect of ads on product sales.

Recall that the identifying assumption is that sales on either side of a border would follow the same trend in the absence of an advertising intervention. To explore whether this assumption is credible, I compare the trend in cigarette sales in border markets before e-cigarette companies began to advertise. Figure 8 plots the total number of packs of cigarettes sold in the border counties in the Louisville and Lexington DMAs in 2010. In the absence of differences in e-cigarette advertising, sales in the two markets seem to follow the same trend. The correlation in 2010 sales is $\rho = 0.54$. Figure 9 plots a histogram of the correlation in 2010 sales across all border markets. The median border in the sample has a correlation in weekly cigarette sales in 2010 of $\rho = 0.53$. However, sales in a small number of bordering markets are un-correlated or even negatively correlated in 2010. In Appendix B I test the sensitivity of the results to the common trends assumption by restricting the sample to only the set of borders with a correlation in 2010 sales above $\rho = 0.5$. The results are consistent in sign and the magnitude of the ad effects becomes slightly larger.

Another way of assessing the validity of the common trends assumption in a differences-in-differences model is to run a placebo check. In Appendix C, I conduct a placebo analysis using advertising data for a seemingly unrelated CPG brand, Angel Soft toilet paper. As expected, the placebo analysis produces insignificant ad effects.

[16] Average weekly e-cigarette GRPs are 63. For 2,889 week-border observations, one DMA is exposed to least 6.3 more GRPs than its neighbor DMA.

Figure 8: Weekly Packs of Cigarettes Sold in Louisville and Lexington DMA Border Counties in 2010

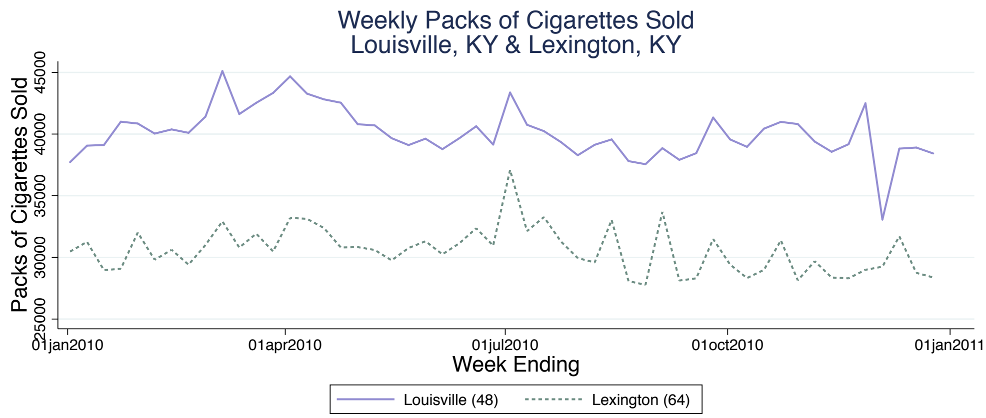
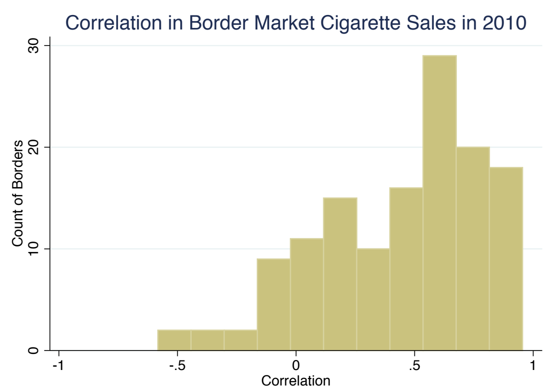


Figure 9: Distribution of Correlation in Weekly Cigarette Sales Across Borders in 2010



4.1.2 Potential Threats and Limitations of the Border Strategy

First, note that the key identifying assumption has only to do with common *trends* and that time invariant differences across bordering markets will ultimately be absorbed by a set of market fixed effects.[17] Thus, the identifying assumption would only be violated if there were an unobserved shock on one side of the border that was correlated with both sales and advertising. I could think of two such shocks that could differentially affect one side of the border and be correlated with sales of cigarettes and advertising for e-cigarettes: i) changes to local excise taxes and ii) changes to local indoor smoking legislation. If a county on one side of a DMA border increased cigarette excise taxes, cigarette sales would have fallen on that side of the

[17]This is true if ad-responsiveness is not a function of population characteristics. To the extent that ad-responsiveness is a function of characteristics, I can check that the bordering markets have similar demographics. This comparison is reported in Appendix E.

border in response to the price increase and e-cigarette companies might have increased their advertising to that DMA. Similarly, if a county on one side of a DMA border approved more stringent indoor smoking bans, demand for cigarettes might have fallen in response to the increased inconvenience of smoking and e-cigarette companies might have increased their advertising to that market. Cigarette excise taxes are included in the prices in my data, so these changes are actually observed. However, as an additional sensitivity check, I tried dropping borders that fall in states that increased their cigarette excise tax during the period 2011-2015 (Campaign for Tobacco-Free Kids (2017)). The results, presented in Appendix D, are similar to the main specification.

Second, an impediment to the identification strategy could arise if cigarette companies strategically respond with their own marketing spending. According to the FTC, the major cigarette manufacturers spent $9.2 billion on cigarette advertising and promotion in 2012. Price discounts paid to cigarette retailers to reduce the price of cigarettes to consumers made up the largest share (85%) of marketing spending (FTC (2015)). These discounts will be reflected in the prices in my dataset and will thus be controlled for in the empirical analysis. The Nielsen advertising database records print advertising expenditures for cigarette companies, but the vast majority of this spending is at the national level. I expect its effect to be uniform on either side of DMA borders and unlikely to be a problem for my identification strategy.

Third, to the extent that consumers watch advertisements in one DMA and make purchases in a neighboring border DMA, the effects I estimate would be attenuated towards 0. In order to explore the extent to which this might be a concern, I take the set of Nielsen households who ever buy cigarettes or e-cigarettes in the household purchase panel and who reside in DMA-border counties, and I calculate the fraction of cigarette or e-cigarette purchases that were made by these households in a DMA other than the one in which they reside.[18] Of the 84,239 cigarette or e-cigarette transactions made by these panelists in stores with location information, only 3% of transactions were made outside of the household's DMA of residence. Given the low incidence of cross-DMA shopping observed in the data, I expect that any attenuation bias due to cross-DMA shopping is likely to be very small.

Finally, the question of the external validity of these estimates must be raised. This border discontinuity identification strategy allows me to measure unbiased causal effects of e-cigarette advertising for a specific sub-population of individuals who reside in border markets.

[18]Not all transactions in the household panel include a store identifier that can be used to identify the county in which the store is located. Thus, this analysis only includes the transactions made at stores for which I observe location information.

The policy question at issue is how demand would respond to a nationwide ban on e-cigarette TV advertising. Thus, when drawing inference from these estimates, it is important to keep in mind how these markets differ from the overall population in the U.S. In Appendix E, I use U.S. Census data to explore the differences in demographics between border counties and non-border counties. I find that individuals residing in border counties on average are slightly older, less educated, and have lower income. Border counties have a lower share of black residents and a lower population density compared to non-border counties. Research by the American Lung Association (Shan, Jump, & Lancet (2012)) shows that rural areas tend to be associated with higher rates of adult and adolescent smoking, and that youth in rural areas tend to start smoking at a younger age. This suggests that to the extent that I am measuring advertising effects for a specific sub-population, this sub-population may be one that policy-makers are especially concerned about. Notably, because rural areas tend to have high smoking rates, these may be areas where the substitution effect has the potential to be large, and thus the generalizability of the results should be considered with this caveat in mind.

4.1.3 Identification of Price Effects

Thus far, the discussion has focused on the identification of ad effects because this is the central focus of this paper. Before introducing the model, a brief discussion of price endogeneity is also warranted. Similar to the case with advertising, a long literature in economics and marketing has pointed out the potential endogeneity of prices, whereby prices may be coordinated with demand shocks that are unobserved to the econometrician. Depending on the nature of the correlation, if not accounted for, this can lead to either an over- or under-estimate of price elasticities. My main strategy to account for endogeneity of prices is to include a rich set of market and time fixed effects. Specifically, the same DMA-border and border-week fixed effects that isolate variation in advertising across borders over time will also isolate similar variation in prices. Naturally, it is important to consider i) whether variation in prices exists after controlling for these fixed effects, and ii) what kind of price variation would be problematic for this identification strategy.

I construct market-week level price series for tobacco cigarettes, e-cigarette cartridges, and smoking cessation products by taking a weighted-average over stores and UPCs. If these prices were identical in neighboring border markets, then the price coefficient would not be separately identified from the border-specific week fixed effects. In Table 3, I summarize the observed variation in cigarette prices across markets, the variation in prices within border markets over time, and the variation in prices that remains after netting out the DMA-border

Table 3: Variation in Cigarette Prices for the Border Market Sample

	N	Min	Median	Mean	Max
Ave Weekly Price Per Pack	282	3.74	5.18	5.39	9.05
SD in Price Per Pack Over Time	141	0.08	0.20	0.25	0.84
SD in Price Per Pack Net of FEs	141	0.01	0.04	0.06	0.32

and border-week fixed effects that are included in the model specified in equation 1. Comparing the last two rows of the table, the fixed effects clearly absorb a significant fraction of the variation in prices that exists across border markets and over time. However, some variation still exists net of these fixed effects. Turning to the second question of whether this variation is problematic, a threat to my proposed identification strategy would require a time-varying shock to demand that is unique to one side of a border, and would require retailers to adjust their prices in response to this shock. Exploring the nature of the price variation, I find that on either side of a border, the price of a given UPC is similar at locations of the same chain, and that prices vary more systematically across chains.[19] This is consistent with the findings of Hitsch, Hortaçsu, & Lin (2017) and helps assuage the concern that local prices might be set based on the DMA's level of advertising. With this understanding of the underlying price variation in the data, I feel confident moving forward under the assumption that the rich set of DMA-border and border-week FEs included in the model sufficiently addresses any concerns about the endogeneity of prices.

4.1.4 Fixed Effects Regressions

In this section I discuss the implementation of the identification strategy and then present the estimation results. At a high level, the approach is to only use data for border markets and to include a rich set of market and border-time fixed effects that allow markets to have different levels of sales and border-specific flexible time trends. I describe these steps below in the context of the descriptive analysis. I later describe in Section 6 how to implement this border discontinuity approach within the context of a more complex non-linear model.

First, the sample is restricted to the set of stores that were active in the full panel from 2010–2015 and are located in a border county. All counties in a given DMA on a given border

[19]Borders that coincide with boundaries between states that have different excise taxes are the exception. However, because these taxes are directly observed in the prices in the data, these differences are controlled for. Appendix D confirms the robustness of the results to dropping borders that include states that changed their excise taxes during the period of study.

are grouped together into a market. For example, the 8 counties in the Louisville DMA that border the Lexington DMA form a market and sales in stores in these counties will be aggregated to form total market sales. The 6 counties in the Lexington DMA that share a border with a county in the Louisville DMA make up the comparison market. The dependent variables of interest are total number of cartridges of e-cigarettes sold and total number of packs of cigarettes sold by stores in each market each week. I focus on sales of refill cartridges and disposable e-cigarettes because these products have similar prices and are a better measure of e-cigarette consumption.[20] To construct price series for each market from the store sales data, I calculate the weighted average price for a pack of cigarettes and price per cartridge of refill and disposable e-cigarettes. I also look at sales of nicotine patches and gum, and I construct the price series for these products as the average price per unit for a patch and piece of gum.

I implement the identification strategy by including a set of market fixed effects and a set of border-week fixed effects. The market fixed effects control for time invariant differences across markets and allow each market to have its own average level of sales. Border-week fixed effects allow each border to have its own flexible trend in sales that will capture the observed seasonality in cigarette sales and will, for example, allow the specific seasonality pattern to differ between borders in New York and borders in Florida.[21]

The difference-in-differences specification is shown in equation 1. The unit of observation is a market-border-week where m denotes market, b denotes border, and t denotes week. Advertising for e-cigarettes and smoking cessation products is denoted by a^e_{mt} and a^q_{mt} and prices for all four products are collected in a vector $\boldsymbol{p}_{mt}$.[22] Equation 1 is estimated separately for e-cigarettes, cigarettes, nicotine patches, and nicotine gum via OLS. Table 4 presents the estimation results.

$$Q_{mt} = \beta_m + \beta_{bt} + \phi_e \log(1 + a^e_{mt}) + \phi_q \log(1 + a^q_{mt}) + \alpha^T \boldsymbol{p}_{mt} + \epsilon_{mt} \tag{1}$$

[20] E-cigarette cartridges are most commonly sold in packs of 3–5.

[21] I regress the log of e-cigarette advertising on the full set of market and border-week fixed effects to confirm that there is sufficient variation in the advertising data to permit this granular level of fixed effects. The mean of the residuals is 0 and the standard deviation is 0.14.

[22] Advertising enters within a log to account for decreasing returns to scale. Models that are linear in advertising produce results that are directionally consistent. I also estimated models that allow past advertising to affect current sales. Appendix F presents ad stock models assuming various carry-over rates. These models produced similar elasticities to the ad flows model presented in Table 4. These results that explicitly control for past advertising show that the estimated ad effect is not biased by a correlation between current and omitted past advertising, which may also be correlated with the stock of addicted consumers in the previous period. These robustness checks reinforce the main finding that e-cigarette ads do not (in the short-run) lead to an increase in cigarette purchases.

First, looking at the first column in Table 4, the positive and significant coefficient on e-cigarette advertising indicates that, as expected, advertising for e-cigarettes increases demand for e-cigarettes. Increasing average e-cigarette advertising by 10% results in a 0.8% increase in sales relative to the mean quantity of e-cigarettes sold. The effect of advertising for the Nicorette and Nicoderm CQ smoking cessation products is not significantly different from 0. The e-cigarette price coefficient is negative and significant as expected. The cigarette and nicotine patch cross-price coefficients are estimated to be positive and statistically significant, suggesting that these products are substitutes to e-cigarettes.

Column 2 of Table 4 regresses the number of packs of cigarettes sold in each market on the set of independent regressors and fixed effects. In column 2 there is a negative and significant effect of e-cigarette advertising on demand for traditional cigarettes. Contrary to all of the arguments that have been made as to why e-cigarette advertising might increase cigarette sales, I find evidence that e-cigarette advertising is actually *decreasing* demand for traditional cigarettes. The magnitude of this effect does appear small (a 10% increase in e-cigarette advertising is associated with a 0.2% decrease in sales relative to the mean quantity of cigarettes sold), but it is economically significant when compared to the fact that volume sales of traditional cigarettes were decreasing by 2.2% per year during this period (FTC (2015)). Furthermore, the positive coefficient on e-cigarette price provides additional evidence that smokers treat e-cigarettes as a substitute to traditional cigarettes. The coefficient on advertising for smoking cessation products is negative and much smaller in magnitude compared to the point estimate for e-cigarette advertising.[23] The estimates imply an own-price elasticity of -1.9 for traditional cigarettes, which is larger than the range of cigarette price elasticities of -0.4 and -0.8 that have been found in previous work (Chaloupka (1991)).[24] The entry into the market of e-cigarettes, a perhaps close substitute to traditional cigarettes, could explain this increase in the price-elasticity of cigarettes. To check this hypothesis, I use data from 2010–2011 to estimate the price elasticity of cigarettes before e-cigarette sales took off. I estimate an own-price elasticity of -0.9 during this period, which is more in-line with estimates from previous studies.[25]

[23]The effects of Nicorette and Nicoderm CQ advertisements remain largely insignificant in the models with less granular fixed effects reported in Appendix A.

[24]Because prices have been aggregated to the market and category level, this elasticity should be interpreted as a market/category level elasticity, not a UPC-level elasticity. This is similar to other estimates of cigarette elasticities that use changes in taxes to estimate how aggregate demand changes in response to a change in the average price of cigarettes.

[25]The estimates from the full model with heterogeneity and addiction imply an even smaller average price elasticity of -0.64 for tobacco cigarettes (Table 6).

Table 4: Difference in Differences Regression Results

	(1) E-Cig Cartridges	(2) Cigarette Packs	(3) Total Nicotine	(4) Nicotine Patches	(5) Nicotine Gum
E-Cigarette Log Ads	29.77***	-631.7***	-572.1***	-7.871***	-116.8***
	(5.644)	(217.4)	(216.5)	(2.866)	(43.14)
Smoking Cessation Log Ads	-6.005	-28.19	-40.20	-	-
	(4.473)	(83.81)	(84.44)	-	-
Nicotine Patch Log Ads	-	-	-	3.651	13.74
	-	-	-	(3.854)	(48.18)
Nicotine Gum Log Ads	-	-	-	-2.679	41.84
	-	-	-	(3.129)	(45.82)
Price E-Cigarette Cartridge	-8.166***	68.48***	52.14***	0.942***	2.889
	(0.988)	(10.89)	(11.61)	(0.246)	(4.525)
Price Cigarette Pack	86.98***	-10,128***	-9,954***	17.55***	-371.5***
	(12.24)	(826.5)	(827.5)	(5.219)	(76.55)
Price Nicotine Patch	7.013***	-47.87	-33.84	-10.12***	-22.81**
	(1.980)	(38.28)	(38.27)	(0.819)	(11.61)
Price Nicotine Gum	-25.23**	87.47	37.00	-20.31***	-1,971***
	(12.12)	(255.2)	(254.8)	(7.186)	(94.83)
DMA-Border FE	Y	Y	Y	Y	Y
Week-Border FE	Y	Y	Y	Y	Y
N Obs	63,952	63,952	63,952	63,952	63,952
Mean D.V.	381	28,414	29,176	219	6,961
E-Cigarette Ad Elasticity	0.08	-0.02	-0.02	-0.04	-0.02

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

The observed increase in e-cigarette consumption and decrease in cigarette consumption as a result of e-cigarette advertising raises the question, what happens to total nicotine consumption in response to an increase in e-cigarette advertising? Under the conservative assumption that each e-cigarette cartridge is equivalent to 2 packs of cigarettes in terms of nicotine content, I calculate total nicotine consumption as the total number of "equivalent" packs of cigarettes and e-cigarettes purchased in each market-week.[26] Column 3 reports how total nicotine consumption varies in response to e-cigarette advertising. An increase of 1 e-cigarette ad GRP results in a net decrease in total nicotine consumption, and some of that nicotine consumption is now coming in the less harmful form of e-cigarettes. This result indicates that there is not one-to-one substitution from cigarettes to e-cigarettes within stores tracked in the sample. This could be because some consumers are reducing their nicotine consumption, or it could be due to differential coverage of cigarette and e-cigarette sales in the Nielsen data.[27] In either case, when interpreting these results it is important to keep in mind that nicotine itself is not the component of tobacco cigarettes that has been strongly linked with adverse health effects and mortality. The medical literature is careful to draw this distinction. For example, Benowitz & Gourlay (1997) notes, "It is important to recognize that cigarette smoke is a complex mixture of chemicals that includes not only nicotine but also potentially cardiotoxic substances, such as carbon monoxide, oxidant gases and polycyclic aromatic hydrocarbons. The role of nicotine, if any, in causing acute or chronic cardiovascular disease has not been definitely demonstrated" (pp. 1422-1423). Thus, while the documented decrease in total nicotine consumption is interesting and suggestive about consumption patterns, the reduction in purchases of tobacco cigarettes shown in column 2 is more concrete and by itself an important finding for health policy.

The analysis thus far has considered the effect of e-cigarette advertising on demand for cigarette products. Given that the results suggest that consumers treat e-cigarettes as a substitute to traditional cigarettes, it is also informative to look at the effect of e-cigarette advertising on demand for traditional nicotine replacement therapies – the nicotine patch and nicotine

[26]Nicotine content per cigarette pack and per e-cigarette cartridge may vary across brands. I abstract away from these differences for the purpose of this "back of the envelope" analysis.

[27]While the Nielsen data has coverage of purchases of e-cigarettes made in traditional retail channels, the data does not record purchases of e-cigarettes made at local "vape" shops. In a note published in August 2015, Wells Fargo analyst Bonnie Herzog writes "Because a large portion of VTM [vaporizer, tank and mod] sales occur online and in vape shops – neither of which are tracked by Nielsen – the Nielsen data is no longer capturing the full e-vapor category. [...] While Nielsen's data is useful directionally we believe the e-cigarette unit and pricing data remains difficult to rely on given Nielsen is not yet reporting 'equivalent' units in this category" (Haar (2015)). Thus, it is possible that the Nielsen data is underestimating the increase in e-cigarette consumption as a result of e-cigarette advertising.

gum. Columns 4 and 5 present the regression results for the nicotine patch and gum products. The dependent variables are number of nicotine patches and number of pieces of nicotine gum sold in a market-week. I find that e-cigarette advertising has a business stealing effect on these smoking cessation products. The coefficient on e-cigarette advertising is negative and has a statistically significant effect on demand for both nicotine patches and gum. Additionally, the coefficient on e-cigarette price is positive. These results indicate that consumers are using e-cigarettes as a substitute to the nicotine patch and gum. This could be a concern for policy makers because it suggests that e-cigarette advertising shifts consumers away from clinically proven smoking cessation aids to e-cigarettes, which have not yet been proven to be effective in helping smokers quit. In columns 4 and 5, I separate out advertising for the patch and gum in order to capture any cross-product effects. Again, I don't find any significant advertising effects for these products. Interestingly, the cross-price effects between cessation products are negative, suggesting that nicotine patches and gum may be complements.[28]

Together these results lead to the following conclusions. (1) E-cigarette advertising increases demand for e-cigarettes and reduces demand for traditional cigarettes and smoking cessation products. (2) Consumers treat e-cigarettes, traditional cigarettes, and smoking cessation products as substitutes. In the next section, I further explore the substitution patterns between products using household purchase panel data.

4.2 Substitution Patterns and Addiction in Household Data

Thus far, the aggregate data indicates that e-cigarette advertising increases demand for e-cigarettes and reduces demand for traditional cigarettes. Cross-price effects also suggest that e-cigarettes and tobacco cigarettes are substitutes. In this section, I examine household panel data to explore whether households increase or decrease their consumption of cigarettes after buying e-cigarettes, and whether there are patterns consistent with nicotine addiction. Relative to the aggregate data, the household data is more transparent in revealing these substitution patterns over time. The exploratory analysis below is intended to motivate the modeling assumptions made in Section 5.

I analyze the weekly purchases of cigarettes and e-cigarettes for the 25,159 households who ever buy a cigarette or e-cigarette product. 2,288 (9%) of these households ever buy an e-cigarette. To test for patterns consistent with addiction, I model current purchases as a

[28]In their clinical practice guidelines, the U.S. Department of Health and Human Services (2008) reports that using nicotine gum and patches together leads to higher long-term abstinence rates relative to other treatments.

function of past purchase history. This framework for modeling addiction is consistent with existing models of addiction that allow past consumption to be complementary to current consumption. The model framework is:

$$c_{it} = \beta^T \boldsymbol{y}_{it-1} + \mu_i + \mu_t + \epsilon_{it} \tag{2}$$

$$e_{it} = \beta^T \boldsymbol{y}_{it-1} + \mu_i + \mu_t + \epsilon_{it} \tag{3}$$

where c_{it} and e_{it} are dummy variables indicating whether household i purchased at least one pack of cigarettes and at least one e-cigarette product in week t. $\boldsymbol{y}_{it-1}$ is a vector containing four dummy variables that indicate whether household i purchased the following products in week $t-1$: 1) at least one pack of cigarettes, 2) any kind of e-cigarette product, 3) nicotine gum, or 4) a nicotine patch.[29] Finally, the regressions include household fixed effects μ_i, such that the coefficients are identified off of within-household variation over time, and week fixed effects μ_t, which capture aggregate trends and seasonality in cigarette sales. Standard errors are clustered at the household level.

The first column of Table 5 presents the regression results when the binary decision to purchase tobacco cigarettes is the dependent variable. The coefficient on the indicator of a cigarette purchase in the previous week is positive and significant, indicating that households are more likely to buy in the current period if they purchased in the past. This is a pattern which is consistent with addiction and, more generally, with positive state dependence. The coefficients on the variables recording past purchase incidence of e-cigarettes and smoking cessation products are negative and significant, indicating that individuals are less likely to purchase a cigarette product when they have purchased one of these alternative nicotine products recently. If households were using e-cigarettes as a complement to traditional cigarettes, we might expect to see a positive relationship between recent purchases of e-cigarettes and current purchases of traditional cigarettes. Empirically, this is not the case.

The second column presents the regression results when the e-cigarette purchase indicator is the dependent variable. The results are similar, again showing patterns consistent with substitution and addiction or state dependence at the product level. The coefficient on

[29]In Appendix G, I report an alternative specification that includes controls for current prices of all four products and current advertising for e-cigarettes and smoking cessation products. The coefficients on the lagged purchase dummies are very stable in both sign and magnitude. Notably, the advertising effects in these household-level regressions are not statistically significant. The lack of significant advertising effects using the household data ties into the main motivation for using the aggregate data to pin down advertising effects. The aggregate data pools information across many consumers, which makes it easier to measure small advertising effects.

Table 5: Household Addiction and Substitution Patterns Between Cigarettes and E-Cigarettes

	Cig Purchase Incidence	E-Cig Purchase Incidence
Cig Purchase in Previous Week	0.100***	-0.001***
	(0.003)	(1.36e-4)
E-Cig Purchase in Previous Week	-0.038***	0.160***
	(0.007)	(0.015)
Nicotine Gum Purchase in Previous Week	-0.034***	0.001
	(0.009)	(0.002)
Nicotine Patch Purchase in Previous Week	-0.049***	-9.43e-5
	(0.009)	(0.001)
HH FE	Y	Y
Week FE	Y	Y
N Obs	4,609,029	4,609,029
N HHs	25,159	25,159
N E-Cig HHs	2,288	2,288
Mean DV	0.140	0.002
Mean DV if E-Cig Buyer	0.251	0.015
Last Week Cig as % of DV	71.0%	-55.8%
Last Week E-Cig as % of DV for E-Cig Buyers	-15.2%	1,049.1%

Clustered standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: Magnitude of change post e-cigarette reported as percent of average DV for those households who ever purchase an e-cigarette. E-cigarette users are on average heavier smokers than non e-cigarette users. The average weekly cigarette purchase incidence for e-cigarette users is 0.25 and for non e-cigarette users is 0.13.

e-cigarette incidence in the previous week is positive, consistent with addiction or state dependence, while the coefficient on cigarette purchase incidence in the previous week is negative. If e-cigarettes and tobacco cigarettes act as complements because they both reinforce the nicotine addiction stock, we would expect that past purchases of tobacco cigarettes would increase dependence on nicotine, which would lead to an increase in demand for e-cigarettes. The fact that the coefficient on past tobacco cigarette purchase incidence is negative suggests that households treat these products more as substitutes. The coefficients on the variables recording past purchase incidence of smoking cessation products are not statistically significant.

A key result of this analysis is that addiction or state dependence appears to operate at the product level, rather than the category level. Although we cannot interpret these results as causal, these substitution patterns are consistent with e-cigarettes acting as a substitute to traditional cigarettes, as opposed to as a gateway through nicotine addiction.

In the preceding sections, I presented reduced form evidence that e-cigarette advertising increases demand for e-cigarettes and reduces demand for cigarettes and smoking cessation products. Analysis of household panel data further showed that households tend to reduce their consumption of cigarettes after they purchase e-cigarettes and that addiction is an important force at play in this market. In the following section, I present a structural model of demand for cigarettes that is motivated by these empirical findings. The model will allow me to i) simultaneously account for advertising effects and addiction, ii) implement more efficient joint estimation using both aggregate and household data, iii) control for unobserved heterogeneity in preferences, and iv) evaluate a counterfactual scenario that predicts how cigarette demand would respond to a proposed ban of e-cigarette TV advertising.

5 An Integrated Micro-Macro Model of Demand

5.1 Overview

My descriptive analysis of market-level sales and advertising data indicates that e-cigarette advertising reduces demand for traditional cigarettes. These results suggest that banning e-cigarette advertising may have unintended consequences and actually lead to an increase in aggregate cigarette consumption. The magnitude of this effect is of great importance to policy makers as they consider whether to impose restrictions on advertising for e-cigarettes. In the following sections, I develop a structural model of demand for cigarettes and use the estimated preference parameters to predict the counterfactual demand for cigarettes that would have been observed in the absence of e-cigarette advertising.

I specify a structural model that i) harnesses the information content of both individual and aggregate data in an efficient and internally consistent way, ii) incorporates dynamic dependencies that arise as a result of nicotine addiction,[30] and iii) identifies advertising effects accounting for endogeneity using the border strategy approach. The existing literature has addressed each of these individually, but this paper is the first to unify these objectives within a single cohesive framework. I discuss each of these aspects of the model in turn below.

In theory, I could use either the aggregate or household-level data to estimate demand for cigarettes. However, each dataset has its relative merits and limitations. The aggregate data measures advertising effects with less noise and can be used to recover unobserved aggregate demand shocks, while the household data is more transparent in revealing patterns of addiction and heterogeneity in the population. For these reasons, I leverage both datasets to estimate demand for cigarettes. Specifically, I propose an individual-level demand model that aggregates in an internally consistent way, such that the equations that govern household and aggregate demand are functions of the same parameters. In order to estimate the model, I adapt an integrated estimation procedure developed by Chintagunta & Dubé (2005), who illustrate how to combine household and aggregate store level data to estimate the parameters of a discrete choice random coefficients model of demand. The intuition behind their estimation approach is to take advantage of the relative merits of each dataset to simultaneously estimate the mean effects of marketing activities, account for endogeneity in prices, and allow for heterogeneity across households. As Chintagunta and Dubé point out, although heterogeneity in the population can be identified using only aggregate data (Berry, Levinsohn, & Pakes (1995)), household panel data is more informative about heterogeneity than store level data.[31] Motivated by these facts, Chintagunta and Dubé propose a method to use aggregate data to estimate mean preference parameters and address the endogeneity problem and household-level data to estimate the distribution of heterogeneity.

I extend this micro-macro demand model to account for dynamic dependencies that arise as a result of nicotine addiction and other forms of structural state dependence.[32] State

[30] I do not model rational addiction in the sense that individuals in my model are not forward-looking (Becker & Murphy (1988), Gordon & Sun (2014)). This modeling assumption yields an individual level demand model that can be aggregated in an internally consistent way. Allowing for forward-looking behavior would make this aggregation intractable and would inhibit my ability to combine both individual and aggregate data in estimation.

[31] Subsequent work has shown that supplementing an aggregate model with household moments can generate more realistic model-predicted substitution patterns (Petrin (2002) and Berry, Levinsohn, & Pakes (2004)).

[32] Dubé, Hitsch, & Rossi (2010) define structural state dependence as occurring when "past purchases directly influence the consumer's choice probabilities for different brands." They distinguish structural state dependence

dependence is not incorporated in the Chintagunta and Dubé approach, but it is key to the analysis of addictive goods. The incorporation of state dependence, however, complicates the aggregate demand system considerably, since demand is no longer independent across time. In order to capture this persistence across time, I adapt a formulation from Caves (2004). Caves presents an aggregate structural model of demand for cigarettes that incorporates addiction as a form of category-level state dependence where a consumer's utility from buying cigarettes in the current period is higher if he purchased cigarettes in the previous period. He allows for heterogeneity in the form of discrete types. I combine Caves' model, which was developed originally for only aggregate data, with Chintagunta and Dubé's estimation strategy, while extending Caves' algorithm to allow for a continuous distribution of heterogeneous preferences. I find that allowing for a rich continuous distribution of heterogeneity is important to correctly separate the impact of addiction – a form of state dependence – from persistent unobserved tastes, an observation well known to econometricians (Heckman (1981)).[33]

The final modeling challenge I face is how to incorporate the identification of advertising effects within the structural model. The same intuition behind identification in the reduced form setting holds in the structural model as well. I estimate the model only using data for stores and individuals located within border markets, and I include market-border and border-time fixed effects. I explain in further detail how the structural model accommodates these fixed effects in Section 6.

In the sections below, I first lay out the equations characterizing individual-level demand and then show how the model aggregates and accommodates unobserved heterogeneity. Next, I describe the estimation procedure in more detail. Finally, I present the estimation results and use the model estimates to consider the impact of a proposed ban on e-cigarette advertising.

5.2 Individual Level Model

I specify an individual-level discrete choice model where consumers choose whether to buy a pack of cigarettes, an e-cigarette, a smoking cessation product, or not to make a purchase.[34] To

from spurious state dependence, which "arises if consumers differ along some serially correlated unobserved propensity to make purchase decisions." Economic mechanisms that can create structural state dependence include addiction, loyalty, search, and learning. I would also classify stockpiling as a form of negative structural state dependence.

[33]In Appendix H, I use model simulations to show that the model is well identified and that combining aggregate and household data leads to increased estimation efficiency.

[34]Initially, a consumer's choice sets includes only cigarettes, cessation products, and the outside option. I allow e-cigarettes to enter the choice set in each market in different periods. Specifically, in the store sales data for each market I find the first week of sustained positive e-cigarette sales, and I assume that e-cigarettes entered the

incorporate addiction, an important characteristic of the cigarette market, I allow utility from consuming in the current period to be increasing in consumption in the previous period.

Consider an individual i who shops in market m. Denote the individual's indirect utility function from consuming product $j \in \{c, e, q, 0\}$ by equations 4 – 7.[35] The indirect utility is a function of observed variables and unobserved product characteristics. Observed variables include current prices p_{jmt} and advertising for e-cigarettes and smoking cessation products, A_{emt} and A_{qmt} respectively. Note that e-cigarette advertising A_{emt} enters the indirect utility for e-cigarettes, as well as the utility for traditional cigarettes and smoking cessation products. This flexible model allows for the possibility that e-cigarette advertising decreases demand for cigarettes through a direct positive effect on demand for e-cigarettes (substitution / business stealing), while simultaneously allowing for the possibility that e-cigarette advertising makes cigarettes more attractive relative to the outside option (renormalization of smoking).[36] Similarly, the model allows for the fact that e-cigarette advertising could decrease demand for smoking cessation products through substitution or increase sales of smoking cessation products by reminding consumers about alternatives to cigarettes or their desire to quit smoking. Also observed is the purchase outcome in each period where $y_{it} \in \{c, e, q, 0\}$ indicates whether individual i in time period t purchased a pack of traditional cigarettes, an e-cigarette, a cessation product, or did not buy in the category. Dummies indicating purchase of an e-cigarette or tobacco cigarette product in the previous period capture addiction in the model.[37] The unobserved (to the econometrician) components of the indirect utility function include ξ_{jmt} which captures systematic shocks to aggregate demand including, for example, unobserved marketing activity,

choice set in that week. I group the nicotine patch and gum together and treat them as a single cessation product because sales of the nicotine patch are relatively infrequent and there are many weeks with 0 sales of patches in the aggregate data. Appendix I explains how I aggregate the cessation products and how I handle observations in the household data that are inconsistent with the discrete choice modeling assumption.

[35]This is a model of choice at the sub-category level, not the UPC-level. Henceforth, I use the term product to reference distinct sub-categories of products, i.e. a cigarette, e-cigarette, or smoking cessation product.

[36]Such a model specification can arise from a direct utility function in which the marginal utility of consuming traditional cigarettes is a function of e-cigarette advertising. Suppose a consumer maximizes the direct utility function $u(x_c, x_e, z) = \psi_c x_c + \psi_e x_e + \psi_z z$ s.t. $p_c x_c + p_e x_e + z = y$. The marginal utility of consuming x_c is then $\frac{\partial u}{\partial x_c} = \psi_c$. Typically we model ψ_j as a function of the attributes of product j. In this case, since I want to allow for the possibility that e-cigarette advertising directly affects the utility from consuming tobacco cigarettes, I allow ψ_c to be a function of e-cigarette advertising A_e.

[37]An alternative way of modeling nicotine addiction would be to have one addiction parameter γ that boosts utility for both e-cigarettes and tobacco cigarettes if the individual purchased either of these nicotine products in the previous period. Because the household-level analysis in Section 4.2 shows that individuals are less likely to buy tobacco cigarettes after buying e-cigarettes in the past, I chose to make γ product specific. I do not allow for a state-dependence parameter for smoking cessation products because there is very little repeat-purchase activity of smoking cessation products in the household data.

and ε_{ijt}, a stochastic error which is assumed to be distributed type I extreme value. Individuals have heterogeneous preferences for tobacco cigarettes and e-cigarettes, denoted by β_{ic} and β_{ie}.[38] The deterministic part of utility from consuming the outside good is normalized to 0.

$$u_{icmt} = \beta_{ic} + \alpha p_{cmt} + \phi_c A_{emt} + \gamma_c \mathbb{I}(y_{it-1} = c) + \xi_{cmt} + \varepsilon_{ict} \tag{4}$$

$$u_{iemt} = \beta_{ie} + \alpha p_{emt} + \phi_e A_{emt} + \gamma_e \mathbb{I}(y_{it-1} = e) + \xi_{emt} + \varepsilon_{iet} \tag{5}$$

$$u_{iqmt} = \beta_q + \alpha p_{qmt} + \phi_q A_{emt} + \psi A_{qmt} + \xi_{qmt} + \varepsilon_{iqt} \tag{6}$$

$$u_{i0mt} = \varepsilon_{i0t} \tag{7}$$

The observed price and advertising variables, together with a set of product intercepts, are grouped into a matrix X. Integrating out the distribution of stochastic errors ε_{ijt}, define the probability that an individual i in market m will buy j at time t if they bought k at time $t-1$ as $\pi_{imt}(y_{it} = j \mid y_{it-1} = k)$. For ease of exposition, I use the short-hand $\pi_{imt}(j \mid k)$:

$$\pi_{imt}(j \mid k) = \frac{e^{X_{jmt}\theta_i + \xi_{jmt} + \gamma_j \mathbb{I}(j=k)}}{1 + \sum_{l \in \{c,e,q\}} e^{X_{lmt}\theta_i + \xi_{lmt} + \gamma_l \mathbb{I}(l=k)}} \tag{8}$$

5.3 Aggregate Homogeneous Model

Conditional on past consumption status, the probability that an individual will buy product j is the logit probability given by equation 8. If consumers are homogeneous, $\pi_{imt}(j \mid k) = \pi_{mt}(j \mid k)$. Let s_{jmt} denote the market share of product j in market m in week t and s_{0mt} denote the market share of the outside good. Aggregate market shares can then be expressed as the weighted sum of purchase probabilities conditional on consumption status, where the weights are the probability of having that consumption status. Furthermore, in the model with homogeneous consumers, the probability of being in a given consumption state is just equal to the market share of that good in the previous period.

$$s_{jmt} = \sum_{k \in \{c,e,q,0\}} \pi_{mt}(j \mid k) s_{kmt-1} \tag{9}$$

[38]In theory, consumers could have heterogeneous preferences over smoking cessation products as well. In practice, I observe very few households buying smoking cessation products and little variation in purchase frequency conditional on purchase, so there is not a lot of variation with which to pin down heterogeneity in preferences.

5.4 Incorporating Unobserved Heterogeneity

Thus far, I have shown how to derive aggregate demand from a homogenous demand model with state dependence. In this section I extend the derivation to a model with unobserved heterogeneity in preferences.

The key insight in deriving aggregate demand in a model with heterogeneity is that the joint distribution of heterogeneity and state dependence is not stationary; rather, it evolves over time. For example, if consumers vary in their preference for cigarettes, then an increase in price will differentially decrease the probability that consumers of all types buy in the current period. This will affect the joint distribution of consumer types and consumption states in the next period. In particular, prices and advertising in the current period affect the joint distribution of state dependence and heterogeneity in all subsequent periods.

In order to obtain aggregate market shares, I need to integrate out unobserved heterogeneity and the stochastic demand shocks. In the model with heterogeneity, I calculate aggregate shares by integrating the purchase probabilities conditional on consumption status and consumer type against the joint distribution of consumption status and heterogeneity.

$$s_{jmt} = \int_{\Theta \times \{c,e,q,0\}} \pi_{imt}(j \mid k) dF_{mt}(\theta, k) \tag{10}$$

The discussion above does not assume any particular joint distribution of unobserved heterogeneity and state dependence. In the estimation section below, I make specific assumptions about that distribution and show how to numerically evaluate the above integral.

Discussion

Before moving on to the estimation procedure, I first discuss some of my modeling assumptions. First is the decision to use a discrete choice model instead of explicitly modeling purchase quantities. Past work on addiction has assumed that addiction operates through the effect of past purchase quantities on current purchase quantity (Becker & Murphy (1988), Gordon & Sun (2014)). The household panel data would in theory allow me to model quantities; however, the panel is thin. The aggregate data is richer and allows me to identify advertising effects with more precision, but it limits my ability to model purchase quantities.[39] In order to be able to harness the richness of the aggregate data, I choose to model purchase incidence in a discrete

[39]Hendel & Nevo (2013) model purchase quantities using aggregate data, but need to impose other simplifying assumptions in order to make their model tractable with aggregate data.

choice framework.[40]

A separate but related assumption is that only the previous week's purchase decision affects current period consumption and that consumers are not forward looking. An assumption closer to observed consumer behavior and patterns of addiction might allow additional lags of purchase decisions to affect current choices. I choose to work with the simpler one period lag because the model with state dependence can be estimated using aggregate data.

6 Estimation and Results

6.1 Estimation with Unobserved Heterogeneity

The model discussion above did not rely on any specific assumptions about the distribution of unobserved heterogeneity. In my model implementation, I allow the cigarette and e-cigarette intercepts to vary across the population, as described in the individual level model in Section 5.2. In the final specification, I assume that unobserved heterogeneity follows a normal distribution, but to facilitate exposition, I first introduce the model with R discrete types. Within each market, the marginal probability of being a certain type r is $Pr(\theta_r)$. For each type r, the probability of purchasing product j after purchasing k in the prior period is the familiar logit probability $\pi_{rmt}(j \mid k)$.

In each market and period, the population of consumers is distributed across these types and consumption states according to the joint distribution $Pr(\theta_r, y_{rmt-1} = k)$ where $k \in \{c, e, q, 0\}$. Note that the distribution in the current period is determined by the consumption choices made in the previous period. The distribution in the initial period is $Pr(\theta_r, y_{rmt_0} = k)$.[41] In subsequent periods, the marginal probability of being a certain type $Pr(\theta_r)$ remains constant, but the joint distribution of consumer types and consumption status evolves as the heterogeneous population responds to variation in prices and advertising.[42] The joint distribution updates each period according to the recursion in equation 11.

$$Pr(\theta_r, y_{rmt} = j) = \sum_{k \in \{c,e,q,0\}} \pi_{rmt}(j \mid k) \times Pr(\theta_r, y_{rmt-1} = k) \tag{11}$$

[40] Appendix I discusses how I rationalize the data with this modeling assumption.

[41] I discuss how I resolve this initial conditions problem with a burn-in period in Section 6.2.

[42] Note that across markets the marginal distribution $Pr(\theta_r)$ is the same, but the joint distribution can vary as consumers in different markets are exposed to different prices and advertising.

The recursion shows that the probability of being a specific type r and smoking in the current period, $y_{rmt} = c$, is equal to (a) the probability that a smoker of type r continues smoking in the current period, plus (b) the probability that an e-cigarette user of type r takes up cigarettes in the current period, plus (c) the probability that a smoking cessation user of type r takes up cigarettes in the current period, plus (d) the probability that a non-smoker of type r begins smoking in the current period.

Finally, aggregate market shares are obtained in the model with R latent types by weighting the logit probability of purchase for each type by the joint distribution of types and consumption states in the population. Specifically, the integral describing market shares in equation 10 becomes a summation over discrete types and consumption states, as shown in equation 12.

$$s_{jmt} = \sum_{r=1}^{R} \sum_{k \in \{c,e,q,0\}} \pi_{rmt}(j \mid k) \times Pr(\theta_r, y_{rmt-1} = k) \tag{12}$$

Now, I discuss how to extend the model to allow for a continuous heterogeneity distribution. I assume that the distribution of random coefficients follows a normal distribution, and I estimate the mean $\bar{\theta}$ and variance Σ of the distribution. Let ν_i be standard normal and Λ be the Cholesky decomposition of Σ s.t. $\theta_i = \bar{\theta} + \Lambda \nu_i \sim N(\bar{\theta}, \Sigma)$. The consumer's indirect utility function can be decomposed into common and individual-specific components, as shown in equation 13, where $\delta_{jmt} = X_{jmt}\bar{\theta} + \xi_{jmt}$ captures the mean aggregate utility level, including the average effects of prices and advertising, and $\mu_{ijmt}(X_{jmt}, y_{it-1}; \Sigma, \gamma) = X_{jmt}\Lambda \nu_i + \gamma_j \mathbb{I}(y_{it-1} = j)$ represents deviations from the mean utility level. Specifically, μ captures heterogeneity and the increase in utility generated by addiction and other forms of structural state dependence.

$$u_{ijmt} = \delta_{jmt}(X_{jmt}, \xi_{jmt}; \bar{\theta}) + \mu_{ijmt}(X_{jmt}, y_{it-1}; \Sigma, \gamma) + \varepsilon_{ijt} \tag{13}$$

The additional layer of complication in incorporating a continuous distribution of unobserved heterogeneity is in how to evaluate the integral in equation 10 and how to update the joint distribution of unobserved heterogeneity and state dependence each period. Like in a standard random coefficients model, I integrate out unobserved heterogeneity by taking draws from the latent distribution and using Monte Carlo integration. Conditional on R draws from the latent normal, the Monte Carlo integral is equivalent to an R-type latent class model. Equation 11 approximates the joint distribution of heterogeneity and state dependence and equation 12 can be used to obtain the model-predicted aggregate market shares.

6.2 Estimation Procedure

At a high level, I estimate the mean utility parameters $\bar{\theta}$ and recover unobserved demand shocks ξ_{jmt} from aggregate data and estimate the heterogeneity distribution Σ and addiction parameters γ_c and γ_e from household panel data. The estimation steps are described below.

6.2.1 Step 1: Aggregate Data Step

Given a guess of the heterogeneity and addiction parameters $(\tilde{\Sigma}, \tilde{\gamma}_c, \tilde{\gamma}_e)$, for each market m, product j, and time period t, I compute $\delta_{jmt} = X_{jmt}\bar{\theta} + \xi_{jmt}$ that equates the model predicted market share to the observed market share in the aggregate data.[43] The model-predicted market share $s(X_{jmt}, \delta_{jmt}; \Sigma, \gamma_c, \gamma_e)$ is given by equation 10. In practice, I approximate the integral over the joint distribution of consumer heterogeneity and state dependence using Monte Carlo integration. I take $R = 200$ standard normal draws ν_r and for the given guess of $\tilde{\Sigma}$ calculate $\theta_r = \bar{\theta} + \tilde{\Lambda}\nu_r \sim N(\bar{\theta}, \tilde{\Sigma})$. Then I use equations 11 and 12, plugging in the guess of non-linear parameters $(\tilde{\Sigma}, \tilde{\gamma}_c, \tilde{\gamma}_e)$, to calculate the model-predicted aggregate market shares. The recursion in equation 11 relies on knowing the joint distribution of heterogeneity and consumption status in the initial period. I use the first quarter of data for each market to forward simulate the joint distribution.[44] I then use the remaining weeks of data in estimation.

With the equations describing model-predicted shares in hand, I calculate the values of δ_{jmt} that equate observed and model-predicted shares using the BLP contraction mapping algorithm shown in equation 14 (Berry et al. (1995)). The values of δ_{jmt} must be calculated iteratively each period because state dependence causes the current period share to depend on the previous period's prices, advertising, and unobserved demand shocks ξ_{jmt-1}. The resulting

[43] I calculate observed market shares by dividing total store sales in each market by the adult smoking population of that market. Because I do not observe the fraction of total sales that are covered by Nielsen stores in different markets, I re-scale county-level adult population measures by state level smoking prevalence and intensity rates. I then adjust this measure to make the observed shares in the data consistent with the purchase probabilities observed in the household data.

[44] The literature has typically resolved this type of initial conditions problem by either treating the initial probability distribution as parameters of the model to estimate, or by using an initial period of data as a burn-in period to forward simulate the distribution (Erdem, Imai, & Keane (2003)). I take the second approach. For each guess of the parameters, I re-calculate the series of probabilities governing the evolving joint distribution of heterogeneity and state dependence for the initial burn-in period. I assume equal probabilities of smoking and not smoking for each type in the first week of the burn-in period, such that the probability of having a given type and smoking consumption status at the beginning of the burn-in period is equal to $\frac{1}{2R}$. I have tried a variety of different starting values and found that the joint distribution converges to the same steady state within the burn-in period.

vector becomes the current guess of δ, which is denoted as $\tilde{\tilde{\delta}}$ in Step 2 below.

$$\delta^{h+1}_{jmt} = \delta^{h}_{jmt} + \ln S_{jmt} - \ln s(X_{jmt}, \delta^{h}_{jmt}; \tilde{\Sigma}, \tilde{\gamma}_c, \tilde{\gamma}_c) \tag{14}$$

6.2.2 Step 2: Household Data Step

Given the current guess of $\tilde{\tilde{\delta}}$, I estimate Σ, γ_c and γ_e via maximum likelihood with household data. Each household is matched to its aggregate data counterpart.[45] Substituting the appropriate $\tilde{\tilde{\delta}}_{jmt}$ into the household's indirect utility function, the probability that a household buys a given product in a given period is given by equation 15. Integrating out the distribution of unobserved heterogeneity, the likelihood for each individual is then given by equation 16, where Y_i is a vector indicating the choices made by consumer i over time.

$$P_{ijmt}(X_{mt}, \tilde{\tilde{\delta}}_{mt}, y_{it-1}, \Sigma, \gamma_c, \gamma_e) = \frac{\exp[\tilde{\tilde{\delta}}_{jmt} + X_{jmt}\Lambda\nu_i + \gamma_j \mathbb{I}(y_{it-1} = j)]}{1 + \sum\limits_{k\in\{c,e,q\}} \exp[\tilde{\tilde{\delta}}_{kmt} + X_{kmt}\Lambda\nu_i + \gamma_k \mathbb{I}(y_{it-1} = k)]} \tag{15}$$

$$L_i(Y_i|X_m, \tilde{\tilde{\delta}}_m; \Sigma, \gamma_c, \gamma_e) = \int \prod_{t=1}^{T_i}\prod_{j=1}^{J} P_{ijmt}(X_{mt}, \tilde{\tilde{\delta}}_{mt}, y_{it-1}, \Sigma, \gamma_c, \gamma_e)^{\mathbb{I}(y_{it}=j)} dF_{\nu} \tag{16}$$

In practice, I approximate the integral using a Monte Carlo simulation using the same $R = 200$ draws from the standard normal that I used in Step 1, and I estimate the parameters Σ, γ_c, and γ_e by maximizing the log likelihood in equation 17 via simulated maximum likelihood.

$$\mathscr{L}(Y|X, \tilde{\tilde{\delta}}; \Sigma, \gamma_c, \gamma_e) = \sum_{i=1}^{N} \log[L_i(Y_i|X_m, \tilde{\tilde{\delta}}_m; \Sigma, \gamma_c, \gamma_e)] \tag{17}$$

6.2.3 Iterate Until Convergence

I iterate steps 1 and 2 until the estimated parameters and mean utility terms $(\Sigma, \gamma_c, \gamma_e, \delta)$ differ by less than 10^{-6}.

[45]The matched sample contains 6,861 households who ever make a purchase of a cigarette or e-cigarette and who reside within a border county. This is out of the 25,077 households who ever purchase a cigarette product.

6.2.4 Estimate Linear Parameters from Aggregate Data

After the model parameters have converged, I then use the fact that $\delta_{jmt} = X_{jmt}\bar{\theta} + \xi_{jmt}$ to estimate the linear parameters $\bar{\theta}$. Specifically, I estimate $\hat{\bar{\theta}} = (X'X)^{-1}X'\hat{\delta}$.

6.2.5 Inference

I calculate standard errors for Σ, γ_c and γ_e, the model parameters identified off of the household data, by inverting the Hessian at the optimum of the likelihood function. Standard errors for the remaining linear parameters are calculated using a bootstrap procedure that takes into account the fact that the dependent variable δ was estimated in a first stage. Specifically, I take $N = 1,000$ draws from the asymptotic distribution of the non-linear parameters $\Omega = (\Sigma, \gamma_c, \gamma_e)$, and for each draw ω_n I calculate the implied vector $\delta(\omega_n)$ that equates observed and model-predicted shares. For each iteration of the bootstrap, I draw B borders with replacement from the data for all borders $\{(\delta_1(\omega_n), X_1), ..., (\delta_B(\omega_n), X_B)\}$ and stack the resampled blocks to create a bootstrapped dataset (δ^*_n, X^*_n). I then estimate $\hat{\bar{\theta}}_n = (X^{*'}_n X^*_n)^{-1} X^{*'}_n \delta^*_n$. The standard deviation of the distribution of the bootstrapped $\hat{\theta}_n$ estimates gives standard errors for the linear parameters. Intuitively, the first component of this bootstrap procedure captures estimation error from the non-linear first stage and the block bootstrap component captures typical sampling error.

6.3 Identification

Before presenting the model estimates, I first discuss identification and highlight how I incorporate the border counties identification strategy into the estimation of the structural model. I estimate the model using aggregated store data for *only* those stores in border county markets and household data for *only* those households who reside within border counties. Thus, the same regression discontinuity identification from the linear model applies here — the nonlinear estimator is also only based on the behavior of marginal consumers at borders. In total I have data for 232 markets and 6,861 households. The fact that the linear parameters are estimated in a simple linear regression allows me to continue to include a rich set of border-market and border-time fixed effects like in the descriptive regressions. Specifically, I include a set of almost 14,000 product-border-market and product-border-month fixed effects in the structural estimation. It would be impossible to include this many parameters in a typical non-linear optimization routine. The linear regression stage is thus an important component of the model that allows me to incorporate regression discontinuity identification into the structural model.

Finally, the household purchase data identifies the parameters pinning down the heterogeneity distribution and state dependence, while the aggregate data identifies the mean utility parameters, including the price and advertising coefficients. Combining the household and aggregate data in this way requires the maintained assumption that the households living in border regions are a representative sample from the population of consumers that shop at stores in the border regions.

6.4 Estimation Results

Table 6 presents the estimated model parameters. The first column reports estimates from a homogenous aggregate logit model without addiction. The second column reports estimates for the homogeneous joint model with addiction. The third column reports estimates for the heterogeneous joint model without addiction. The fourth column reports estimates for the heterogeneous joint model with addiction. The heterogeneity in models 3 and 4 consists of a random coefficient on the traditional cigarette intercept and two-type discrete heterogeneity on the e-cigarette intercept. The two-type heterogeneity for e-cigarettes is motivated by the observation that the majority of households never buy an e-cigarette, while a small segment of households buy frequently. Consumers are modeled as either having high or low preference for e-cigarettes, $\beta_e \in \{\beta_{eL}, \beta_{eH}\}$, where the probability of being a high-type is allowed to be correlated with the consumer's taste for traditional cigarettes. In particular, I assume $Pr(\beta_e = \beta_{eH}) = \frac{exp[\lambda_H + \rho_{ce} \nu_r]}{1+exp[\lambda_H + \rho_{ce} \nu_r]}$ where ν_r is a draw from the heterogeneity distribution for preference for tobacco cigarettes and λ_H and ρ_{ce} are parameters to be estimated. For each market, β_{eL} is treated as the mean utility level, and in the household maximum likelihood step, I search over $\Delta\beta_e$ to pin down the preference of the high type where $\beta_{eH} = \beta_{eL} + \Delta\beta_e$.

Focusing on the estimates for the model with heterogeneity and addiction (column 4), I first discuss the linear parameters that are identified off of the aggregate data. The coefficient on price is estimated to be negative and statistically significant. The e-cigarette own-ad coefficient ϕ_e is positive but not statistically significant. The cross-ad effect ϕ_c which would allow for positive spillovers from e-cigarette advertising to demand for tobacco cigarettes is negative and significant. This negative coefficient reflects an additional reduction in demand for tobacco cigarettes, on top of the reduction in traditional cigarette market share implied by the positive effect of e-cigarette advertising on e-cigarette demand in the logit model. Similarly, the cross-ad effect ϕ_q which allows for spillovers from e-cigarette advertising to demand for smoking cessation products is negative and marginally significant. The negative sign implies

Table 6: Model Estimation Results

		(1) Aggregate Logit	(2) Joint Homogeneous w/ Addiction	(3) Joint Heterogeneous w/out Addiction	(4) Joint Heterogeneous w/ Addiction
Price	α	-0.2523***	-0.2448***	-0.2666***	-0.2590***
		(0.0017)	(0.0090)	(0.0104)	(0.0095)
E-Cig Ads on E-Cigs	ϕ_e	0.0015	0.0016	0.0017	0.0019
		(0.0019)	(0.0014)	(0.0016)	(0.0015)
E-Cig Ads on Cigs	ϕ_c	-0.0038**	-0.0034***	-0.0064***	-0.0061***
		(0.0019)	(0.0005)	(0.0010)	(0.0010)
E-Cig Ads on Cess	ϕ_q	-0.0039**	-0.0039*	-0.0038*	-0.0038*
		(0.0019)	(0.0023)	(0.0023)	(0.0023)
Cessation Ads	ψ	-0.0044***	-0.0044***	-0.0043***	-0.0043***
		(0.0007)	(0.0007)	(0.0007)	(0.0007)
Cig State Dependence	γ_c	-	3.5730***	-	0.5472***
		-	(0.0049)	-	(0.0078)
E-Cig State Dependence	γ_e	-	5.5494***	-	2.4458***
		-	(0.0574)	-	(0.0852)
Cig Intercept Std. Dev.	σ_{β_c}	-	-	2.5639***	2.3102***
		-	-	(0.0141)	(0.0126)
E-Cig Intercept Hetero	$\Delta\beta_e$	-	-	4.9151***	3.9109***
		-	-	(0.0535)	(0.0549)
Prob High E-Cig Type	λ_H	-	-	-5.0592***	-4.6969***
		-	-	(0.1781)	(0.1680)
Corr Pref Cig E-Cig	ρ_{ce}	-	-	-0.5690**	-0.6314**
		-	-	(0.2450)	(0.2094)
Product-DMA-Border FEs		Y	Y	Y	Y
Product-Border-Month FEs		Y	Y	Y	Y
Aggregate Data		✓	✓	✓	✓
Individual Data		-	✓	✓	✓
Median Cig Price Elasticity		-1.15	-0.90	-0.66	-0.64
Median Cig Ad Elasticity		-0.004	-0.003	-0.004	-0.003
N Markets		232	232	232	232
N Aggregate Obs		66,584	66,584	66,584	66,584
N Households		-	6,861	6,861	6,861
N Household Obs		-	1,125,410	1,125,410	1,125,410

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

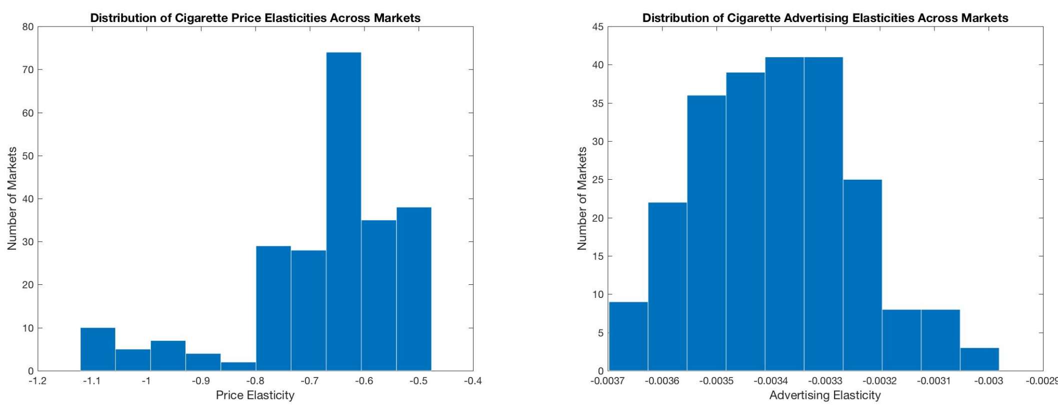
Note: Models with heterogeneity estimated with R=200 simulation draws. Standard errors for the non-linear parameters governing state dependence and heterogeneity are obtained by inverting the Hessian at the optimum of the likelihood function. Standard errors for the linear parameters are calculated using a bootstrap procedure that accounts for estimation of the mean utilities, δ.

an additional reduction in demand for smoking cessation products, on top of the reduction implied by the positive effect of e-cigarette advertising on e-cigarette demand in the logit model. Putting it all together, the estimated ϕ_j coefficients, combined with the model's functional form, indicate that e-cigarette advertising leads to an increase in demand for e-cigarettes and a decrease in demand for tobacco cigarettes and smoking cessation products. These predictions are consistent with the descriptive results shown in Table 4. Note that the model's functional form does not by itself impose that shares of any of the products will be increasing or decreasing in e-cigarette advertising. The ϕ_j coefficients give the model the flexibility to capture increases or decreases in the shares of all of the inside goods. The fact that not all of the ϕ_j coefficients are significantly different from zero suggests that, combined with the model's functional form, not all of the e-cigarette ad coefficients are needed to explain the observed share patterns in the data. The fact that the cross-ad effect ϕ_c is significant and the own-ad effect ϕ_e is not is likely due to the fact that e-cigarette shares are very small relative to tobacco cigarette shares, so there is more variation in tobacco cigarette shares for the model to leverage. In order to evaluate the impact of e-cigarette advertising on demand in this "saturated" model, it is necessary to consider the combined effect of all e-cigarette ad parameters on market shares of the inside goods. Finally, the negative estimate for the last advertising coefficient ψ predicts that sales of cessation products decrease when there is more advertising for cessation products. This result is surprising and differs from the insignificant own-effect of cessation advertising that was shown in the descriptive regressions in Table 4.

Turning to the non-linear parameters, the estimated standard deviation of the cigarette intercept random coefficient σ_{β_c} is large, reflecting the substantial heterogeneity in purchase probabilities observed in the household data. The two-type heterogeneity distribution on the e-cigarette intercept predicts a very high probability of being a low type and a small probability of being a high type. For a consumer with average preferences over cigarettes (i.e. $\nu_r = 0$), $Pr(\beta_e = \beta_{eL}) = 0.99$ and $Pr(\beta_e = \beta_{eH}) = 0.01$. The high types have a much higher preference for e-cigarettes than the low types ($\Delta\beta_e = \beta_{eH} - \beta_{eL}$) and are thus more likely to buy. These results are consistent with the purchase patterns in the household data; amongst those who ever buy in the category, a small share ever buy an e-cigarette, and the majority of those households that do buy e-cigarettes only buy once. In the full model, the negative estimate for ρ_{ce} indicates that the probability of being a high-type for e-cigarettes is negatively correlated with a consumer's preferences over tobacco cigarettes. Once heterogeneity is included in the model, the magnitude of the addiction parameters γ_c and γ_e decrease. In the model without heterogeneity, any serial correlation generated by unobserved heterogeneity is absorbed into γ

(Dubé et al. (2010)). Notably, γ_e is roughly 4 times larger than the analogous state-dependence parameter for cigarettes, γ_c. I hypothesize that this is the case because e-cigarettes are a new product, and therefore consumers learn about their preference for e-cigarettes through their prior purchases. In addition, the indicator for having bought e-cigarettes last week may proxy for consumer awareness of e-cigarettes as a product category. Thus, γ_e captures the total effect of different forms of structural state dependence including addiction, learning, loyalty, and stockpiling, and it may also incorporate a component of spurious state dependence, which is individual consumers' awareness of e-cigarettes.[46] It is important to account for this serial dependence in choices in order to get a clean read on the primary object of interest – the effect of e-cigarette advertising on cigarette demand.

Figure 10: Distribution of Average Cigarette Price and Ad Elasticities Across Markets



In order to build intuition around the estimation results, I calculate the implied short-run price and advertising elasticities for each market. The short run elasticity captures the responsiveness of demand to a one-time increase in price or advertising in the same week. Figure 10 shows the distribution across markets of the average short-run own-price and cross-ad elasticities of a pack of cigarettes. The median price elasticity across markets is -0.64, which is in line with previous estimates in the literature (Chaloupka (1991)). The median cross-ad elasticity across markets is -0.003. Comparing the implied elasticities across the four columns in Table 6, the price elasticity of cigarette demand becomes more inelastic as heterogeneity and addiction are incorporated into the model. This is consistent with the fact that consumers who are addicted or have a strong preference for cigarettes will be relatively insensitive to changes

[46] Stockpiling is a consumer behavior that could create negative serial dependence in choices. To the extent that consumers engage in stockpiling, the parameters γ_c and γ_e will reflect the net effect of these positive and negative forms of structural state dependence.

in price. The elasticity of cigarette demand with respect to e-cigarette advertising decreases slightly when state dependence is incorporated into the model. Comparing the model-predicted elasticities to the ad and price elasticities reported in Section 4.1.4, the full model that accounts for addiction and heterogeneity yields smaller price and advertising elasticities.

7 Counterfactual E-Cigarette Ad Ban

In April 2015, the American College of Physicians (ACP) published an opinion paper on e-cigarettes in the *Annals of Internal Medicine* that, among other regulatory requests, called for a prohibition on e-cigarette television advertising (Crowley (2015)). The ACP cited concerns that youth exposure to e-cigarette advertisements has increased dramatically in recent years and that e-cigarette advertising may help contribute to a re-normalization of smoking that will "reverse the progress made to stigmatize smoking and reduce its appeal among young people." The ACP further expressed concern that e-cigarette advertising will give smokers a false perception that e-cigarettes are a clinically accepted means of smoking cessation, and may therefore divert smokers who are looking to quit away from clinically proven cessation products. To date, there exists little to no empirical evidence that supports these arguments.

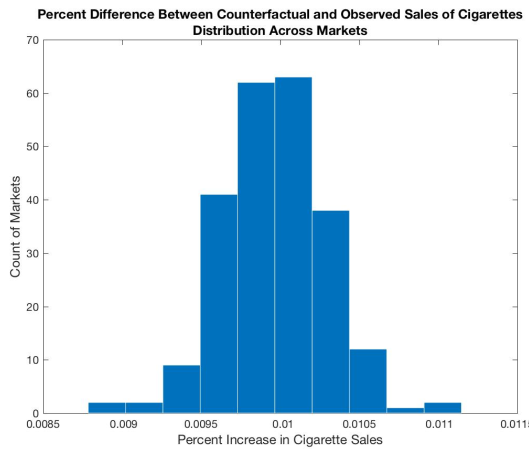
The previous sections provided empirical evidence that e-cigarette advertising has led to a reduction in sales of both tobacco cigarettes and smoking cessation products. In this section, I use the demand model estimates from Section 6 to predict the effect that an e-cigarette TV advertising ban would have on demand for cigarette and cessation products. Specifically, I impose a counterfactual ban on e-cigarette advertising beginning in 2012 and use the model estimates to forecast weekly demand over the next four years. Using the estimated parameters $\hat{\theta}, \hat{\Sigma}, \hat{\gamma}$ and demand shocks $\hat{\xi}_{jmt}$ and setting weekly e-cigarette advertising to 0, I calculate the counterfactual market shares for cigarettes, e-cigarettes, and smoking cessation products.[47] Multiplying by the market size gives me the predicted change in quantity of each product under the ad ban.

Banning e-cigarette advertising leads to an increase in the market share of tobacco cigarettes because the effect of e-cigarette advertising on cigarette demand $\hat{\phi}_c$ is negative and the effect of e-cigarette ads on e-cigarette demand $\hat{\phi}_e$ is positive. The overall percent increase in sales over the four year counterfactual is 1.0%. Though small, this is an economically significant increase given that between 2011 and 2012, the total number of cigarettes sold by the 5 major

[47] Recall that the advertising variable includes both local and national advertising. Both are set to 0 in the counterfactual analysis.

U.S. manufacturers fell by 2.2% (FTC (2015)). Although the advertising effects are estimated on a specific sub-set of markets and thus generalizations of the effects should be made with caution, if I assume that these estimates apply for the U.S. as a whole, the model predicts that approximately 130 million more packs of cigarettes would have been sold in the U.S. each year if there had been no e-cigarette advertising from 2012-2015.[48]

Figure 11: Cross-Sectional Variation in Response to Counterfactual Ad Ban



The counterfactual analysis also shows significant variation across markets in the predicted response to an e-cigarette ad ban because markets differ in their baseline preference for cigarettes, as well as in the intensity of e-cigarette advertising they were exposed to. Figure 11 shows the distribution across markets in terms of the overall response to the ban for the four year period from 2012-2015. The median percent increase in tobacco cigarette sales as a result of the ban is 1.00%. The minimum percent increase is 0.88% in Portland / Auburn, ME DMA border counties,[49] and the maximum percent increase is 1.11% in Miami / Ft. Lauderdale border counties.[50]

Turning to the effect of the ban on e-cigarettes, the counterfactual analysis predicts that sales of e-cigarettes would decrease by 0.9%. This decrease is comparable in magnitude to the increase in sales of tobacco cigarettes. However, because the baseline volume sales of tobacco cigarettes are much larger than baseline sales of e-cigarettes, the overall effect of the ban is a

[48]This back of the envelope calculation assumes baseline sales of 13.385 billion packs of cigarettes a year, which is the number of packs of cigarettes sold by the top 5 manufacturers in the U.S. in 2012 (FTC (2015)).

[49]Specifically, counties in the Portland/Auburn Maine DMA that share a border with counties in the Burlington NY/Plattsburgh VT DMA.

[50]Specifically, counties in the Miami Ft Lauderdale, FL DMA that share a border with counties in the Ft. Myers / Naples, FL DMA.

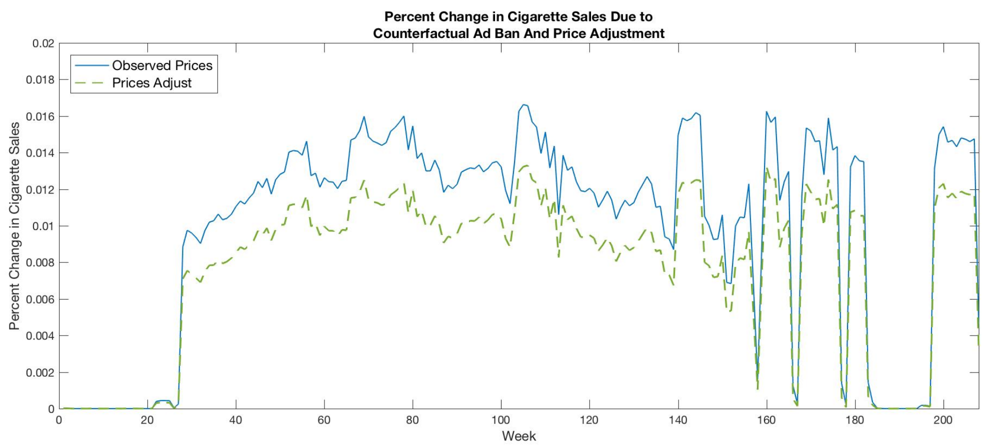
0.96% increase in the combined sales of tobacco and e-cigarettes.

Finally, the counterfactual suggests that e-cigarette advertising has led to a small decrease in sales of traditional smoking cessation products. The advertising ban is predicted to increase sales of cessation products by 1.0%. This result supports policymakers' concerns that e-cigarette advertising may divert smokers away from clinically accepted smoking cessation products (Crowley (2015)), though the effect size is relatively small.

Thus far, I have assumed that prices would remain at the observed levels when advertising is banned. However, it is possible that prices would adjust in the absence of advertising, and these price changes could off-set the effect of the ad ban on demand (Dubois, Griffith, & O'Connell (2018)). A full supply-side model would be very complex in this setting because of the presence of multi-product firms who are potentially forward-looking in their pricing decisions. Given that the focus of this paper is the effect of e-cigarette advertising on demand for tobacco cigarettes, such a complex supply-side pricing model is beyond the scope of this paper. However, to build intuition as to how prices might adjust under an ad ban, I specify a simple static supply-side pricing model and use it to predict how prices might change at the category-level. Appendix J details how counterfactual prices are estimated and discusses the credibility of these predicted price changes in more detail. The analysis predicts that prices would change very little. Directionally, cigarette prices are predicted to increase in a counterfactual world with no e-cigarette advertising, but across all markets and weeks, the mean predicted increase in cigarette prices is 1.6 cents per pack and the maximum predicted increase is 4.0 cents per pack. These changes are small compared to the average observed price per pack of $5.50 during the counterfactual period. Prices of e-cigarettes are predicted to decrease, but again these decreases are very small in magnitude. The mean (maximum) decrease in e-cigarette prices across all markets and weeks is a decrease of 0.3 (1.0) cents. The predicted change in the price of smoking cessation products is even smaller, with a maximum increase in price of less than 0.5 cents.

Next, I predict new market shares with these counterfactual prices and with e-cigarette advertising set to 0. Figure 12 graphs the predicted change in the total sales of tobacco cigarettes in the counterfactual with price adjustments and compares it to the setting where prices are held fixed. The predicted price adjustments partially off-set the increase in cigarette sales that results from the removal of e-cigarette advertising. Overall, allowing for price adjustments, tobacco cigarette sales over the four year counterfactual are predicted to be 0.8% higher than the observed sales in the data. Relative to the simulation when prices are held fixed, allowing prices to adjust reduces the increase in cigarette sales by 0.2 percentage points, which implies that price adjustments may off-set approximately 20% of the direct effect of banning e-cigarette ads. While this supply-side analysis is a simplified version of a very complex industry, this robustness helps guide intuition as to the direction in which prices might adjust and the magnitude of potential price changes. Altogether, this counterfactual analysis indicates that banning e-cigarette advertising would lead to a small but economically significant increase in cigarette sales.

Figure 12: Counterfactual Percent Increase in Cigarette Sales under Different Price Responses



8 Conclusions and Future Work

This paper is among the first to empirically analyze the effects of e-cigarette advertising on demand for traditional cigarettes, e-cigarettes, and smoking cessation products. Using both descriptive and structural methods, I show that e-cigarette advertising *decreases* demand for cigarettes and smoking cessation products. My research contributes to the ongoing policy debate as to whether e-cigarette TV advertising should be banned and suggests that a ban on e-cigarette advertising may have unintended consequences. More generally, my approach contributes to the study of advertising in categories with state dependence and to the analysis of substitution and complementarities in demand across categories.

Although this paper takes an important first step towards better understanding the role of e-cigarette advertising in the market, my analysis thus far is limited by the availability of data that would allow me to study additional questions that are of considerable interest to academics and policy makers. For example, I am not able to address the impact of e-cigarette advertising on teenagers' long-run demand for cigarettes and other nicotine products. This is an important area for future research that requires both data on youth consumption, which is not well covered in my dataset, as well as a long panel to track long-run consumption patterns. In addition, many of the pro-regulation arguments made by researchers, clinicians, and regulators are based on concerns about the long-term consequences of e-cigarette consumption. As time goes by and as individual states begin to pass new legislation concerning e-cigarette use indoors, tax policies, minimum purchase ages, and restrictions on e-cigarette advertising, new opportunities to study this growing market will likely arise. This paper has focused on category-level substitution between tobacco cigarettes and e-cigarettes. Future work could explore competition and substitution at the brand level. In addition, the advertising effects in this paper are identified off of variation in sales and advertising in counties that lie on the borders between DMAs. Future research should explore the generalizability of these results to other markets, especially urban areas where more young people may reside.

Future work could also address the supply side of the market. In the absence of regulatory intervention, the future of e-cigarettes will be largely shaped by industry manufacturers and vendors. Initially the industry was composed of many small, independent producers who had no interest in perpetuating tobacco consumption. However, with the entry of the Big Tobacco companies into the arena in recent years, the incentives for producers have changed. The industry has been growing more concentrated with the largest emerging players being the big cigarette manufacturers. Rather than encourage users to quit smoking, cigarette companies are incentivized to maintain smoking as the status quo[51] and invest in e-cigarettes as a long-term hedge in the event that the market for tobacco cigarettes dissolves in the future. With the rapid growth of e-cigarette sales in the market, Solomon (2014) even argues that the 2015 merger between Reynolds and Lorillard was partially motivated by fear of the rapidly growing e-cigarette market and the disruption this new technology will cause going forward.

References

Baltagi, B. H., & Levin, D. (1986). Estimating dynamic demand for cigarettes using panel data: The effects of bootlegging, taxation, and advertising reconsidered. *Review of Economics and*

[51] As my analysis finds that e-cigarette advertising reduces cigarette sales, one may question the tobacco companies' short-run incentives to advertise. During the period of my data, Lorillard was the dominant cigarette company in the e-cigarette market and the third-largest tobacco cigarette company with about 12% market share. One explanation for their observed decision to advertise could be that Blu advertising increases Blu e-cigarette sales largely at the expense of other firms' cigarette brands. To explore this possibility, I check whether Blu advertising was higher in markets with lower Lorillard market share. I do not find any significant correlation to suggest that Lorillard coordinated its advertising to target markets where they had relatively less to lose.

Statistics, *68*(1), 148–155.

Becker, G., & Murphy, K. (1988). A rational model of addiction. *Journal of Political Economy*, *96*(4), 675–700.

Benowitz, N. L., & Gourlay, S. G. (1997). Cardiovascular toxicity of nicotine: Implications for nicotine replacement therapy. *Journal of the American College of Cardiology*, *29*(7), 1422–31.

Berry, S., Levinsohn, J., & Pakes, A. (1995). Automobile prices in market equilibrium. *Econometrica*, *63*(4), 841–890.

Berry, S., Levinsohn, J., & 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.

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

Campaign for Tobacco-Free Kids (2017). Cigarette tax increases by state per year 2000-2017.

Card, D., & Krueger, A. (1994). Minimum wages and employment: A case study of the fast-food industry in New Jersey and Pennsylvania. *The American Economic Review*, *84*(4), 772–793.

Caves, K. (2004). State dependence and heterogeneity in aggregated discrete choice demand systems: An example from the cigarette industry. *Working Paper*.

CDC (2018). About electronic cigarettes (e-cigarettes).
URL `https://www.cdc.gov/tobacco/basic_information/e-cigarettes/about-e-cigarettes.html`

Chaloupka, F. J. (1991). Rational addictive behavior and cigarette smoking. *Journal of Political Economy*, *99*(4), 722–742.

Chaloupka, F. J., & Warner, K. E. (1999). The economics of smoking. In J. Newhouse, & A. Culyer (Eds.) *The Handbook of Health Economics*. North-Holland.

Chintagunta, P., & Dubé, J. P. (2005). Estimating a stockkeeping-unit-level brand choice model that combines household panel data and store data. *Journal of Marketing Research*, *42*(3), 368–379.

Crowley, R. (2015). Electronic nicotine delivery systems: Executive summary of a policy position paper from the American College of Physicians. *Annals of Internal Medicine*, *162*(8), 583–4.

Dubé, J. P., Hitsch, G., & Rossi, P. (2010). State dependence and alternative explanations for consumer inertia. *RAND Journal of Economics*, *41*(3), 417–445.

Dubois, P., Griffith, R., & O'Connell, M. (2018). The effects of banning advertising in junk food markets. *Review of Economic Studies*, *85*, 396–436.

Duke, J., Lee, Y., Kim, A., Watson, K., Arnold, K., Nonnemaker, J., & Porter, L. (2014). Exposure to electronic cigarette television advertisements amoung youth and young adults. *Pediatrics*, *134*(1), 1–7.

Erdem, T., Imai, S., & Keane, M. P. (2003). Brand and quantity choice dynamics under price uncertainty. *Quantitative Marketing and Economics*, *1*, 5–64.

Friedman, A. S. (2015). How does electronic cigarette access affect adolescent smoking? *Journal of Health Economics*, *44*, 300–308.

FTC (2015). Federal Trade Commission cigarette report for 2012.

Gordon, B., & Sun, B. (2014). A dynamic model of rational addiction: Evaluating cigarette taxes. *Marketing Science*, *Articles in Advance*, 1–19.

Haar, M. (2015). Nielsen: Electronic cigarette sales growth declines. Is it time to rethink how we track e-vapor?

Heckman, J. J. (1981). Heterogeneity and state dependence. In S. Rosen (Ed.) *Studies in Labor Markets*, (pp. 91–140). University of Chicago Press.

Hendel, I., & Nevo, A. (2013). Intertemporal price discrimination in storable goods markets. *American Economic Review*, *103*(7), 2722–2751.

Hitsch, G., Hortaçsu, A., & Lin, X. (2017). Prices and promotions in u.s. retail markets: Evidence from big data. *Working Paper*.

Houthakker, H., & Taylor, L. (1970). *Consumer Demand in the United States, 1929-1970: Analyses and Projections*. Cambridge (MA): Harvard University Press, 2nd ed.

Huang, J., Tauras, J., & Chaloupka, F. J. (2014). The impact of price and tobacco control policies on the demand for electronic nicotine delivery systems. *Tobacco Control*, *0*, 1–7.

Ippolito, R. A., Murphy, R. D., & Sant, D. (1979). *Staff Report on Consumer Responses to Cigarette Health Information*. Washington: Federal Trade Commission.

Kantar Media (2014). The impact of electronic cigarettes on the smoking cessation industry: What pharma marketers need to know.

Kim, A. E., Arnold, K. Y., & Makarenko, O. (2014). E-cigarette advertising expenditures in the U.S., 2011-2012. *American Journal of Preventive Medicine*, *46*(4), 409–412.

Maloney, E., & Cappella, J. (2015). Does vaping in e-cigarette advertisements affect tobacco smoking urge, intentions, and perceptions in daily, intermittent, and former smokers? *Health Communications*, (pp. 1–10).

McNeill, A., Brose, L., Calder, R., Hitchman, S., Hajek, P., & McRobbie, H. (2015). *E-cigarettes: An Evidence Update*. Public Health England.

Mullahy, J. (1985). *Cigarette Smoking: Habits, Health Concerns, and Heterogeneous Unobservables in a Micro-Econometric Analysis of Consumer Demand*. Ph.D. thesis, University of Virginia, Charlottesville (VA).

Pesko, M. F., & Currie, J. M. (2016). The effect of e-cigarette minimum legal sale age laws on traditional cigarette use and birth outcomes among pregnant teenagers. *NBER Working Paper*.

Petrin, A. (2002). Quantifying the benefits of new products: The case of the minivan. *Journal of Political Economy*, *110*(4), 705–729.

Porter, R. H. (1986). The impact of government policy on the U.S. cigarette industry. In P. M. Ippolito, & D. T. Scheffman (Eds.) *Empirical Approaches to Consumer Protection Economics*. Washington: US Government Printing Offices.

Royal College of Physicians (2016). *Nicotine without smoke: Tobacco Harm Reduction*. London.

Schneider, L., Klein, B., & Murphy, K. (1981). Government regulation of cigarette health information. *Journal of Law and Economics*, *24*, 575–612.

Seldon, B., & Doroodian, K. (1989). A simultaneous model of cigarette advertising: Effects on demand and industry response to public policy. *Review of Economics and Statistics*, *71*, 673–7.

Shan, M., Jump, Z., & Lancet, E. (2012). Urban and rural disparities in tobacco use. In *National Conference on Health Statistics*. American Lung Association.

Shapiro, B. T. (2018). Positive spillovers and free riding in advertising of prescription pharmaceuticals: The case of antidepressants. *Journal of Political Economy*, *126*(1).

Solomon, S. D. (2014). With merger, tobacco takes on technology. *The New York Times*.

Tavernise, S. (2016). Smokers urged to switch to e-cigarettes by british medical group. *The New York Times*.

Tremblay, C. H., & Tremblay, V. J. (1995). The impact of cigarette advertising on consumer surplus, profit, and social welfare. *Contemporary Economic Policy*, *13*(1), 113–124.

U.S. Department of Health and Human Services (2008). Treating tobacco use and dependence: 2008 update. *Clinical Practice Guideline*.

WHO (2014). Electronic nicotine delivery systems.

Zheng, Y., Zhen, C., Nonnemaker, J., & Dench, D. (2016). Advertising, habit formation, and U.S. tobacco product demand. *American Journal of Agricultural Economics*, *98*(4), 1038–1054.

A Fixed Effects Regressions using All Counties

As a comparison to the border strategy results, in this appendix I estimate regressions using data from all counties (border and non-border) in the top 100 DMAs. The unit of observation in this analysis is a county-week, and I include county fixed effects and common week fixed effects. If firms target advertising as a function of market and time varying unobservables, these regressions could suffer from an endogeniety bias. Comparing the border strategy ad elasticities in Table 4 with the fixed effects elasticities in Table 7, the elasticities from the fixed effects regressions appear to have a slight positive bias. Relative to the border strategy analysis, the fixed effects regressions estimate a larger positive effect of e-cigarette advertising on e-cigarette demand and a positive but not statistically significant effect of e-cigarette ads on cigarette demand. These patterns are consistent with firms advertising more in markets during periods of relatively high demand.

Table 7: Fixed Effects Regression Results

	(1) E-Cig Cartridges	(2) Cigarette Packs	(3) Nicotine Patches	(4) Nicotine Gum
E-Cigarette Log Ads	15.49*	39.60	-3.893**	-16.23
	(8.835)	(58.15)	(1.916)	(25.44)
Smoking Cessation Log Ads	-4.594**	26.15	-	-
	(1.868)	(19.58)	-	-
Nicotine Patch Log Ads	-	-	-0.224	-44.27
	-	-	(1.948)	(28.59)
Nicotine Gum Log Ads	-	-	1.088	-20.81
	-	-	(0.835)	(16.89)
Price E-Cigarette Cartridge	-5.273***	16.37*	-0.159	1.265
	(0.968)	(8.342)	(0.106)	(3.028)
Price Cigarette Pack	1.271	-2,225**	5.730***	-217.7
	(37.83)	(941.2)	(1.603)	(140.2)
Price Nicotine Patch	-5.559***	4.302	-14.89***	-119.6***
	(1.647)	(14.57)	(2.296)	(20.44)
Price Nicotine Gum	-44.71***	163.1	-18.42***	-1,068***
	(12.04)	(114.7)	(5.638)	(123.8)
County FE	Y	Y	Y	Y
Week FE	Y	Y	Y	Y
N Obs	324,865	324,865	324,865	324,865
E-Cigarette Ad Elasticity	0.09	0.003	-0.04	-0.004

Clustered standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

B Common Trends Sensitivities

Recall that the difference-in-differences identification strategy relies on the assumption that sales in bordering markets would follow a parallel trend in the absence of differences in treatment. In this section, I re-estimate the descriptive difference-in-differences regressions, restricting to the subsample of markets that have a correlation in weekly cigarette sales in 2010 above $\rho = 0.5$. This is the set of border markets that most closely satisfy the parallel trends assumption in the year before e-cigarettes were first advertised on TV. As shown in Table 8, the effect of e-cigarette advertising is directionally consistent and the magnitude of the effect increases relative to the estimates for the full sample.

Table 8: Difference in Differences Regression Results for the Restricted Sample

	(1) E-Cig Cartridges	(2) Cigarette Packs
E-Cigarette Log Ads	85.40***	-1,546***
	(13.50)	(387.6)
Smoking Cessation Log Ads	-12.27*	-90.92
	(6.253)	(111.9)
Price Controls	Y	Y
DMA-Border FE	Y	Y
Week-Border FE	Y	Y
N Obs	32,222	32,222
E-Cigarette Ad Elasticity	0.16	-0.04

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: Price controls include prices of e-cigarettes, tobacco cigarettes, and smoking cessation products.

C Placebo Test

As an additional robustness check, I conduct a placebo test where I regress e-cigarette and tobacco cigarette sales on TV advertising GRPs for Angel Soft toilet paper.[52] As expected, I find no economically or statistically significant relationship between cigarette sales and advertising for this seemingly unrelated product.

Table 9: Placebo Difference in Differences Regression Results

	(1) E-Cig Cartridges	(2) Cigarette Packs
Angel Soft Log Ads	-4.68	59.9
	(5.11)	(105.7)
Price Controls	Y	Y
DMA-Border FE	Y	Y
Week-Border FE	Y	Y
N Obs	48,968	48,968

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: Price controls include prices of e-cigarettes, tobacco cigarettes, and smoking cessation products.

[52] I was able to obtain data on Angel Soft advertising from 2011–2014, so this regression is on a slightly smaller sample.

D Sensitivity to Changes in Cigarette Excise Taxes

The identification section discussed the fact that changes to cigarette excise taxes could pose a threat to my identification strategy. To check the sensitivity of the results to this potential omitted variable, I tried dropping observations for DMA borders located in states that increased their cigarette excise tax during the period 2011–2015 (Campaign for Tobacco-Free Kids (2017)). I consider two procedures. First, I drop observations for border-weeks corresponding to years and states with excise tax changes. Specifically, I drop all observations for the year in which the border's tax change occurred. Almost all tax changes occurred in June, July, or August, so this procedure would take care of any correlations between advertising and sales leading up to and following the tax change. In the event of a tax change in January, I drop observations for the preceding year too. These results are shown in columns 1 and 2 in Table 10. Second, I drop all observations for all borders that are located within a state that ever changed its cigarette excise tax between 2011–2015. The results are presented in columns 3 and 4. The estimates are consistent with the full sample results shown in Table 4 and if anything, the ad effects are larger in this restricted sample. These results indicate that the estimates for the full sample do not seem to be driven by changes in excise taxes.

Table 10: Sensitivity to Excise Tax Changes

	(1) E-Cig Cartridges	(2) Cigarette Packs	(3) E-Cig Cartridges	(4) Cigarette Packs
E-Cigarette Log Ads	34.08***	-647.1***	56.77***	-1,021***
	(6.063)	(246.8)	(7.921)	(317.1)
Smoking Cessation Log Ads	-0.183	38.84	-4.374	72.64
	(3.935)	(61.95)	(4.598)	(73.87)
Price Controls	Y	Y	Y	Y
DMA-Border FE	Y	Y	Y	Y
Week-Border FE	Y	Y	Y	Y
N Obs	54,850	54,850	36,151	36,151
E-Cig Ad Elasticity	0.09	-0.03	0.15	-0.04

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

E Border Sample Demographics

Two questions arise with respect to the profile of border markets. First, are bordering markets similar on observed demographics? Market fixed effects in the model control for any time invariant differences across bordering markets, but to the extent that ad-sensitivity could be a function of demographics, it is informative to compare demographics for bordering markets. Second, how do border counties compare to the larger DMAs in which they are located? This second question relates to the generalizability or external validity of the estimates. If the demographics of the individuals living in border markets are similar to the general population, then it may be reasonable to think that the causal effect of e-cigarette advertising estimated on the border sample can be extrapolated when making policy decisions.

E.1 Comparison of Bordering Markets

In order to check whether neighboring markets are similar on observed demographics, I calculate border market level demographics by taking the population-weighted average of county-level U.S. census data. For each characteristic I calculate the absolute deviation for each pair of bordering markets and normalize this statistic by the standard deviation of that characteristic across all 282 border markets.[53] The resulting statistic measures the distance in standard deviations between bordering markets. The distributions of these statistics are reported in Table 11. The median pair of bordering markets is within less than half of a standard deviation of each other for most characteristics.

Table 11: Normalized Absolute Deviations in Demographics Across Bordering Markets

	N	Min	Median	Mean	Max
Percent Female	141	0.00	0.71	0.92	5.08
Percent Population Under 18	141	0.00	0.60	0.81	3.63
Percent HS Diploma	141	0.01	0.43	0.60	3.42
Percent White	141	0.00	0.34	0.52	2.74
Percent Black	141	0.01	0.18	0.38	2.60
Per Capita Income	141	0.00	0.41	0.65	4.47
Population Per Square Mile	141	0.00	0.17	0.48	8.06

[53]Absolute deviation in characteristic x for markets i and $j = |x_i - x_j|$. Normalized absolute deviation calculated as $\frac{|x_i - x_j|}{\sigma_x}$.

E.2 Comparison of Border Markets to Non-Border Markets

I compare county-level demographics for border counties to the demographics of non-border counties. The results in Table 12 show that the population of border counties is on average slightly older, less educated, and lower income. Border counties have a lower share of black residents and a lower population density than non-border counties.

Table 12: Average Characteristics in Border and Non-Border Markets

	Border Counties	Non-Border Counties	*p* value
Percent Female	50.18	50.07	0.217
Percent Population Under 18	22.11	22.82	0.000
Percent HS Diploma	83.25	85.35	0.000
Percent White	86.32	84.91	0.037
Percent Black	8.95	10.17	0.056
Per Capita Income	23,085	24,582	0.000
Population Per Square Mile	167.0	524.1	0.000
N Counties	847	1,130	

F Sensitivities to Advertising Stock Carry-Over Rate

This appendix considers the robustness of the results to allowing for the effect of past advertising on current sales. Specifically, I replace current period advertising a_{mt} in equation 1 with a discounted cumulative stock of advertising $A_{mt} = \sum_{\tau=t-52}^{t} \delta^{t-\tau} a_{m\tau}$. Tables 13 and 14 report results for various advertising carry-over parameters δ. Note, column 1 with $\delta = 0$ corresponds to the specification reported in Table 4. For both e-cigarettes and tobacco cigarettes, the implied elasticities are quite consistent across specifications for the models with $\delta \leq 0.6$. The models with $\delta = 0.9$ differ, though the models with $\delta = 0.9$ have a higher root-mean squared error, indicating worse model fit. Based on these analyses, the estimated elasticities reported in Table 4 appear robust to allowing for the effect of past advertising on current sales.

An alternative approach to modeling past ad effects would be to include current and lagged advertising separately. The correlation between current and lagged e-cigarette advertising is high ($\rho = 0.89$), making it hard to separately identify the effects of current and lagged ads. The ad stock model shown in this appendix is frequently used in the literature because it serves as a parsimonious way to capture the combined effect of current and past advertising.

Table 13: E-Cigarette Regressions by Ad Stock Carry-Over Rate δ

	(1) E-Cigarette Cartridges ($\delta = 0$)	(2) E-Cigarette Cartridges ($\delta = 0.3$)	(3) E-Cigarette Cartridges ($\delta = 0.6$)	(4) E-Cigarette Cartridges ($\delta = 0.9$)
E-Cigarette Log Ad Stock	29.77**	37.05***	32.82***	1.162
	(5.644)	(6.072)	(5.593)	(7.147)
Smoking Cessation Log Ad Stock	-6.005	-3.976	-6.406	-20.54
	(4.473)	(5.615)	(8.759)	(23.13)
Price E-Cigarette Cartridge	-8.166***	-8.166***	-8.138***	-8.131***
	(0.988)	(0.988)	(0.987)	(0.985)
Price Cigarette Pack	86.98***	87.35***	87.57***	87.25***
	(12.24)	(12.24)	(12.24)	(12.18)
Price Nicotine Patch	7.013***	6.979***	6.944***	7.043***
	(1.980)	(1.981)	(1.982)	(1.982)
Price Nicotine Gum	-25.23**	-24.96**	-25.14**	-26.53**
	(12.12)	(12.12)	(12.11)	(12.09)
Observations	63,952	63,952	63,952	63,952
RMSE	213.18	213.16	213.17	213.26
Market FE	Y	Y	Y	Y
Week-Border FE	Y	Y	Y	Y
E-Cigarette Ad Elasticity	0.08	0.10	0.09	0.003

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 14: Cigarette Regressions by Ad Stock Carry-Over Rate δ

	(1) Cigarette Packs ($\delta = 0$)	(2) Cigarette Packs ($\delta = 0.3$)	(3) Cigarette Packs ($\delta = 0.6$)	(4) Cigarette Packs ($\delta = 0.9$)
E-Cigarette Log Ad Stock	-631.7***	-791.9***	-773.5***	101.3
	(217.4)	(245.9)	(238.6)	(198.9)
Smoking Cessation Log Ad Stock	-28.19	167.8	416.6*	285.9
	(83.81)	(132.7)	(238.8)	(649.1)
Price E-Cigarette Cartridge	68.48***	68.28***	67.58***	68.10***
	(10.89)	(10.89)	(10.88)	(10.85)
Price Cigarette Pack	-10,128***	-10,134***	-10,140***	-10,124***
	(826.5)	(826.6)	(826.8)	(828.3)
Price Nicotine Patch	-47.87	-47.48	-46.94	-48.88
	(38.28)	(38.31)	(38.32)	(38.20)
Price Nicotine Gum	87.47	83.29	84.30	118.8
	(255.2)	(255.5)	(255.5)	(254.6)
Observations	63,952	63,952	63,952	63,952
RMSE	5207.68	5207.09	5207.03	5209.06
Market FE	Y	Y	Y	Y
Week-Border FE	Y	Y	Y	Y
E-Cigarette Ad Elasticity	-0.02	-0.03	-0.03	0.004

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

G Household-Level Regressions with Prices and Advertising

This appendix explores the robustness of the household-level regressions to including controls for current prices and advertising. Because I only observe prices paid by households when they make a purchase and I don't have any household-level ad exposure data, I bring in price and advertising data from the aggregate data by merging on average prices in each household's county and ad GRPs for each household's DMA. The results are shown in Table 15. Notably, the coefficients on the lagged purchase dummies are very consistent with the coefficients in the more parsimonious model reported in Table 5.

Table 15: Household Addiction and Substitution Patterns Between Cigarettes and E-Cigarettes

	Cig Purchase Incidence	E-Cig Purchase Incidence
Cigarette Purchase in Previous Week	0.080***	-0.001***
	(0.003)	(0.0002)
E-Cig Purchase in Previous Week	-0.034***	0.140***
	(0.008)	(0.012)
Nicotine Gum Purchase in Previous Week	-0.033***	0.002
	(0.010)	(0.003)
Nicotine Patch Purchase in Previous Week	-0.053***	-0.001
	(0.013)	(0.002)
E-Cigarette Log Ads	0.0002	0.0003*
	(0.001)	(0.0002)
Smoking Cessation Log Ads	-0.002	2.92e-05
	(0.001)	(0.0002)
Price Cigarette Pack	-0.005***	-0.0003
	(0.002)	(0.0002)
Price E-Cigarette Cartridge	0.0002	3.6e-05
	(0.0002)	(2.7e-05)
Price Nicotine Gum	-0.002	-0.0003
	(0.006)	(0.0009)
Price Nicotine Patch	0.0005	7.36e-05
	(0.0006)	(9.79e-05)
HH FE	Y	Y
Week FE	Y	Y
N Observations	3,261,333	3,261,333
N HHs	21,808	21,808
N E-Cigarette HHs	2,145	2,145
Mean DV	0.132	0.002
Mean DV if E-Cig Buyer	0.242	0.019
Last Week Cig as % of DV	60.5%	-51.3%
Last Week E-Cig as % of DV for E-Cig Buyers	-13.8%	738.9%

Clustered standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: Magnitude of change post e-cigarette reported as percent of average DV for those households who ever purchase an e-cigarette. E-cigarette users are on average heavier smokers than non e-cigarette users. The average weekly cigarette purchase incidence for e-cigarette users is 0.24 and for non e-cigarette users is 0.12. Standard errors clustered at the household level.

H Model Simulations

I carry out a simulation exercise to illustrate the model's ability to recover the parameters of interest. The steps of the simulation are described below.

In each period consumers decide whether to smoke cigarettes ($c = 1$) or not ($c = 0$). Addiction is captured by allowing today's consumption decision to be related to the consumption state in the previous period through the parameter γ. I assume the following data generating process at the individual level.

$$u_{ict} = \beta_i + \alpha p_t + \phi a_t^e + \gamma \mathbb{I}(c_{it-1} = 1) + \xi_t + \varepsilon_{ict} \tag{18}$$

$$u_{i0t} = 0 + \varepsilon_{i0t} \tag{19}$$

Consumers are assumed to be heterogenous in their preference for cigarettes. The distribution of product intercepts β_i is assumed to be normal, with mean $\bar{\beta}$ and variance σ^2. The parameters of interest are the "linear" parameters $\theta_1 = (\bar{\beta}, \alpha, \phi)$ and "non-linear" parameters $\theta_2 = (\sigma, \gamma)$. Consistent with the full model, I include unobserved aggregate demand shocks in the simulation. In a first simulation, I assume that ξ_t is normally distributed and e-cigarette advertising is uncorrelated with the aggregate demand shocks. In a second simulation, I assume $\xi_t = \beta_t + \eta_t$ can be decomposed into a component β_t that varies systematically over time and a component η_t that is normally distributed. In order to illustrate the joint model's ability to account for endogeneity using the aggregate data, I assume that demand for cigarettes is decreasing over time ($\beta_t \geq \beta_{t+1}$) and advertising is increasing over time such that $Corr(a_t^e, \xi_t) < 0$, making advertising endogenous. Finally, I assume the ε shocks are distributed type 1 extreme value. The model-predicted aggregate market share of cigarettes is given by equation 20.

$$s_{ct} = \int_{\Theta \times \{0,1\}} \pi_{it}(c \mid k) dF_t(\theta, k) \tag{20}$$

In estimation, I approximate the distribution of heterogeneity with $R = 100$ draws from the standard normal distribution $\nu_r \sim N(0, 1)$ s.t. $\beta_r = \bar{\beta} + \sigma \nu_r \sim N(\bar{\beta}, \sigma^2)$ and evaluate the integral in equation 20 using Monte Carlo integration.

I simulate purchase decisions for 10,000 consumers in each of $T = 150$ periods. Aggregate market shares in each period are calculated using the full set of households. A 1% random sample of households makes up the household-level dataset used for estimation. I estimate the model parameters (i) via maximum likelihood using only the household data and (ii) using the joint estimation procedure and both the aggregate and household datasets. I include time fixed effects that control for the endogeneity of advertising in the final linear regression step in the joint estimation procedure. Because of the parameter proliferation problem, including these fixed effects in the household model is intractable. I carry out the simulation $NS = 1,000$ times and compare the results across models.

As shown in Table 16 and Figures 13 and 14, both estimation procedures perform quite well in recovering the “non-linear" model parameters $\theta_2 = (\sigma, \gamma)$. In the simulation with exogenous advertising, the joint procedure is more efficient in recovering the “linear" model parameters $\theta_1 = (\bar{\beta}, \alpha, \phi)$ because it incorporates the full information contained in the aggregate data. In the simulation with endogenous advertising, the joint procedure recovers an unbiased estimate of the advertising coefficient, while the model using only household data recovers biased estimates because of the persisting advertising endogeneity.

Table 16: Model Simulation Results

		Exogenous Ads		Endogenous Ads	
	True Values	HH ML	Joint Est	HH ML	Joint Est
$\bar{\beta}$	-0.5	-0.4844	-0.4999	-0.6968	-0.4967
		(0.2941)	(0.1810)	(0.3206)	(0.1819)
α	-0.6	-0.5982	-0.6009	-0.6358	-0.6012
		(0.0957)	(0.0536)	(0.1054)	(0.0538)
ϕ	-0.02	-0.0200	-0.0201	-0.0212	-0.0201
		(0.0020)	(0.0010)	(0.0022)	(0.0010)
σ	0.2	0.1881	0.1927	0.1763	0.1899
		(0.0488)	(0.0495)	(0.0656)	(0.0624)
γ	1.75	1.7428	1.7562	1.8073	1.7554
		(0.0653)	(0.0724)	(0.0819)	(0.0905)

Figure 13: Distribution of Estimates in Simulation w/ Exogenous Ads

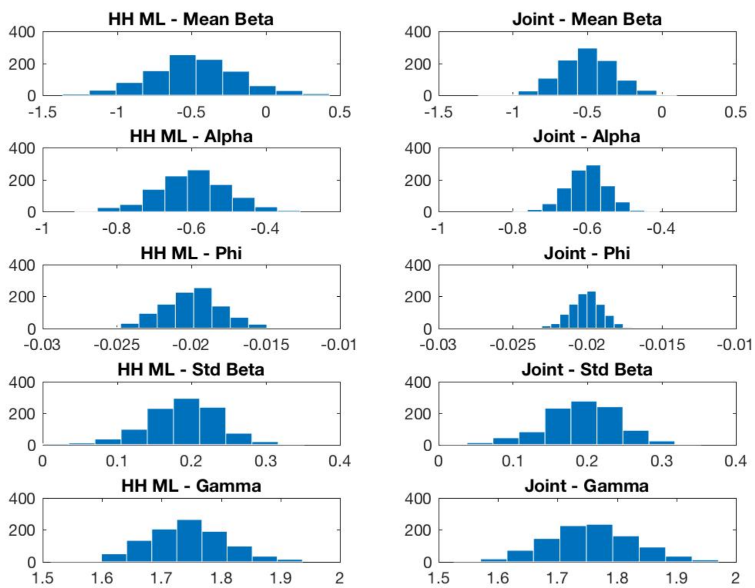
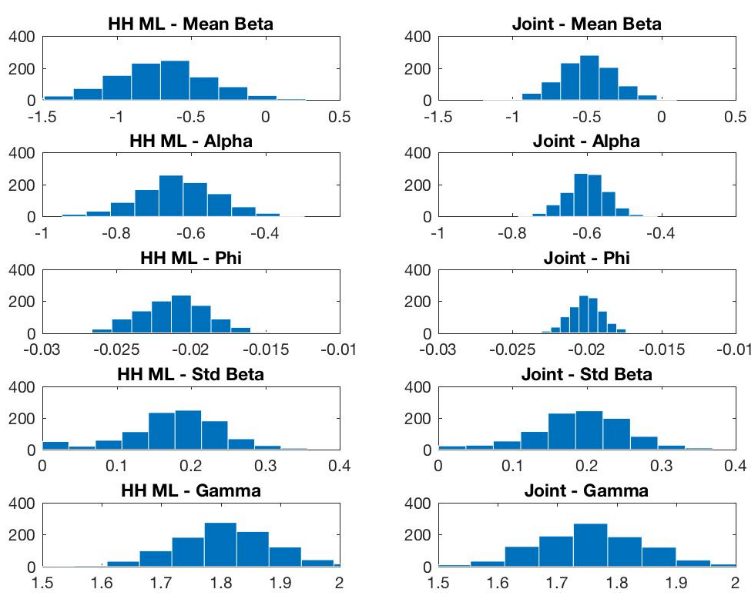


Figure 14: Distribution of Estimates in Simulation w/ Endogenous Ads



I Data Appendix

I.1 Aggregate Price Series Construction

A market-level price series is constructed for packs of cigarettes, refill cartridges / disposable e-cigarettes, and smoking cessation products. In each case, the price is constructed on a per unit basis, where a unit is a pack of cigarettes, a single cartridge / disposable e-cigarette, and a single piece of nicotine gum or a single nicotine patch. The price series is calculated as the quantity-weighted price of products across all UPCs and stores in a given market.

In the structural model, I further aggregate the two types of smoking cessation products together because both products have small market shares. Here, I define an equivalent unit for nicotine gum as 10 pieces of gum and for nicotine patches as a single patch because these quantities yield roughly the equivalent nicotine content to a single pack of cigarettes. Thus, I calculate the aggregate market share for smoking cessation products as the total number of nicotine patches sold plus 0.1 times the number of pieces of gum sold, divided by the market size. The aggregated smoking cessation price series is computed as the price for such an equivalent unit, and the aggregation is again carried out using quantity sales as weights.

I.2 Rationalizing Multiple Purchases with Discrete Choice

In some cases, a household will buy multiple different products from the choice set or multiple units of a given product in a given period. The former case happens very infrequently in the household-level data: a purchase of multiple different inside goods occurs in only 0.4% of weeks that have a purchase. In the estimation dataset, I create duplicate entries for those weeks, one for each inside good that was purchased. If both an e-cigarette and a tobacco cigarette were purchased in the same week, I randomly set the state variable to either the e-cigarette or the tobacco cigarette state for the following week's observation, since it is inconsistent for a household to be in both states at once. If a consumer buys both a smoking cessation product and either an e-cigarette or a tobacco cigarette in the same week, I again create duplicate entries for those weeks, one recording each purchase. For the following week's observation, I set the state variable to the cigarette product that was purchased in the previous week.

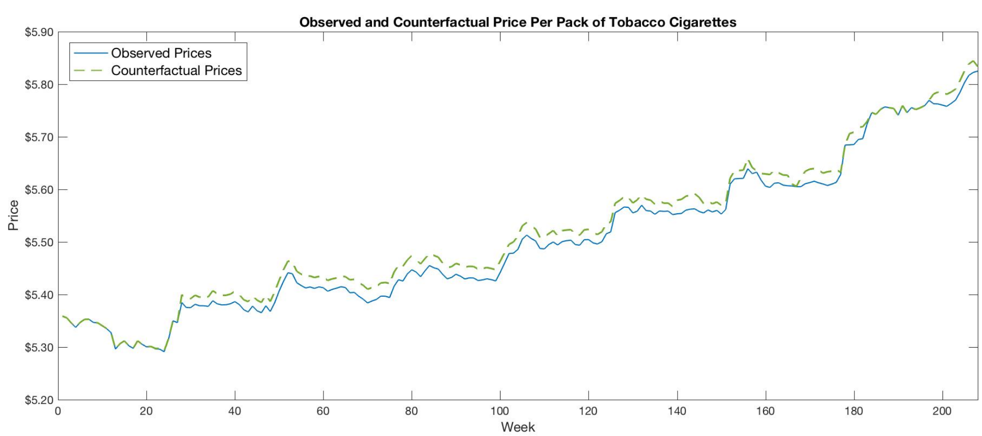
In the weekly household data, a single pack of tobacco cigarettes is purchased in 11% of weeks in which at least one pack was purchased. This suggests that it is common for smokers to purchase more than one pack in a week. In the structural model, I abstract away from quantity and focus on incidence – whether at least one unit was purchased. Thus, a household that buys one or more packs of cigarettes in a given week is recorded as having purchased a tobacco cigarette product that week. In the aggregate data, I cannot observe how many consumers purchased in a given store-week. This is a standard challenge when working with aggregate data. The discrete choice model essentially thinks of each unit in the aggregate data as being purchased by a single consumer. Thus, the sale of 10 packs of cigarettes in a given week would imply that there are 10 "addicted" consumers in that market for the next period. I model households' purchase incidence rather than purchase quantity because the aggregate data is necessary to measure ad effects, but models of purchase quantity require strong assumptions to make the models tractable with aggregate data.

J Counterfactual Prices

This appendix first explains how I predict counterfactual prices and then discusses how reasonable these counterfactual prices are. First, under the assumption that observed prices are set optimally, I infer marginal costs given the observed prices, market shares, and price sensitivity parameter α. A limitation of this analysis is that tobacco cigarette price elasticities are predicted to be inelastic (between -1 and 0) in many markets, which can lead to negative marginal cost estimates. This is a problem that previous researchers have resolved by either i) adding supply-side moments that require marginal costs to be positive and thus force demand to be more elastic or ii) specifying a dynamic model that can justify why firms might set low prices. The first approach is not well-suited to my estimation procedure that utilizes maximum likelihood as opposed to GMM and the latter approach would add significant complexity to this robustness check. Balancing these trade-offs, I move forward with the static supply side model and think of the negative marginal cost estimates as costs that are adjusted for discounted future revenue streams. Given the estimated costs for all three products, I then solve for the new prices that would maximize profits when e-cigarette ads are banned. I compare the resulting counterfactual prices to the observed prices and find that the predicted change in prices is small. Below I discuss how credible the resulting predicted prices are.

Figure 15 plots the observed and counterfactual average price of a pack of cigarettes, where the simple average is taken across all markets. The blue line in the graph shows the observed prices and the dashed green line shows the predicted counterfactual prices.

Figure 15: Observed and Counterfactual Average Price Per Pack of Tobacco Cigarettes



First looking at the observed price series, there are clear price increases that occur roughly twice a year. These correspond to manufacturers' bi-annual changes to wholesale prices. During the four year counterfactual period (2012–2015), Altria increased cigarette list prices 8 times.[54] Each increase was either 6 or 7 cents per pack. When Altria announced list price increases, Reynolds and Lorillard typically responded within a couple of days by raising their wholesale prices by a similar amount. The graph shows that these list price increases were largely passed on to consumers at the point of sale.

Turning to the counterfactual price series in green, the supply-side analysis predicts that the price of cigarettes would have been on average 1.6 cents higher if there were no e-cigarette advertising on TV. In evaluating whether the magnitude of the difference between counterfactual and observed prices seems reasonable, cigarette pricing before e-cigarettes gained popularity can serve as a point of comparison. The observed wholesale price increases in 2010–2011 were 1 to 2 cents larger than the observed wholesale price increases between 2012–2015.[55] The slightly higher counterfactual prices would thus be consistent with the observed pricing behavior in the period before e-cigarette advertising became prevalent. These additional facts suggest that the predicted counterfactual prices are reasonable.

The counterfactual predicts that the average price of smoking cessation products would have been slightly higher in the absence of e-cigarette advertising, but the maximum predicted price increase is less than 0.5 cents per unit. The observed average price for a unit of cessation product (1 patch or 10 pieces of gum) was about 4 cents lower in 2012 (after e-cigarette advertising picked up) compared to an average price of $4.26 in 2010–2011. Thus, the counterfactual suggests that only a small part of the observed price decrease was driven by e-cigarette advertising.

It is harder to evaluate how credible the predicted counterfactual prices are for e-cigarettes because there is very little data on e-cigarette prices in the absence of e-cigarette advertising. However, the directional prediction that counterfactual e-cigarette prices would be lower than observed prices seems credible.

[54] Price increases went into effect on 6/18/2012, 12/3/2012, 6/10/2013, 12/1/2013, 5/11/2014, 11/16/2014, 11/16/2014, 5/17/2015, and 11/15/2015.

[55] Altria increased their list prices by 8 cents on 5/10/2010 and 12/6/10, by 9 cents on 7/8/11, and by 5 cents on 12/12/11. All 8 observed list price increases between 2012–2015 were either 6 cents or 7 cents.