THE UNIVERSITY OF CHICAGO

ESSAYS ON THE INFLUENCE OF MEDIA

A DISSERTATION SUBMITTED TO
THE FACULTY OF THE UNIVERSITY OF CHICAGO
BOOTH SCHOOL OF BUSINESS
IN CANDIDACY FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

BY
MICHAEL THOMAS

CHICAGO, ILLINOIS
AUGUST 2017

CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

I thank my advisers Günter Hitsch, Pradeep Chintagunta, Bradley Shapiro and Sanjog Misra for their support and guidance. I also thank Matthew Gentzkow and Jesse Shapiro whose training set the ground work for the chapter on smoking and television.

ABSTRACT

CHAPTER 1: Consumers' response to mass-media can be difficult to assess because individuals choose for themselves the amount of media they consume and that choice may be correlated with their other consumption decisions. To address this important confound, this paper examines the introduction of television to the U.S. during which some cities gained access to television years before others. This natural experiment makes it possible to estimate the causal impact of television on the decision to start smoking, a consumer behavior with important public-health implications. Difference-in-differences analyses of television's introduction indicate that (1) television did cause people to start smoking, (2) 16-21 year-olds were particularly affected by television, and (3) the response to television occurred within a couple of years of its introduction. Estimates from this analysis indicate that television increased the share of smokers in the population by 3%-21% in cohorts exposed to television through age 25. More broadly, these results offer causal evidence that (1) mass-media can have a large influence on consumers, potentially affecting their health, (2) media exerts an especially strong influence on teens, and (3) mass media can influence consumers more than typical changes in prices.

CHAPTER 2: Many advertisements, such as national television ads, are purchased for large populations, preventing marketers from perfectly targeting all subsets of the population. Additionally, researchers increasingly have the data required to measure which populations have received too much or too little advertising. In this paper, I argue that, together, these events generate a natural experiment which can be used to obtain consistent estimates of the response to advertising. I present a formal model for exploiting this natural experiment and report the results of its application to multiple product categories. Estimates from this "coarseness" strategy are consistent with recent literature, suggesting many standard

approaches to estimating the response to advertising may produce misleading results due to unobservables.

CHAPTER 1

WAS TELEVISION RESPONSIBLE FOR A NEW GENERATION OF SMOKERS?

1.1 INTRODUCTION

Smoking for the first time can initiate a lifetime of addiction with potentially severe health consequences. For this reason, academics have taken a keen interest in understanding the factors that drive smoking initiation, including the role of the media. A positive association between media and smoking has been reported in numerous studies from across disciplines, employing diverse data sources and methodologies, investigating various types of media and populations.[1] In addition to reporting positive associations, these studies also all rely on (non-experimental) observational data – and with good reason. As described in a meta-analysis of 51 such studies by Wellman et al. (2006), "Ethical considerations preclude experimental studies that would expose youth to marketing and media to see whether they would initiate tobacco use."

Unfortunately, without experimental evidence, researchers cannot conclusively determine whether media exposure causes individuals to start smoking. This limitation is not just a technical quibble. With the existing evidence, researchers simply cannot distinguish between a world in which media causes people to start smoking, and another plausible world in which people who like to consume a lot of media also tend to like to smoke. Further raising doubt as to the causal link between media and smoking are field experiments on television advertising that demonstrate a small or non-existent response in the consumption of other products (see Lodish et al. [1995] and Hu et al [2009]). If the observed associa-

1. See section 2.2 for a literature review.

tions between media and smoking are simply a result of correlated preferences, the existing, wide-spread bans on cigarette advertising would be useless.

To address this gap in the literature and provide compelling evidence of mass media's influence on smoking, I examine a natural experiment. Using this natural experiment, I can exploit cross-sectional differences in the timing of television's introduction across cities, which allows me to compare groups of individuals who either did or did not have access to television at a given age and point in time. I can then observe whether television access influenced their smoking behavior. This approach allows for estimates of the effect of television on smoking without the standard concern that individuals chose for themselves the amount of media they consumed. This analysis indicates television did have a significant impact on smoking, raising the probability of becoming a smoker by 0.6%-2.8% for each year of exposure between the ages of 16 and 21.

These estimates are produced using data from the 1965-66 National Health Interview Survey (NHIS). The NHIS offers data on individuals' smoking status, age, and geographic location, which allows us approximate when each individual first gained access to television and therefore estimate its influence on their smoking behavior. I examine these data using a difference-in-differences analysis that differs from the standard approach in a couple of ways: first, I disaggregate the model to make use of the rich, individual-level data set; second, I include additional controls to ensure the non-random order of television entry across cities does not confound the estimates. This approach allows me to generate plausible control and treatment groups of individuals whose age of first television access was determined by exogenous delays in television's introduction.

This research design allows identification of the cross-age differences in smoking-uptake responses to media exposure. Estimating these heterogeneous responses to media across ages provides better predictions of the long-run impact of media than existing aggregate estimates. Aggregate estimates fail to account for the influence of high-risk ages,

which I estimate to be about 16-21, in which people experience much larger responses to the media than they do at other ages. Because all members of the population pass through these high-risk ages, aggregate estimates understate the true effect of media exposure by assuming the entire population has a common response.

Alternatively, I find television had a profound effect on smoking initiation. My estimates predict the share of smokers in cohorts that were exposed to television up to age 25 were 3%-21% higher than they would have been without television, meaning television had generated an additional 3-19 million smokers by 1971 when the advertising ban was imposed. These magnitudes are not only economically meaningful, but also large relative to the estimated influence of cigarette prices. Indeed, much of the literature finds prices do not influence initiation. Even the largest estimated price response I am aware of (Sen and Wirjanto [2009]) suggests a price drop of 34% in cigarettes would be required to match television's impact.

To further understand television's influence on smoking, I investigate the timing of the response using a hazard analysis. Here I examine the 1970 NHIS, which reports the age respondents started smoking, and find that after controlling for age and calendar-year effects, the risk of becoming a smoker increased suddenly when television became available in an individual's city. More formally, my hazard model estimates the rate at which non-smokers converted to smokers increased by 10%-30% over the two years following TV's introduction in a city. Estimates for years prior to television entry show no evidence of pre-trends. These results indicate the response to the new medium was quite fast, and offer supporting evidence of television's influence on smoking.

The historical setting examined in this study presents both strengths and weaknesses relative to the existing literature. Unlike the body of existing work, this study estimates the impact of access to just one medium: television. Furthermore, the approach is not able to denominate television's impact in terms of hours watched, or empirically differen-

tiate between the possible channels of televisions influence: spot advertisements, product placement, non-sponsored smoking by stars, etc. Such nuances require richer data environments. Instead, this study makes its contribution by avoiding the selection concerns that arise when individuals choose for themselves the amount of media they consume, concerns that, to our knowledge, arise in all existing studies. By presenting causal evidence of television's influence, this study complements existing work by supporting the possibility that the associations they report reflect causal relationships, and therefore their findings offer useful guidance for policy decisions. Additionally, this paper's results confirm the long-standing suspicion that teens' smoking behaviors may be particularly sensitive to media influence. Finally, this paper demonstrates the response to mass-media communications can be quite fast, and can be large relative to the influence of prices.

The remainder of this paper is organized as follows. Section 2.2 presents an overview of the existing literature on the relationship between media and smoking. Section 1.3 describes the natural experiment used for this study, along with its historical context. Section 1.4 provides details on the data I use. Section 1.5 examines at which ages people are more likely to start smoking as a result of media exposure, and section 1.6 analyzes the hazard of becoming a smoker around the time of television entry. In section 1.7, I estimate the long-run effects of television on the share of smokers in the population. I discuss the implications of my estimates in section 1.8, and conclude in section 1.9.

1.2 LITERATURE REVIEW

Although studies of the cigarette industry have come from a variety of disciplines with different objectives, many have focused on whether media exposure affects smoking behavior. This section briefly reviews a number of these literatures: studies of aggregate consumption

using time-series analysis and cross-country data, plus studies of individual-level consumption from both the economics/marketing literatures and the medical literature.[2] Many of these studies find a positive association between media and smoking, but provide limited evidence that it is a causal relationship.

A large econometrics literature has estimated advertising elasticities in the cigarette industry using time-series data and found a mix of significant and non-significant effects. Earlier papers in this literature estimated the marginal advertising elasticity, but later work has questioned the relevance of marginal-effects estimates and looked to other means of estimating advertising effectiveness by introducing more sophisticated modeling assumptions, panel data, and instrumental variables.[3] The literature as a whole has estimated an average advertising elasticity of 0.8 according to a meta-study by Andrews and Franke (1991). Although the 1971 ban on television and radio advertising might have offered a means of estimating advertising elasticities with weaker assumptions than those required in time-series studies, the ban's impact is confounded by the simultaneous termination of large-scale anti-smoking advertisements that the Federal Communications Commission (FCC) required from 1967 under the Fairness Doctrine.

As potentially more robust alternative to these time-series studies, cross-country studies of media bans offer contemporaneous control and treatment groups. Unfortunately, these also suffer from a limited number of observations and have also produced mixed results.[4]

---

2. See Chaloupka and Warner (2000) for a survey of the economics literature. See Wellman, et al. (2006) for a meta-study of the medical literature.

3. See Telser (1962), Schmalensee (1972), Hamilton (1972), Schneider et al. (1981), Abernethy and Teel (1986), Baltagi and Levin (1986), Roberts and Samuelson (1988), Seldon and Doroodian (1989), and Wilcox and Vacker (1992).

4. Hamilton (1975) finds no effect of bans prior to 1973, and similarly, Laugesen and Meads (1991) find no effect of bans prior to 1973, but find greater drops in consumption associated with their "adverting restriction score" from 1973 onward. Stewart (1993) finds no significant change in consumption in countries that imposed complete media bans from 1964 to 1990.

Most recently, Saffer and Chaloupka (2000) estimate marginally significant effects of comprehensive bans on advertising from 1984-1992, but not for earlier time periods or partial bans. They argue that, unlike partial bans, comprehensive bans leave firms no outlets to substitute their advertising, and estimate that comprehensive bans reduced total cigarette consumption by 7.1%.

Micro-data have been more successful at revealing significant associations between media exposure and smoking. In the economics literature, Lewit et al. (1981) study the effect of anti-smoking advertising on teen smoking. They find the positive correlation between TV-viewing hours and the likelihood of smoking as reported by teens weakened after 1967, and the authors suggest the significant increase in televised anti-smoking advertisements from 1967 under the Fairness Doctrine is responsible. Also, in the marketing literature, Pollay et al.'s (1996) findings suggest advertising may be more influential on teens than adults. They estimate brand-specific advertising elasticities that are three-times larger for teens than for adults, based on correlations between brand shares and advertising shares between 1979 and 1993. Finally, in the medical literature, Pierce et al. (1994) use the NHIS to estimate smoking uptake across time, and find an increase in uptake by women in the late 1960s and early 1970s, about the time sales and advertising of women's cigarette brands also increased.

The medical literature is perhaps the largest source of micro-based observational studies of the association between childhood television exposure and smoking outcomes. Gidwani et al. (2002) find a positive correlation between the number of hours watching TV and being a smoker as reported by 10-15-year-olds, even after controlling for numerous demographic characteristics. Gutschoven and Bulck (2004) report a positive association for teens between hours of television watched and quantity of cigarettes smoked for teens. Hancox, Milne, and Poulton (2005) examine longitudinal data over the course of the life of 1,000 New Zealanders and find an association between an individual's childhood television

exposure and his or her smoking status at age 26. For the influence of movies see Dalton et al. (2003), Dalton et al. (2009), and Sargent (2005).

Despite the repeated findings of a positive association between media and smoking in a variety of empirical environments, clear evidence of a causal relationship has remained elusive. The difficulty is highlighted by large meta-study of the medical literature by Wellman et al. (2006), which examines the association between smoking and a variety of types of media exposure including advertisements and positive depictions of smoking in television, movies, and videos. Although the authors report strong evidence of an association between smoking and media exposure, they must rely on circumstantial evidence to interpret these findings as causal. For example, they point to experiments that show media exposure leads to more positive attitudes toward smoking, and that violence in films and video games leads to violent behavior in children.

This paper aims to fill this gap in the literature, and provide more direct, convincing evidence of mass media's influence on smoking behavior.

1.3 A Brief History of Television and Tobacco

In this section, I present evidence of the significant presence of tobacco from the earliest days of television. I also detail television's introduction to the United States and key events that caused delays for some cities – delays I will later use to examine television's influence on smoking behavior. Finally, I discuss other events over this period that may have influenced smoking behavior, and argue the observed volatility in smoking behavior underscores the need for a natural experiment such as the one provided by television's entry to the United States.

Viewers experienced cigarettes as an integral part of television from early broadcasts in the 1940s at least until the ban on cigarette advertising in 1971 (see Figure 1.2 for adver-

tising expenditures over time). Undoubtedly, viewers experienced cigarettes on television outside of advertising, but advertising offers some of the clearest evidence of its presence. For example, the "Dancing Pack" commercial from 1948 featured a woman dressed in a pack of cigarettes, with only her legs showing as if it were one of the acts on the Original Amateur Hour. Similarly, the eponymous host of Arthur Godfrey and Friends from 1949 could be seen chain-smoking Chesterfield cigarettes on the air, and – during an occasional sponsor announcement – ostensibly dismiss his advertising script to offer a personal endorsement of the brand. Whereas sponsorship of early television was typically a season-long investment, cigarette advertisers also sponsored sporting events and bought spot advertising, together creating a presence that did not go unnoticed. According to Television Magazine 1948-1949, cigarette brands consistently topped "sponsor identification" tables based on surveys that asked respondents to spontaneously list three advertisers on television.

These experiences did not immediately reach the entire U.S. population, however, due to a couple of interventions by the FCC (see Figure 1.3). The FCC issued the first licenses for commercial television in 1941 for a few cities, but shortly thereafter banned further station development in order to focus resources on World War II. Following the war, those cities with station infrastructure already in place began to rapidly adopt television. By the end of the 1940s, more cities had developed the infrastructure to offer broadcasts, but in 1948, the FCC imposed a "freeze" on any further licenses in order to address issues with spectrum allocation, leaving many cities without access to television. The freeze was not lifted until 1953, at which point nearly all of the remaining population received television within a couple of years, [5] These events produced cross-sectional differences in television

5. See Gentzkow and Shapiro (2008) for further discussion of television's introduction and rapid penetration.

Figure 1.1. TOTAL ADVERTISING EXPENDITURE

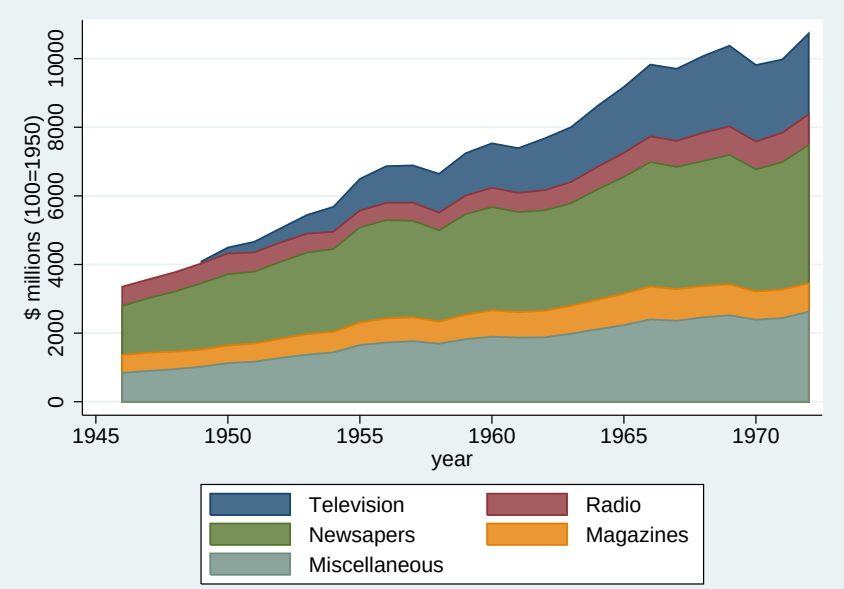
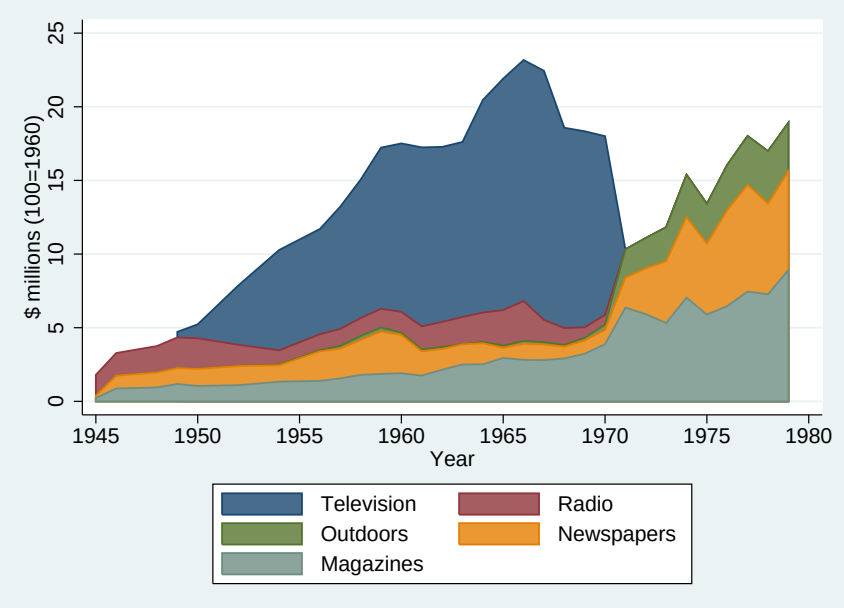
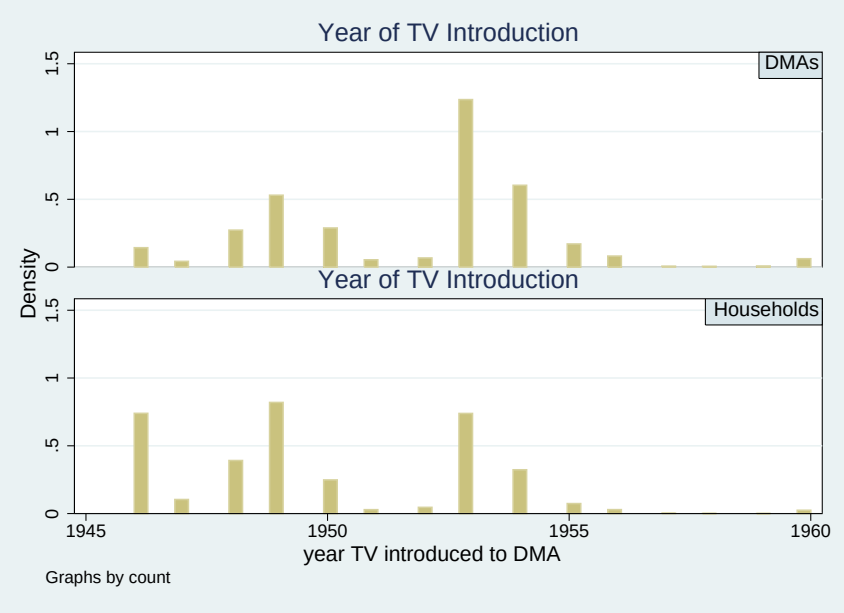


Figure 1.2. TOBACCO ADVERTISING EXPENDITURE



Note: Tobacco Expenditure is total expenditure by the top six cigarette firms, which includes a small amount of non-tobacco advertising. Data are from Advertising Age.

Figure 1.3. TV ENTRY HISTOGRAM



Note: Weighted by DMAs and Number of Households.

access that allow us to construct plausible control and treatment groups we can compare to estimate the effect of television on smoking.

Such natural experiments are rare. For example, we cannot construct such control and treatment groups for many of the other potentially influential events in this era, which include World War II, the "health scare," the cigarette industry's response to health concerns, and later government intervention (see Figure 1.4). Unlike television's introduction, these events affected all of the United States simultaneously. This simultaneity causes difficulty disentangling the influence of these events from one another and from other, unobserved factors that may explain the significant changes in smoking behavior we observe over this period.

I briefly discuss these events because they provide the historical backdrop to my analysis and because significant prior work has attempted to estimate their impact on smoking behavior (see section 2.2). Figure 1.4 shows that during World War II, consumption increased by 44% and continued to inch upward after the war, before reaching a peak in

1952. Then, in 1953, mounting health concerns came to a head as new medical studies strengthened cancer links, and by 1955, per-capita consumption had fallen by 7.8% from its 1952 peak. The cigarette industry responded by funding its own research on health effects, increasing its advertising, and expanding its lines of filtered cigarettes. By the late 1950s, per-capita consumption surpassed the 1952 peak and reached an all-time high in 1963.

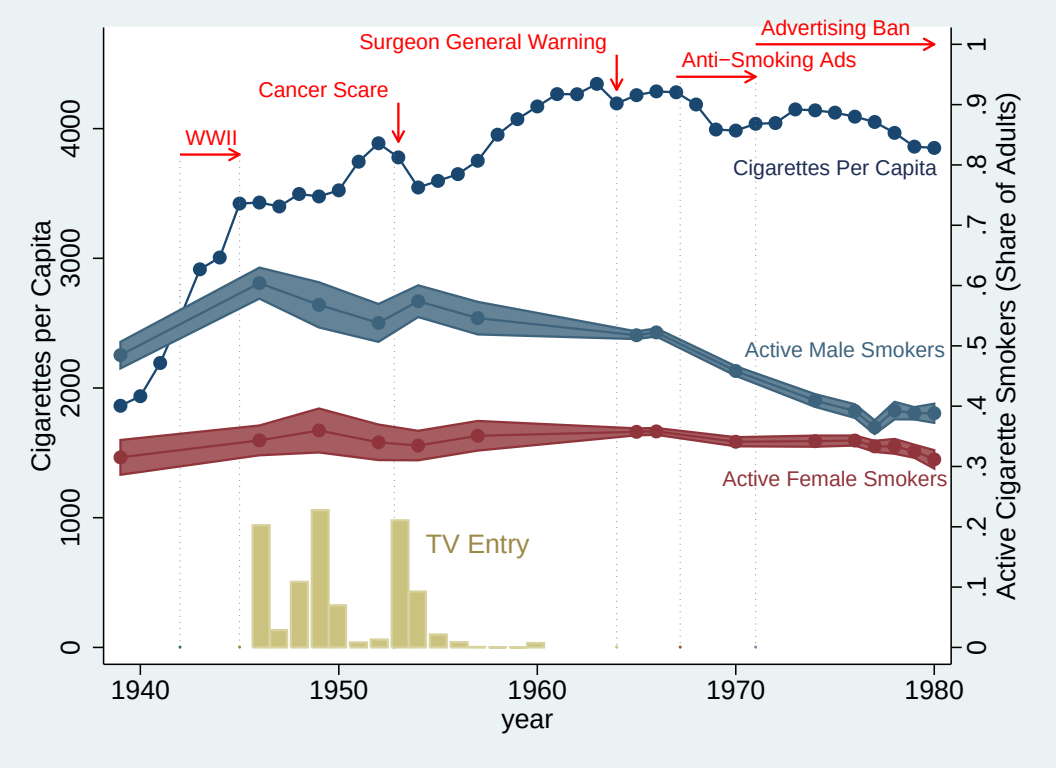
At this point, government involvement may have begun to influence smoking. First, in 1964, the Surgeon General announced the established causal link between smoking and cancer and per-capita consumption dropped for the first time in over a decade. Next, from 1967 to 1971, the FCC's interpretation of the Fairness Doctrine forced television and radio stations to air anti-smoking advertising, and per-capita consumption dropped further still. Finally, in 1971, the FCC ended the anti-smoking advertisements and imposed a complete ban on cigarette advertising on radio and television. The ban remains in place today.

Although these events generate interesting questions about their influence on smoking, I focus on the impact of television's introduction, because of the compelling experimental conditions it offers.

1.4 DATA

The National Health Interview Survey (NHIS) is my primary data source. It reports individuals' smoking status and the age they first smoked regularly, which provide our dependent variables. Additionally, the NHIS reports age and geographic information for individuals, which, combined with the television-entry dates from Gentzkow (2006), allow me to determine the age at which a respondent first had access to television. Together, these data allow me to estimate whether access to television influenced individuals' self-reported smoking status later in life.

Figure 1.4. SMOKING TRENDS



Note: The dotted line represents per-capita consumption in the United States as reported by "The Tobacco Situation," which is produced by the U.S. Department of Agriculture. Above this time trend, the dates of major events that may have affected smoking are indicated in red. The banded series in the middle are the reported share of active smokers in the adult population for each gender. Gallup Polls are used for 1939-1957, and report the share selecting "cigarettes" to answer the question "Do you smoke?" NHIS data are used for 1965-1980 and report the share that report smoking one or more cigarettes per day. 95% confidence intervals are shown for each and are adjusted for the NHIS survey design, but not for the Gallup. Along the bottom (labeled "TV Entry") is a histogram of TV entry dates, weighted by population.

Outcome variables for this study come from two NHIS questions. The 1965-66 surveys asked all respondents 17 and older, "Have you smoked 100 cigarettes in your entire life?" This question provides this paper's definition of a smoker and functions as the dependent variable in section 1.5. The 1970 survey asked, "At what age did you start smoking regularly?" I use this question for the hazard analysis in section 1.6. Each of these surveys come from the NHIS' continuous sampling of the civilian, non-institutionalized U.S. population. Although this NHIS sampling has continued since 1957, the questions and procedures have changed over time. The 1965-66 surveys were conducted between July 1964 and July 1966; the 1970 survey was conducted over its calendar year.

Geographic information from the NHIS allows us to determine the Designated Marketing Area (DMA) an individual lived in at the time of the survey. The NHIS contains Standard Metropolitan Statistical Area (SMSA) identifiers for respondents living in a 1960-definition SMSAs. For surveys before 1973, only the Census Region is known for individuals living outside of an SMSA, so we exclude these individuals from the analysis. We match the known SMSAs to DMAs, geographic regions with common television reception, which Gentzkow and Shapiro (2008) treat as having common television-entry dates. Although nearly all SMSAs lie within the DMA definitions, the few that cross DMA borders are assigned to the DMA that contains most of its population.

The age at which a respondent first had access to television can be approximated using the age and location they report in the NHIS. Specifically, I set age of first television access equal to an individual's age at the time of the survey minus the number of years that had passed since his DMA first received television. This calculation assumes the respondent's DMA at the time of TV entry is the same as when he was interviewed; as a result, migration that occurred between TV entry and the survey biases our estimates towards zero

Respondents were selected for the NHIS with a multi-stage sampling design. The first stage samples from Primary Sampling Units (PSUs) that are composed of SMSAs and other

clusters of one or two contiguous counties. From the universe of 1,900 PSUs, 357 were selected. Within the selected PSUs, further sampling steps were applied, eventually leading to final sampling units called "segments," which consist of six or nine nearby households. All members of these households were interviewed, though some questions could be answered by a "proxy" (i.e., by a family member).

For this analysis, I exclude veterans of World War II and the Korean War because these wars occurred near or during TV's introduction to the United States and caused many servicemen to start smoking while away from home. Bedard and Deschênes (2006) estimate that service in these wars increased the smoking rate by 30%, a large effect that, if included, would obscure the civilian response to television. The excluded veterans represent about 30% of the male respondents in these surveys.

1.5 Age Groups that Responded to Television

In this section, I investigate whether exposure to television at some ages increases the chance of becoming a smoker more than exposure at other ages. The analysis exploits the fact that television entered different parts of the United States at different times due to exogenous policy interventions discussed in section 1.3. These interventions allow for comparison of the smoking status of people who either did or did not have television available to them at a given age and point in time.

I start in section 1.5.1 by presenting model-free evidence suggestive of television's influence on smoking by age groupings. I discuss potential confounds to this analysis, and in section 1.5.2, I introduce a model that allows for more robust estimates of television's influence on smoking, by using only variation from the natural experiment to estimate television's effect. Finally, section 1.5.3 presents estimates using this model.

1.5.1 Model-Free Evidence of Responses by Age Group

To start thinking about how television might have varying effects across ages on people's decision to start smoking, consider Figure 1.5, a histogram of the ages at which people report having started smoking. For both men and women, the vast majority of smokers start in their late teens or early 20s. If these years are the critical ages at which a person decides whether to become a smoker, we may hypothesize that these years are also the ages at which television is most influential on smoking uptake.

First consider simple evidence: the smoking-rate averages by the age at which people first gained access to television. For this step, I create three groups from our sample: those who received television before age 16, between 16 and 21, and after 21. For each of these groups, I compute the share of people in the 1965-66 NHIS who report having smoked 100 lifetime cigarettes. I find that 58% of respondents who received television after age 21 report having smoked 100 cigarettes in their lifetime, but that number goes up to 62% for those who received it between ages 16-21, and rises to 63% for those who received it after age 21. Although suggestive of television's influence, these simple comparisons may ignore some potential confounds. The most obvious confound comes from the strong correlation between birth cohort and age of first television access. This concern can be addressed using the natural experiment described in section 1.3.

As a result of this natural experiment, some DMAs received television years before others, which allows for comparison of the smoking status of people who received television at different ages within identical birth cohorts. Such an analysis is presented in Figure 1.6, which shows each of the age-of-first-access groups by the year the individual was born. Here we can see that whereas the share of smokers tends to increase across birth cohorts, within a cohort, those who were exposed to television earlier in life were more likely to be smokers. This pattern does not hold for every birth cohort, but it holds for the vast majority.

Figure 1.5. AGE STARTED SMOKING REGULARLY, REPORTED

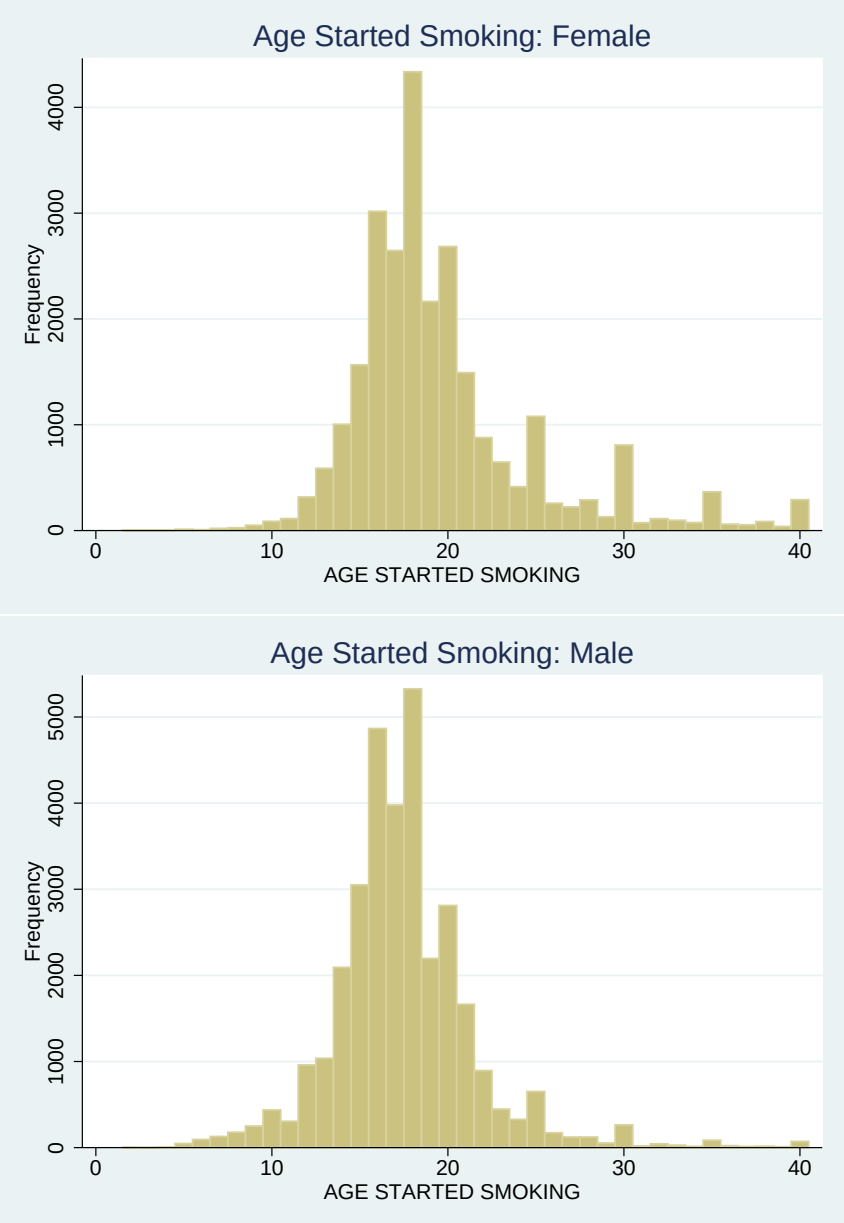
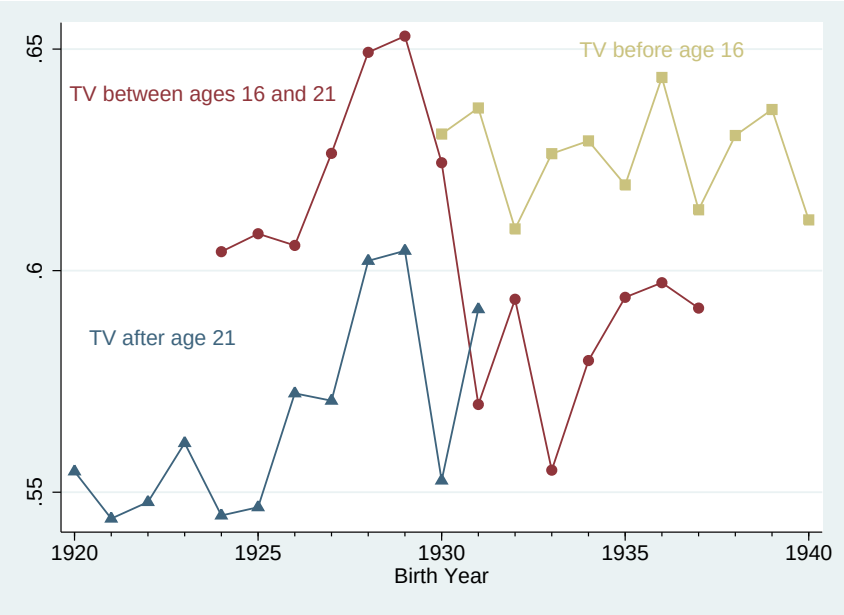


Figure 1.6. SHARE SMOKERS BY THE AGE THEY FIRST HAD ACCESS TO TELEVISION



Note: Share of people answering "yes" to the 1965-66 NHIS question "Have you smoked 100 cigarettes in your entire life?" broken out by birth year and the age people first received television.

Note at this point that whereas the natural experiment described in Section 1.3 varied the timing of television's introduction across the United States, it did not vary the order in which it was received across DMAs. In general, the larger, wealthier DMAs received television before the smaller, poorer DMAs, and this fact introduces a potential confound to the analysis presented in Figure 1.6. For example, if the larger DMAs happened to experience increases in teen smoking before smaller DMAs, we would misattribute these differential trends to television. Gentzkow (2006) argues the main determinants of television's entry order are the DMA's population and wealth, and shows that after controlling for the logs of these values, other observable DMA characteristics are not independently or jointly predictive of the television-entry order. In the next step of the analysis, I introduce a model that will allow me to control for such differential trends, in addition to cohort and DMA effects, so that the effect of television is only identified by the randomness in the timing of television's introduction.

1.5.2 A Model for Estimating Responses by Age Group

Although the simple averages in Figure 1.6 provide a useful look at the data, those results may be confounded by ordering effects discussed in section 1.5.1. The following model allows for a more robust analysis that ensures only the exogenous variation in the timing of television's introduction, and not the order of television entry across cities, estimates the effect of television. This model generalizes a difference-in-differences analysis by including an additional term to control for differential time trends across different types of DMAs, which should also control for any confounding effect of television's entry order. Consider

$$y_{cd} = \rho_d + \tau_c + \phi_c W_d + \alpha \mathbf{p}_{cd} + \varepsilon_{cd}, \tag{1.1}$$

where y_{cd} is the share of smokers in birth cohort c living in DMA d, ρ_d is a DMA fixed effect, τ_c is a birth cohort fixed effect, and ε_{cd} is an error term. Based on the terms introduced so far, this specification is similar to a standard difference-in-differences analysis: I have controlled for location effects with ρ_d, and instead of time effects, I control for birth cohorts, τ_c. However, controlling for differential trends that might be correlated with the TV-entry order requires a more general specification. I attain such a specification by including $\phi_c W_d$, where W_d is a vector containing the log of the population and log of median income of DMA d, and ϕ_c is a vector of cohort-specific coefficients. Now if a certain birth cohort happened to have higher smoking rates in larger DMAs, this trend will not be misattributed to television.

Ultimately, I am interested in the response people had to television at different ages. The vector $\mathbf{p}_{cd}$ provides a count of the number of years cohort c from DMA d was exposed to television in each of the age groups: 9-12, 13-15, 16-18, 19-21, and 22-25. Hence, the values of α provide estimates of the effect of one year of television exposure on the probability of becoming a smoker for each of these age groups. Note that because everyone

in the 1965-66 surveys has access to television, I can only estimate the relative response to television across age groups, not the level response; no one in the survey can serve as the control group that has never been exposed to television. However, some ages in my analysis may have had close to zero response to television and can serve as a reference group. Given the way $\mathbf{p}_{cd}$ is defined, the reference group consists of those who received television before age 9 or after age 25. As a result, assuming the effect of television on the smoking behavior of ages 0-8, and 26+ is zero, each element of α provides an estimate of television's impact on the corresponding age group. Alternatively, if this assumption is false and the people in the reference group experienced relatively large responses to television, my estimates understate the true impact of television on ages 9-25.

I have presented the model in equation (1.1) because it shows the observational groups for which I have identifying variation. However, NHIS allows the flexibility and precision of running regressions at the individual level. Rewriting equation (1.1) for each individual, we have

$$y_{icd} = \rho_{gd} + \tau_{gc} + \phi_{gc} W_d + x_g + \alpha \mathbf{p}_{cd} + \varepsilon_{icd}, \tag{1.2}$$

where y_{icd} is now the smoking status of person i from birth cohort c living in DMA d. Each of the controls from equation (1.1) now contains gender-specific coefficients indicated by g, as does x_g, an indicator for whether the survey response was self-reported or by proxy. These gender-specific controls are motivated by Figure 1.4, which shows very different smoking trends for men and women. When estimating this individual-level model, I cluster the standard errors by DMA (Moulton [1990]) to prevent aggregation bias.

1.5.3 Model-Based Estimates of the Response by Age Group

Estimates of the age-specific TV response, α, for different versions of equation (1.2) are presented in Table 1.1. Consistent with earlier discussion of Figure 1.6, the regression estimates in Table 1.1 suggest a greater response to television from those in their late teens than other ages; the estimates are significant for ages 16-18 across all specifications. In addition, estimates for the 19-21 group increase when I control for heterogeneous birth-year effects across different types of DMAs, a set of controls that was not available in our analysis of Figure 1.6. The estimates do not change much when including DMA fixed effects in the model.

The last two columns of Table 1.1 look at each gender separately. Here I find no evidence of significant differences across the genders. Each appears to be similarly responsive to television in their late teens, but not responsive at other ages.

The magnitude of the estimates in Table 1.1 suggest that the response to television access was large. Estimates from the third column, for example, suggest each year of exposure during ages 16-21 increased the chance of becoming a smoker by around 1.7%, meaning a person with access to television during ages 16-21 was around 10% more likely to become a smoker than someone with no television access.[6]

A graphical representation of these results can be found in Figure 1.7, which plots the estimates from the third column of Table 1.1. Comparing this figure with the histogram of starting ages in Figure 1.5 shows the ages at which people respond most to television do correspond closely with the ages at which people typically start smoking, which is consistent with our original hypothesis.

---

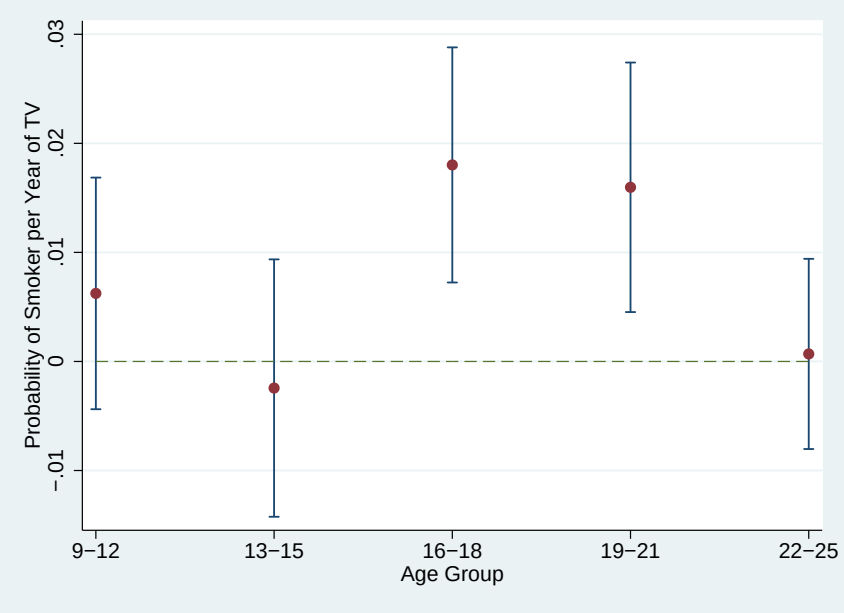
6. This calculation is formalized in section 1.7.

Table 1.1: EFFECT OF TELEVISION ON SMOKING STATUS BY AGE GROUP

	(1)	**(2)**	**(3)**	**(4)**	**(5)**
Age Group: α					
9-12	**.008**	.004	.006	.001	.010
	(.004)	(.004)	(.005)	(.007)	(.006)
13-15	-.008	-.003	-.002	-.001	-.003
	(.005)	(.006)	(.006)	(.009)	(.009)
16-18	**.015**	**.016**	**.018**	**.020**	**.017**
	(.004)	**(.005)**	**(.005)**	**(.009)**	**(.007)**
19-21	.007	**.015**	**.016**	**.027**	.011
	(.006)	**(.005)**	**(.006)**	**(.010)**	(.006)
22-25	.006	-.001	.001	.007	-.001
	(.003)	(.004)	(.004)	(.010)	(.005)
Self-Response Indicator: X_i	X	X	X	X	X
Birth Year x Gender Fixed Effect: τ_{gc}	X	X	X	X	X
DMA Characteristics x Gender x Year: $\phi_{gc}W_d$		X	X	X	X
DMA x Gender Fixed Effects: ρ_{gd}			X	X	X
Gender	Pooled	Pooled	Pooled	Male	Female
Number of observations	83,548	83,284	83,284	30,295	52,989
Number of DMAs	114	113	113	113	113

Note: This table presents estimates of the effect of one additional year of TV exposure on the probability of being a smoker, broken out by age groups, i.e., α from equation (1.2). The regression controls for birth-year fixed effects plus the birth-year fixed effects interacted with DMA characteristics (log of DMA population and median income). Additional controls are DMA fixed effects and respondent type (self or proxy). Each set of controls differs by gender, but the television response is pooled across genders. Estimates are relative to the reference population (people who were exposed to television before age 9 or after age 25) and are obtained from responses to the 1965-66 NHIS question "Have you smoked 100 cigarettes in your entire life?" The regressions use sample weights provided by the NHIS. Standard errors are clustered at the DMA level and reported in parentheses. Coefficients with t-statistics of two or higher are in boldface.

Figure 1.7. AGES THAT RESPOND TO TELEVISION EXPOSURE



Note: This figure presents estimates of the effect of one additional year of TV exposure on the probability of being a smoker, broken out by age groups, i.e., α from equation (1.2). The regression controls for birth-year fixed effects plus the birth-year fixed effects interacted with DMA characteristics (log of DMA population and median income). Additional controls are DMA fixed effects and respondent type (self or proxy). Each set of controls differ by gender, but the television response is pooled across genders. Estimates are relative to the reference population (people who were exposed to television before age 9 or after age 25) and are obtained from responses to the 1965-66 NHIS question "Have you smoked 100 cigarettes in your entire life?" answered by ˜83,000 respondents across 113 DMAs. The regression uses sample weights provided by the NHIS. Standard errors are clustered at the DMA level and the ranges shown are $\pm 2 * SE$.

1.6 SMOKING INITIATION WHEN TELEVISION ENTERS

Another way to evaluate television's impact on smoking is to look at the rate at which non-smokers became smokers, and test whether this rate increased significantly upon television entry. These rates can be estimated using a hazard model, which at the individual-level estimates the risk (or "hazard") an individual will become a smoker at different points in his life. As in the analysis above, one can compare individuals of the same birth cohort who either did or did not have television at a given age to estimate whether television affected their propensity to start smoking. This analysis offers insight into the timing of television's influence.

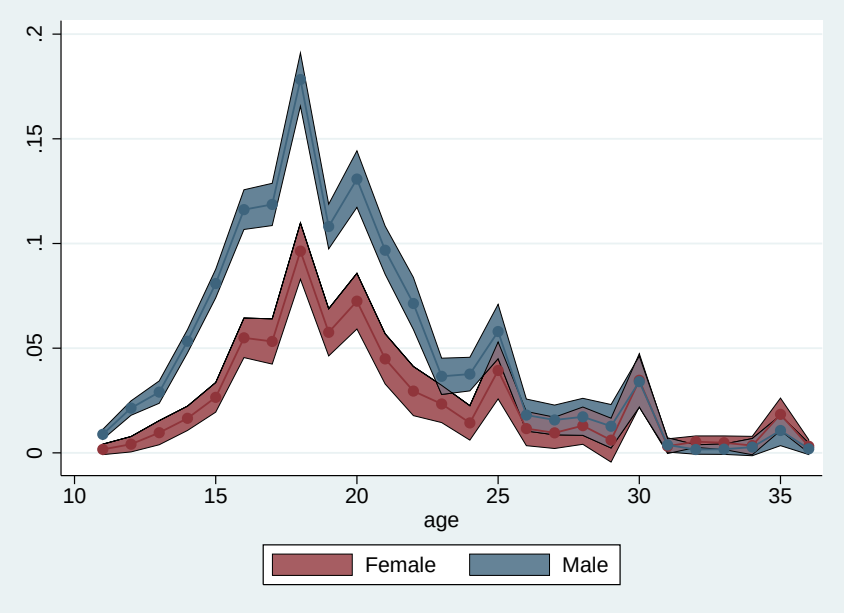
In section 1.6.1, I begin by explaining how I estimate the hazard of smoking at each age and then inspect the residuals from those estimates to look for model-free evidence of television's impact on smoking in section 1.6.2. Section 1.6.3 introduces a model for estimating how television changed the hazard of becoming a smoker, and I discuss estimates from this model in section 1.6.4.

1.6.1 Hazard Estimates for Each Age

The NHIS data on starting ages lend themselves to thinking about the risk of becoming a smoker over discrete periods of time: people's ages. For this analysis, I assume everyone is born a non-smoker but is at risk of becoming a smoker either until they are converted, or their life is over. Once a person becomes a smoker, he can never return to being a non-smoker. In this setting, the hazard is the probability that a person who is not yet a smoker becomes a smoker over the next year of his life.

The hazard of smoking at each age can be estimated by simply calculating the share of all people in the sample who started smoking at each age. In anticipation of later analysis, I perform this calculation by constructing a panel with an observation for each age over each

Figure 1.8. HAZARD OF SMOKING UPTAKE BY AGE



Note: Coefficients are the average hazard of becoming a smoker for ages 11-36 and are estimated from responses to the 1970 NHIS question "At what age did you start smoking regularly?" Standard errors are adjusted for the NHIS survey design and the ranges shown are $\pm 2 * SE$.

respondent's life up until the age he either becomes a smoker, or the last age he completed without becoming a smoker. Then I estimate

$$y_{it} = \gamma_{gt} + \varepsilon_{it}, \tag{1.3}$$

where y_{it} is an indicator that person i reported becoming a smoker at age t, and ε_{it} is an error term. γ_{gt} is a fixed effect for each gender-age, which provides the desired estimates of the hazard associated with each gender at each age. These estimates are presented in Figure 1.8, and the distribution is the same as the histogram of starting ages in Figure 1.5, but contains additional information on the level of risk associated with each age. For this analysis, the age distribution has been truncated to 10-36, a range that accounts for 97% of all starting ages.

1.6.2 Model-Free Evidence of Television's Influence: Hazard Time Trends

Next I inspect the residuals from the regression in (1.3) for evidence of television's influence on smoking.[7] If television was influential on smoking uptake, we might expect that individuals who received television early experienced shocks to their hazards that late-TV receivers did not experience until later. To test this theory, I exploit the three distinct waves of television entry that are apparent in Figure 1.3: "early" DMAs that received it in 1946, "middle" DMAs that received it from 1947 to 1952, and "late" DMAs that received it after the FCC ban was lifted in 1953. To observe whether these groups experienced different shocks, I compute the mean of the residuals from equation (1.3) for each of TV-entry group and calendar year.[8] I can now inspect the trends in these means for evidence of TV-related shocks.

The results from this exercise are presented in Figure 1.9, which plots two lines: the trend in the hazard for early- and middle-television receivers, each relative to the trend for late receivers.[9] These plots allow us to see that people did experience increases in the hazard of smoking that correspond to when they received television. From 1946 to 1947, the trend for early DMAs moved upward, unlike the other groups. Then, from about 1950 to 1952, the trend for middle DMAs moved upward, also unlike the other groups. Finally, in 1954 – one year after most of the late DMAs received television – both trends move

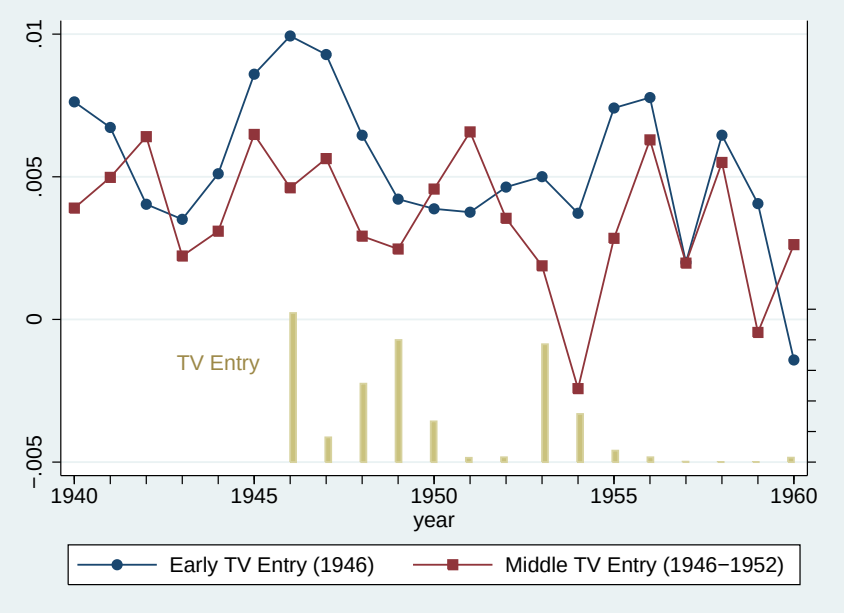
---

7. Because we saw in Figure 1.8 that age is highly predictive of smoking uptake, we remove the effect of age using the regression in (1.3) so that we may observe the influence of television with greater precision.

8. The residuals are $\hat{\varepsilon}_{it} = y_{it} - \hat{\gamma}_{gt}$, as estimated from equation (1.3). Each age, t, of a respondent's life is mapped to the calendar year in which he spent the majority of that age, τ. Each person, i, is mapped to his DMA's TV-entry group, $m \in \{\text{early}, \text{middle}, \text{late}\}$. I compute $\bar{\hat{\varepsilon}}_{m\tau} = \frac{1}{N_{m\tau}} \sum_{i \in m, t \in \tau} \hat{\varepsilon}_{it}$.

9. That is, Figure 1.9 plots $\left[\bar{\hat{\varepsilon}}_{(m=\text{early})\tau} - \bar{\hat{\varepsilon}}_{(m=\text{late})\tau}\right]$ and $\left[\bar{\hat{\varepsilon}}_{(m=\text{middle})\tau} - \bar{\hat{\varepsilon}}_{(m=\text{late})\tau}\right]$. $\bar{\hat{\varepsilon}}_{m\tau}$ is defined in footnote 8.

Figure 1.9. HAZARD OF SMOKING UPTAKE BY DMA TYPE



Note: The plot reports how the hazard in DMAs that received television before 1953 compare to DMAs that received television after 1953. The reported values were obtained as follows. First, I computed the residuals from a regression of indicators for the age that individuals reported starting smoking on age dummies for each gender, i.e., the residuals from the regression that produced the estimates reported in Figure 1.8. The average of the residuals was computed for each calendar year and DMA type, where early DMAs received television in 1946, middle DMAs received it 1947-1952, and late DMAs from 1953 onward. The value reported for each year is the average for the indicated DMA group minus the average for the late DMAs. The data come from responses to the 1970 NHIS question "At what age did you start smoking regularly?" Along the bottom (labeled "TV Entry") is a histogram of TV entry dates, weighted by population.

downward, suggesting the hazard of becoming a smoker had increased in the late DMAs. Thus, evidence shows that each of the three TV-entry groups experienced a shock to its hazard rate that corresponds to when it received television, offering consistent evidence of television's impact on smoking uptake.[10]

10. The reader may also notice that each of these shocks appears to be temporary. See section 1.6.4 for a discussion of the dynamics, which are consistent this observation.

1.6.3 A Model of Television's Impact on the Hazard

Having observed graphical evidence of television's effect on the hazard, the following specification allows for the estimation of the magnitude of television's effect, while incorporating the precise timing of TV entry to each DMA and using the variation from the natural experiment for identification:

$$y_{it} = \gamma_{gt} + \theta_{g\tau} + \phi_{g\tau} W_d + \beta \mathbf{TV} + \varepsilon_{it}. \tag{1.4}$$

y_{it} is again an indicator for whether person i became a smoker at age t. As in equation (1.2), I allow all the controls in this expression to have gender-specific coefficients indicated by g subscripts to account for the very different smoking trends between men and women shown in Figure 1.4. γ_{gt} are fixed effects for each gender g at each age t. To control for time trends in smoking behavior, I include fixed effects, $\theta_{g\tau}$, for each calendar-year τ that are also allowed to differ by gender, g. $\phi_{g\tau} W_d$ controls for differing time trends across DMAs that are larger or wealthier: W_d is the log of the population and log of median income of DMA d, which i lives in, and $\phi_{g\tau}$ is a set of calendar-year and gender-specific coefficients. Finally, **TV** is a vector of dummies corresponding to each of K years before and K years after person i's DMA received television. Specifically, let t^* be the age at which person i received television; then $TV_k = \mathbf{1}\{k = t - t^*\}$.

I am primarily interested in the coefficients on **TV**, β. Coefficients for years prior to TV's entry provide a pre-trend test: TV should have no influence before it enters, and non-zero estimates for these coefficients would suggest some important unobservables are not included in the estimation. Coefficients for the entry year and later provide estimates of how TV affected the hazard of smoking uptake following its arrival.

1.6.4 Model-Based Estimates of Television's Impact on the Hazard

Estimates of how television influenced the risk of becoming a smoker, β, are presented in Table 1.2 for different versions of the model in equation (1.4). The first column, which employs the simplest version of the model, shows a sudden rise in the estimated coefficients at time zero that persists for the next four years. The second column shows that including calendar year fixed effects reduces the number of significant coefficients. However, the third column shows that adding controls for heterogenous calendar year effects to the model gives the strongest coefficient estimates. This observation suggests the heterogenous time trends mask rather than amplify the true effect of television. None of the specifications presented in Table 1.2 recover significant coefficients for years prior to TV entry, which means my estimates do not fail the falsification test – I estimate no response to television until television actually enters.

Each gender is estimated separately in the fourth and fifth columns of Table 1.2, and show significant responses for women, but not for men; the standard errors are about three times larger for men than women. Although this difference in precision across genders presents something of a puzzle, a couple of explanations are possible. First, note the data contain about three times the number of female respondents as male respondents. This difference in sample size follows from a couple of steps taken when cleaning the data: I excluded survey responses given by proxy, because of the difficulty of reporting the precise age a family member started smoking, and I excluded war veterans (see section 1.4). Each of these steps disproportionately reduces the number of men in our samples: proxy responses because in 1970 more women than men were at home to respond to the survey directly, and war veterans because of the selection of men for military service. A second potential explanation for the difference in precision comes from the psychology literature, which suggests possible differences in measurement error across the genders that also favor

Table 1.2: HAZARD OF SMOKING UPTAKE WHEN TELEVISION ENTERS: LINEAR MODEL

	(1)	**(2)**	**(3)**	**(4)**	**(5)**
Year Relative to TV Entry: β					
-5	.002	.002	.001	-.004	.002
	(.002)	(.002)	(.002)	(.009)	(.003)
-4	.002	.002	.003	-.008	.004
	(.002)	(.002)	(.003)	(.009)	(.003)
-3	.001	-.001	.001	.006	.001
	(.002)	(.002)	(.003)	(.009)	(.003)
-2	.003	.001	.003	.003	.003
	(.002)	(.003)	(.003)	(.010)	(.003)
-1	.001	-.002	.001	.003	.002
	(.002)	(.002)	(.002)	(.009)	(.003)
0	**.008**	**.006**	**.010**	.008	**.010**
	(.002)	**(.003)**	**(.003)**	(.011)	**(.003)**
1	**.009**	**.008**	**.010**	.004	**.012**
	(.002)	**(.003)**	**(.003)**	(.010)	**(.003)**
2	**.005**	.003	**.007**	.008	**.007**
	(.002)	(.003)	**(.003)**	(.009)	**(.003)**
3	**.006**	.004	**.007**	-.007	**.009**
	(.003)	(.003)	**(.003)**	(.009)	**(.003)**
4	**.005**	.004	**.007**	-.010	**.009**
	(.002)	(.002)	**(.003)**	(.011)	**(.003)**
5	.004	.002	.005	-.019	**.008**
	(.003)	(.003)	(.003)	(.010)	**(.003)**
Age x Gender Fixed Effects Baseline: γ_{gt}	X	X	X	X	X
Calendar-Year x Gender Fixed Effects: τ_{gj}		X	X	X	X
DMA Characteristics x Gender x Year: $\phi_{gj}W_d$			X	X	X
Gender	Pooled	Pooled	Pooled	Male	Female
Number of Observations	345,968	345,968	343,068	59,638	283,430
Number of Individuals	22,468	22,468	22,468	5,193	17,275
Number of DMAs	114	114	112	112	112

Note: Estimates of the contribution to the hazard of starting smoking in the years relative to television entry. The specification is a discrete-time hazard model in age that includes age fixed effects, calendar-year fixed effects, and interactions of DMA characteristics (log population and log median income) with the year fixed effects as show in equation (1.4). Estimates are from responses to the 1970 NHIS question "At what age did you start smoking regularly?". Regressions use sample weights provided by the NHIS. Standard errors are clustered at the DMA level and reported in parentheses. Coefficients with t-statistics of two or higher are in boldface.

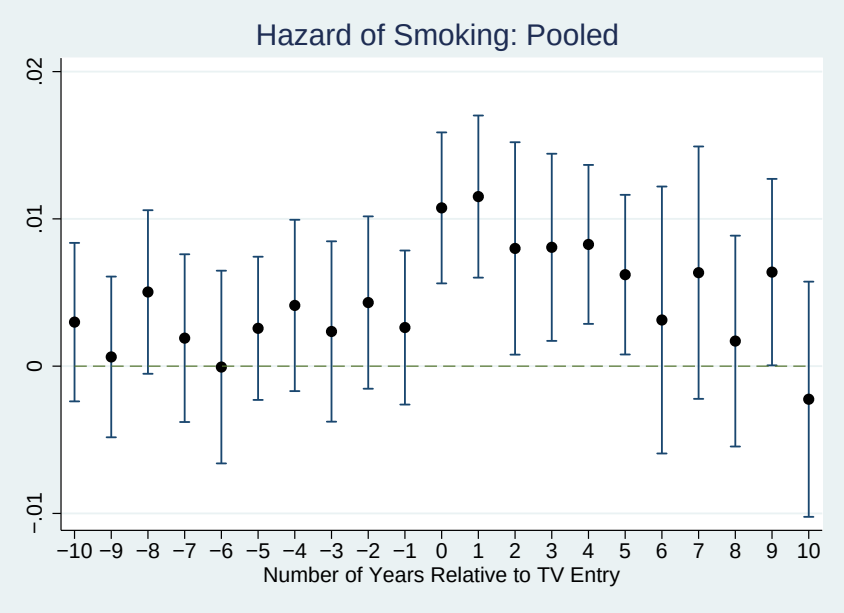
women. Herlitz et al. (1997) find women possess superior episodic memory (recall of autobiographical experiences in a particular place and time), findings that are also suggested by a number of earlier studies they review. Even if these two explanations do not fully resolve this precision puzzle, these estimates do not contradict the findings in section 1.5. There I estimated responses to television for each gender, which were indistinguishable from one another (see Table 1.1), and here the confidence intervals for male responses are large enough that, once again, we cannot distinguish between the estimates for each gender.

A graphical representation of our estimates is presented in Figure 1.10, which shows clearly the sudden jump in the hazard that occurred during the two years after television was introduced. Over these two years, the hazard is about 0.01 higher than prior to entry, which corresponds to an average 20% increase.[11] The speed of the response to television may be surprising given that broadcast availability did not mean all homes suddenly had television. However, according to surveys at the time, people who did not own televisions were frequently exposed to the new medium. A 1949 survey in metropolitan New York conducted by the National Broadcasting Company (NBC) reports that 63% of non-television families were exposed to television at least occasionally, with 41% seeing it once a week or more, typically at the homes of friends or family (Beville 1949). Similarly, a nation-wide survey of women in 1954 reports that 47% of women in non-television homes had watched television in the last month (Simmons & Associates Research, Inc., 1954). The quick response to television may also be explained by network effects: if a few people were converted as result of television exposure, others follow not because they were exposed to television, but because others in their community had started smoking.

Figure 1.10 also shows that after the initial jump in hazard rates after television entry, the rates ease downward toward zero, for example, in years 2-5. However, this evidence

11. Figure 1.8 shows the average baseline hazard is around 0.05.

Figure 1.10. THE HAZARD OF BECOMING A SMOKER WHEN TELEVISION ENTERS



Note: Coefficients are β in equation (1.4), which estimate the contribution to the hazard of starting smoking in the years relative to television entry. The specification is a discrete-time hazard model in age that includes age fixed effects, calendar-year fixed effects, and interactions of DMA characteristics (log population and log median income) with the year fixed effects. The estimates were obtained from responses to the 1970 NHIS question "At what age did you start smoking regularly?", consisting of ˜22,000 respondents across 112 DMAs. The regression uses sample weights provided by the NHIS. Standard errors are clustered at the DMA level and the ranges shown are $\pm 2 * SE$.

does not necessarily mean television became less effective in the years after it first entered. To understand why, remember that (1) the hazard estimates the *rate* of smoker conversion in the population, and (2) the stock of people who might be responsive to television but have not yet been exposed is never higher than when television first enters. To illustrate these dynamics with a simple example, imagine only a portion of 16-21-year-olds respond to television and everyone else is completely unaffected by it. When TV first enters a DMA, a large stock of 16-21-year-olds exist who have never seen television but are inclined to start smoking once they see it. As a result, when TV enters, the rate of smoking uptake spikes upward as this groups suddenly starts smoking. Then, with this TV-responsive population converted, the rate will decrease, but not to its initial level: new cohorts are continuously turning 16, and under our assumption, some of them are becoming responsive to television. As a result, the long-run hazard rate is higher than it was before television, but lower than it was when television was first introduced. These dynamics, which are consistent with the evidence in Figure 1.10, in play out even when all 16-21-year-olds, across all cohorts, are equally affected television.

The estimates in this section were produced using a linear model, but the hazard literature often uses a proportional hazards model. Such a model and its estimates are presented in Appendix A.1 as a robustness check. This approach produces similar but slightly less precise estimates.

1.7 The Share of Smokers Generated by Television

The previous two sections, using different data and types of analysis, presented evidence that television caused people to start smoking. Each of these analyses have implications for how television changed the share of smokers in the population. This section investigates these implications.

First, consider the analysis in section 1.5, where I estimated the impact of one year of television access on the likelihood of becoming a smoker for different age groups. Using these estimates, I can approximate the difference between the share of smokers in populations that always had access to television and the share of smokers in populations that never had access to television by summing across the televisions effects estimated for each age group. That is,

$$\Delta S_{TV} = \sum_j n_j \hat{\alpha}_j,$$

where ΔS_{TV} is the change in the share of smokers as a result of television, $\hat{\alpha}_j$ is the coefficient estimated for age group j, from (1.2), and n_j is the number of years covered by age group j.

I can produce the same estimate using the hazard analysis from section 1.6, but doing so requires more assumptions. Now the hazard estimates must be converted to cumulative probabilities. Additionally, I assume the entire increase in the hazard following TV entry is the result of people initiating smoking who would not have otherwise done so. This assumption could artificially raise the predicted increase in the share of smokers if a significant fraction of people who were going to smoke anyway started sooner because of television. On the other hand, the data on smoking-initiation ages are likely to contain significant reporting errors from respondents, given that they are reporting on an event that occurred around two decades in the past; these reporting errors attenuate the estimates. Although I can only speculate as to how the magnitude of these opposing effects net out, the calculation is still useful for comparison purposes. Appendix A.2 shows the change in the share of smokers from television can be estimated as

$$\Delta S_{TV} = exp\{-H_0\} - exp\left\{-H_0 - \sum_{j=0}^{4} \beta_j\right\},$$

where ΔS_{TV} is again the estimated change in the share of smokers as a result of television, H_0 is the cumulative hazard of becoming a smoker in the absence of television, and β_j is the estimated change in the hazard associated with the j-th year after television's introduction (1.4) for TV-year j.

Taking these expressions to the data and estimating confidence intervals using the delta method shows that the age-group analysis implies $\Delta S_{TV} = 0.122 \pm 0.089$, and the hazard analysis implies $\Delta S_{TV} = 0.030 \pm 0.012$, 95% confidence ranges with a small amount of overlap. As discussed above, estimates from the hazard model required strong assumptions, which may explain the difference in these estimates. The estimates derived from the age-group analysis are more likely to be accurate given they offer more direct measurement of the change in the share of smokers in the population.

1.8 DISCUSSION

The media environment has undoubtedly changed since the 1950s, and data sources have become increasingly rich as reflected in many of the studies presented in section 2.2. These and other studies are able to look at different types of media separately (e.g., advertising, television, movies, and other promotional activity) in ways that are simply not possible using data from the mid-20th century. On the other hand, the ubiquity of modern media has made establishing a causal link between smoking and media difficult, due to selection concerns, and it is here that the historical events presented in this paper make their contribution.

The past work discussed in section 2.2 reveals a strong association between media exposure and smoking, but these studies critically lack random assignment of media exposure. The micro-based studies must assume that after controlling for available demographic characteristics, the remaining self-selected level of media consumption is orthogonal to smoking behavior. If this assumption is incorrect, their estimates contain a selection bias, the size of which is unknowable, but is possibly responsible for all of the correlation they report. In other words, the existing literature leaves open the possibility that media exposure does not cause people to smoke, but those who choose more media consumption also tend to smoke. By exploiting the quasi-random assignment of television access, this study produces estimates without this limitation, and in this way provides more robust evidence that media exposure caused teens to start smoking than was previously available.

Although the natural experiment I examine does not reveal which features of television are responsible for the estimated response, a number of factors suggest advertising was responsible for at least some of it. Section 1.3 provided evidence that the cigarette industry quickly adopted television as its primary advertising medium, and consumers were quite aware of the cigarette brands promoted on television. Indeed, firms had strong incentives to support any aspect of television that encouraged smoking. Although unsponsored features of television might also have encouraged smoking – for example, seeing famous people smoke on television – economic theory suggests firms would have found ways to promote such features, whether through spot ads, product placement, or by other means. Under the assumption that advertising was responsible for some of the effect we measure, this paper provides new support for the effectiveness of tobacco advertising bans. Some readers may find this assumption more palatable than the assumption required by the existing literature that associations between smoking and advertising reveal a causal effect.

The significant age heterogeneity in media responses found in this paper also has policy implications. Most immediately, it helps justify prevention programs that are focused on

youths' exposure to media. More generally, the correspondence between "high-risk" ages and "high-response" ages provides support for targeting at-risk groups that may be particularly influenced by external factors such as the media. This offers new support for the U.S. Surgeon General's focus on tobacco use by young people (U.S. Department of Health and Human Services [1994]).

The age heterogeneity also has implications for how mass media affects the population in the short and long term. Section 1.7 provided an estimate that each fully-exposed cohort had a predicted 12% more smokers than it would have had without television. The short-run impact on the whole population is likely to be small, however, because each cohort represents a small portion of the total population. For example, in the span of five years, these estimates suggests television would change the share of smokers by about 1% (assuming 7% of the population had turned over through births and deaths) – an effect size likely to be undetectable by many studies that use aggregate measures of consumption. These dynamics helps explain why existing studies of aggregate data, primarily in the economics literature, have had difficulty detecting the impact of media on smoking, and suggests these estimates may dramatically understate the true effect of the media.

The evidence for television's influence on smoking does not hinge on a single data source, which helps demonstrate robustness. Using a separate survey question to measure the response to television, our analysis in section 1.6 demonstrates the sudden bump in smoking initiation that accompanied television's entry to a city. That this analysis detects any effect at all may be surprising given that (1) these estimates are likely to be biased towards zero due to errors from respondents reporting on an event that occurred decades in the past, and (2) similar analysis of this NHIS question reveals no response to prices (Douglas and Hariharan [1994]; Douglas [1998]).

Indeed, this study's estimated response to television is large relative to that typically found for prices. Bader et al. (2011) review the literature estimating how prices influence

smoking initiation, and find that of 22 studies, nine found no effect of prices, six found effects in some cases, and seven found significant effects. Some of the largest estimates in this literature come from Sen and Wirjanto (2011), who estimate a price elasticity of smoking initiation of -0.2 to -0.5. As a result, our respective point estimates suggest a 34% drop in prices would be required to produce the same increase in smoking initiation as television. Note that this estimate may dramatically underestimate the influence of television relative to prices because (1) this study's television-response estimates are known to be attenuated by migration, and (2) I purposely selected the largest estimates from the price literature. Because prices are known influence a wide variety of decisions, I believe this comparison helps to underscore the importance of mass media on consumer behavior.

1.9 CONCLUSIONS

This study makes use of a natural experiment to investigate the relationship between media and smoking while avoiding the selection concerns that trouble existing investigations. This study provides consistent evidence that television did affected smoking behavior, particularly on 16-21-year-olds, an age group that is already of high risk smoking initiation. By offering causal evidence of television's impact on smoking initiation, this study offers robust support for the view that mass media can influence the behavior of individuals in ways that are detrimental their health. Comparisons of this study's estimates to those in the literature suggests television was more influential than prices on the decision to initiate smoking.

CHAPTER 2

MASS ADVERTISING AS A NATURAL EXPERIMENT

2.1 INTRODUCTION

Estimating the response to advertising can be difficult because firms strategically adjust their advertising levels in ways that may not be known to the researcher. To address this difficulty, recent work by Shapiro (2016) and Sinkinson and Starc (2015) offer methods of achieving unbiased estimates of the response to television advertising, by isolating differences in local advertising levels that are thought to be determined in a quasi-random manner. Estimates presented in these papers underscore the potential for biased advertising estimates if the researcher is not careful to isolate quasi-random variation for the purpose of identification. However, because these methods rely on local differences in advertising levels, they are rendered ineffective for product categories that rely heavily on national advertising.

To address this limitation, I propose a novel method that is more effective for products that rely on mass advertising such as national television. This identification strategy derives from two simple observations. First, any given advertising intensity is not perfectly optimal for all populations that view it, especially when advertising decisions are made for very large populations. This occurs frequently because firm are often attracted to the price discounts associated with mass advertising. Second, the current data proliferation makes measuring the degree to which some populations receive sub-optimal advertising levels increasingly feasible. These new data may alter firms' behavior to some degree, but as long as the price discount is large enough, they will continue to purchase mass-advertising despite the knowledge that it is sub-optimal for some populations. Together, these observations suggest a natural experiment: the advertising any given population subgroup experiences

is in part determined by the optimal level forecasted for other populations. These natural experiments present a new opportunity to identify the response to mass advertising, without requiring firms to alter their behavior and without relying on the idiosyncrasies of a particular technology. In this paper, I will provide a formal presentation of this "coarseness" strategy and provide examples of its use for a variety of product categories, using a variety of data sources.

For an intuitive illustration of this strategy, consider advertising for antihistamines. Antihistamines treat a variety of allergies; chief among them are pollen allergies, which are highly seasonal. These "organic" changes in demand are well known to firms that target their advertising to match the high-demand spring and fall allergy seasons with national advertising. As a side effect of this behavior, many cities receive more or less advertising than is optimal for them. For example, Chicagoans may see a lot of advertisements for Claritin in mid-March while they are still huddled inside from the snow, and completely unaffected by spring allergies. Meanwhile, sneezy Atlantans see the same amount of advertising for antihistamines, which, for them, is less than optimal given their need for allergy relief. In each of these examples, the expected pollen levels in other places determined the amount of advertising received locally. If expected pollen levels in other places only affect local sales through national advertising, we can obtain unbiased estimates of the advertising elasticity using instrumental variables (IVs).

Although demand for most products is not driven by particles in the air, the concept of "organic demand" – demand driven by factors outside the control of the firm – applies to many other product categories. Frequently, firms adjust their advertising levels in anticipation of some event, for example, holidays, weather forecasts, sporting events, or special-interest festivals, that will alter the level of the organic demand for their product. The coarseness strategy presented in this paper applies to a product if (1) forecasted changes in organic demand drive at least some of the changes in advertising levels over time, (2)

the changes in organic demand are not uniform across all observable units (e.g., markets or individuals), and (3) the researcher can measure at least some of the local changes in organic demand that influenced the mass advertising decision. Opportunities for such conditions being met are improving as large data sets with observations at small levels of aggregation are increasingly available. This paper demonstrates its use for the antihistamine, lip medication, moisturizers, and sunscreen categories because these categories offer non-proprietary measures of organic demand.

Similar to Shapiro (2016) and Sinkinson and Starc (2015), the "coarseness" strategy presented in this paper suggests the use of quasi-random identifying variation can significantly influence advertising estimates relative to other methods. For example, failing to account for the role of forecasted pollen levels in antihistamine advertising decisions can bias estimates of advertising elasticity upwards by a factor of two. Additionally, the evidence in this paper both supports and contradicts Rossi's (2014) critique of the use of IV estimators in marketing applications. In that paper, Rossi rightly points out that IV estimates are less efficient than correctly specified OLS estimates and that using invalid instruments for IV can produce highly biased estimates. Each of these phenomena can be observed in this paper's results. However, he also argues that marketing environments offer enough supporting data that IV estimators typically fail to be useful. The categories considered in this paper were selected based on the availability of non-proprietary measures of organic demand, and yet simply adding additional control(s) for organic demand, as Rossi recommends, produces negative advertising elasticity estimates for two of the four categories: lip medications and sunscreens. These estimates may be biased as a result of additional unobservables, which may not be knowable or measurable. Alternatively, the coarseness strategy, like other IV approaches, does not require the econometrician to have measures of all confounding variables to obtain consistent estimates. IV approaches allow the econometrician to specify exactly which part of the variation in advertising is quasi-

random, and remain agnostic about the rest. The possible utility of the coarseness strategy suggests IV estimators may have a place in marketing applications, especially as data sets continue to grow and mitigate the problems with applying IV to small samples.

That said, the coarseness strategy is not a panacea and will not work properly without a suitable measure of organic demand, and that measure of organic demand cannot contain responses to advertising. For example, Google Trends and other crowd-sourced data sets may at first appear to offer useful measures of organic demand. However, comparisons in this paper with estimates derived from "true" measures of organic demand suggest Google Trends data contain too much response advertising to be useful.

The remainder of this paper is organized as follows. Section 2.2 presents a review of the literature related to the topics covered in this paper. Section 2.3 presents the data used in this paper. Section 2.4 presents a detailed explanation of how the coarseness strategy works, and presents the formal model. Section 2.5 presents applications of the method, both at the category and brand levels, and examines the problems with using Google Trends data. Section 2.6 discusses implications of the method for firms and section 2.7 concludes.

2.2 LITERATURE REVIEW

This paper relates to a number of literatures that include characterizations of advertising carryover, endogeneity bias in marketing-model estimates and the theories of optimal advertising. I discuss each of these below.

The question of exactly how past advertising carries over to a current impact on sales has been explored for some time. Bass and Clarke (1972) lay out a number of options for the distributed lag model, the most popular of which has become the Koyck model (1954), which assumes a geometric decay of advertising's impact over time. Helmer and Johansonn (1977) present the Box-Jenkins transfer function analysis, which provides a

means of estimating which functional forms best represent the true lag structure. Tellis, Chandy, and Thaivanich (2000) investigate a marketing environment that allows them to plot the effect of past advertising without parametric assumptions. Using the precise timing of television advertisements for a telephone number providing medical services, they find the call-in responses dissipate within a few hours after an advertisement has aired. In a similar quasi-experimental setting, Liaukonyte, Teixeira, and Wilbur (2015) look at cross-media responses: seeing an ad on television and making a purchase online. Similarly, they find evidence of little effect two hours after having aired, but pronounced effects in the first two minutes.

Studies employing observational data have frequently relied on the Koyck model of advertising carryover, and have estimated that advertising's effect on sales takes considerably longer to drop to 10% of its same-period impact. Palda (1964) was the first such study and estimated that advertising's effect could last several years, though his results may be biased by the use of annual data. Clarke (1976) estimates carryover for several different product categories and finds that advertising in most categories lasts three to nine months. Assmus, Farley, and Lehmann (1984) report similar carryover estimates in a metastudy. Long durations are also revealed in an experiment by Lodish et al. (1995), who randomized how much advertising consumers received for specific products, by manipulating their cable-television feeds. They found that whereas most advertising was ineffective, advertisements that were effective continued to have an impact three years after the original campaign had ended.

Concerns about the endogeneity bias in marketing-response estimates have also been raised and may be especially acute in studies of advertising. Endogeneity concerns arise because the econometrician frequently does not observe all factors that went into determining how marketing variables were set, and some factors may have affected sales even without having adjusted the marketing variable. For example, if the econometrician used

the change in sales and advertising around July 4th to estimate the effect of advertising on hot dog sales, the estimates are likely to be confounded by the increased organic demand that drove the observed increase in advertising – sales would have increased even if advertising had not changed. Villas-Boas and Winer (1999) present evidence of an endogeneity bias for estimates of price elasticities, using past prices as instruments. Elberse and Eliashberg (2003) present evidence of endogeneity bias in the advertising elasticities for movies, using a simultaneous-equations model. They find that by correcting for how films were supplied, their estimates of the advertising elasticity drops from 0.58 to 0.2. Others have proposed specific identification strategies to produce unbiased estimates of the response to advertising. Shapiro (2016) looks at the different advertising experiences on either side of a DMA boundary, Sinkinson and Starc (2015) propose using political campaigns as an instrument that affects the local cost of advertising, and Hartmann and Klapper (2014) use local preferences for watching the Superbowl. A series of articles in 2005 further discuss the concerns with endogeneity and its implications for the policy-relevance of marketing models (Franses [2005a,b]; Bronnenberg, Rossi, and Vilcassim [2005]; and Van Heerde, Dekimpe, and Putsis [2005]). Also see Besanko, Gupta, and Jain (1998), Bronnenberg and Mahajan (2001) and Manchanda, Rossi, and Chintagunta (2004).

Current theories of advertising are typically based on Nerlove and Arrow (1962), who propose advertising can be viewed as investments in an ever-depreciating stock of goodwill. A large literature in operations research has solved for optimal advertising based on different assumptions of how it is influenced by product quality, prices, S-shaped responses, and competitive forces (see Feichtinger Hartl and Sethi [1994] for a survey). The marketing literature has also investigated the nature of the advertising response curve: whether evidence for an S-shaped response exists (Vakratsas et al. [2004]), how the firm can optimally schedule given such a shape (Naik, Mantrala, and Sawyer [1998]; Feinberg [2001]; Freimer and Horsky [2012]), and justification for pulsed advertising (Dube, Hitsch, and

Manchanda [2005]). Other moderators of advertising responses have included annoyance (Teixeira, Wedel, and Pieters [2010]), the amount of time between sequential exposure (Sahni [2015]), frequency of ad exposure (Johnson, Lewis, and Reiley [2014]) and, for eBay, the frequency and recency of eBay purchases (Blake, Nosko, and Tadelis [2015]). Andrews et al. (2015) show commuters are more responsive to mobile-phone ads when they are on crowded subway trains relative to non-crowded trains. Luo (2016) finds consumers are more responsive to standard mobile-ad promotions in sunny weather, but prevention-based copy (e.g., "Don't miss the opportunity to...") is more effective in rainy weather.

Much of the literature on targeted advertising is theory based and has become increasingly motivated by the information provided by new technologies (Goldfarb [2014]). The model proposed by Ganesh, Soberman, and Villas-Boas (2005) suggests advertisers should target the groups that are most loyal to their product, and not those that have other preferences or are comparison shoppers. Anand and Shachar (2009) consider advertising as a potentially noisy signal. Li and Du (2012) propose a method of targeting advertisements to mobile phones based on preferences that are partially inferred from time and location data. A related literature has looked at how to use web-based information to customize the content of communications. Murthi and Sarakar (2003) propose a framework for personalization, whereas Ansari and Mela (2003) calibrate their model using internet-usage data to improve the design of emails such that they increase website traffic.

To my knowledge, relatively little work has investigated how a firm might adjust its advertising based on predicted future demand. The literature on managing consumer relationships has produced some related studies: Gönül and Shi (1998) and Bitran and Mondschein (1996) analyze the role of recency and frequency of purchase on responses to catalogs and propose models for how such insight should affect the firm's long-run strategy. Ching et al. (2004) consider the problem of which customers to target with promotions based on their recent activity.

2.3 DATA

The data required for this identification strategy include data on sales, advertising, and organic demand. As discussed earlier, organic demand measures factors that influence demand which are outside the control of the firm. Such data will often be available to firms, since these are factors that influence their decision to adjust advertising levels. In this paper, however, I will demonstrate the use of this methodology with data that are non-proprietary, specifically, pollen and weather data.

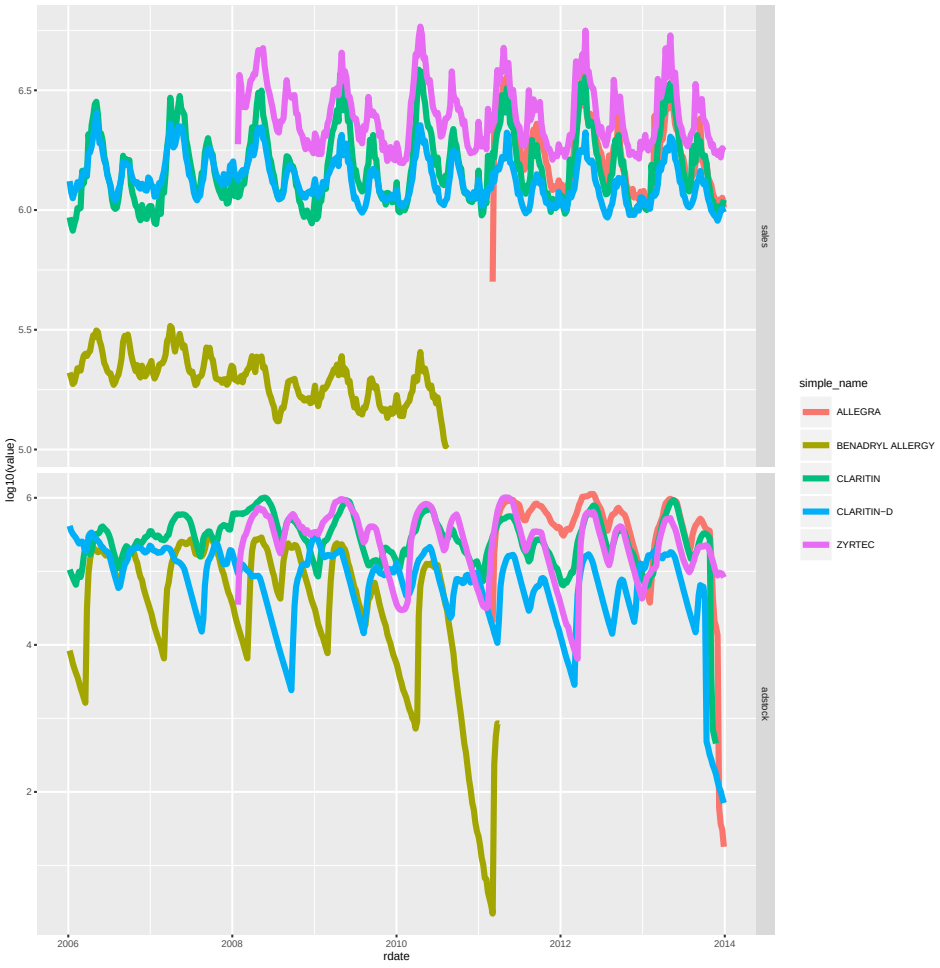
Pollen data are provided by the National Allergy Bureau (NAB), which is administered by the American Academy of Allergy Asthma & Immunology (AAAAI). These data were collected by 51 stations covering 39 DMAs across the United States from 2004 to 2013. The stations are predominantly independent allergists' offices, each of which samples pollen from the air, inspects it microscopically, and reports the number of pollen particles, typically broken out by plant genus. These contributors are listed in Appendix Table B.1, and their locations are shown in Appendix Figure B.1. Pollen data that are missing for DMA-weeks between 2006 and 2013 were imputed using a random forest algorithm. Details on this imputation can be found in Appendix B.1.1.

Weather data are used for the imputation of the pollen data, and separately as measures of organic demand for some product categories. These data are provided by the National Oceanic and Atmospheric Administration (NOAA), collected from 1,335 stations across the United States from 2003 to 2015. Weather data were interpolated for each county-day using bicubic interpolation.

Sales and advertising data come from Nielsen. Sales data are derived from the Retail Measurement Services (RMS), covering stores across the United States for 2006-2013, which are reported by UPC-store-week. Television advertising data are reported at the copy-day level. I aggregate the advertising data to total exposures by week and DMA to

match the sales data. See Appendix B.2.1 for estimates of advertising carryover and the calculation of advertising stock. Brand-level summaries for antihistamines are presented in Figure 2.1 for sales and advertising; summaries for other products can be found in Appendix Figure B.6. Product markets analyzed in this paper are defined using Nielsen's "product module" definitions, except in the case of antihistamines. The antihistamine market was defined by the group of products that are highly sensitive to local pollen levels and have sales levels of national importance. See Appendix B.1.2 for details on the antihistamine market definition.

Figure 2.1. BRAND-LEVEL SALES: ANTIHISTAMINES



Notes: Summary statistics for brands in the antihistamine market. The top panel reports $log_{10}(\text{sales})$ by brand and year, the bottom panel reports $log_{10}(\text{Advertising Stock})$ for each brand and year.

2.4 BASICS OF THE STRATEGY

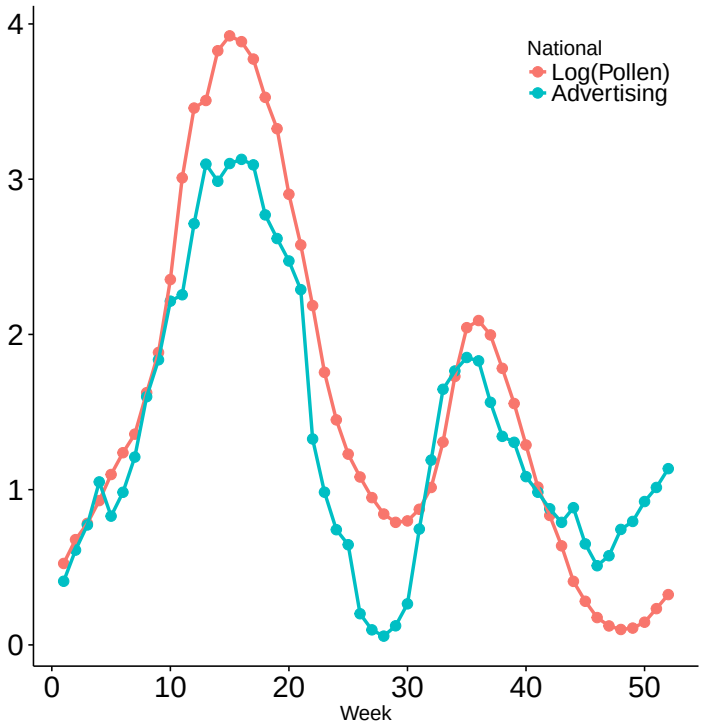
Estimating the effect of advertising is often made difficult by factors that are unobserved to the econometrician, but drive changes both in demand and firms' advertising decisions. This concern frequently arises for academic researchers, and may also arise for firms wishing to understand the influence of their competitor's advertising, or even their own. In this section, I develop a method for econometricians to obtain unbiased estimates when they have only partial knowledge of the factors that drove advertising decisions. The method relies on the observation that advertising levels are often locally sub-optimal, even for firms that behave optimally.

2.4.1 Overview of the Identification Strategy

The mechanics of the coarseness strategy are especially tangible in the antihistamine market. This market has a clear and tangible measure of organic demand: pollen counts. Indeed, pollen drives much of the variation in antihistamine sales, and the typical pollen cycle appears to also drive changes in advertising levels over the course of a year. For evidence of the relationship between the pollen and advertising cycles, see Figure 2.2 which shows the national averages of pollen and advertising levels in the United States averaged across 2006-2013 for each of the 52 weeks of the year. In this figure, we can see that average pollen levels and advertising track each other closely, each rising steadily from the beginning of the new year and reaching their annual peak in the spring before falling back down in the middle of the summer. Each trend rises again in the fall, but to only a fraction of the level reached in the spring. That these two trends track each other so closely suggests firms are targeting their advertising on the forecasted level of pollen.

Given this behavior, one can see how advertising levels may not always be optimal. The first reason for sub-optimality is fairly apparent: when firms purchase their advertising for

Figure 2.2. POLLEN AND ADVERTISING



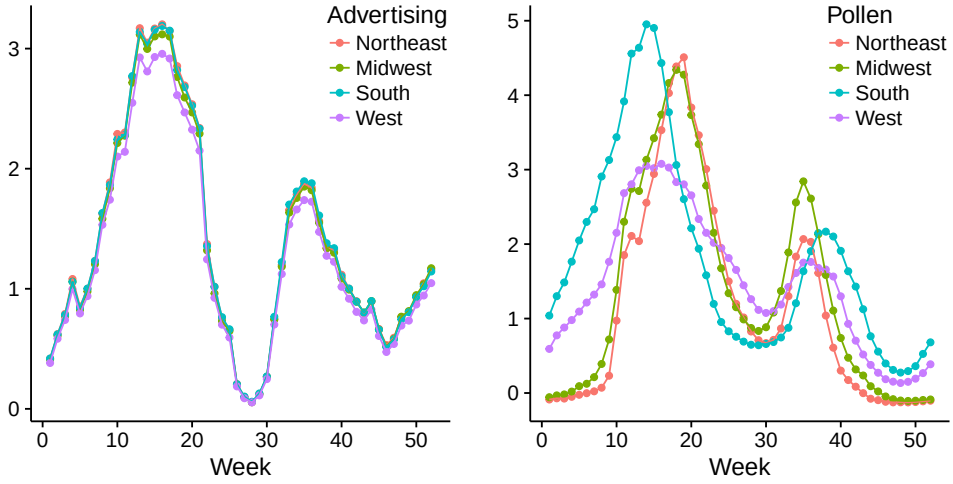
Notes: Nationwide advertising and log of pollen levels by week of the year, averaged across 2006-2013.

a coming year, they cannot perfectly predict what pollen levels will be. Year to year, pollen levels may be higher or lower than expected, or the bloom may occur earlier or later. By definition, firms have no control over this prediction error, and as a result, it provides a good source of quasi-random variation that can be used to estimate the response to advertising.

The second reason for sub-optimal advertising levels may be less apparent, but turns out to be an even stronger source of quasi-random variation for many products, including antihistamines. These sub-optimal advertising levels also result from the fact that firms will often purchase most or all of their television advertising nationally because it is cheaper. Although this strategy may be globally optimal for the firm, a side effect is that all locations experience the same advertising intensity, even if their local demand conditions are different. As a result, the advertising levels are not locally optimal.

For a dramatic illustration of this effect, see Figure 2.3. This figure shows the same data as Figure 2.2, but now disaggregated to census-region level. Figure 2.3 shows that the advertising in each of the regions is virtually identical, whereas the pollen levels are quite different. For example, the South typically experiences spring several weeks before the other regions, and reaches a slightly higher peak than any of the others. Alternatively, the fall allergy season hits the Midwest the earliest, and the hardest. Further differences in pollen levels exist for even smaller geographic definitions. See Appendix B.1.3 for more on these regional differences.

Figure 2.3. REGIONAL ADVERTISING AND POLLEN



Notes: Advertising and log of pollen levels by Census Region and week of the year, averaged across 2006-2013.

To develop a deeper understanding of the identifying variation that arises in the coarseness strategy, consider the advertising and pollen experiences of two very different cities: Houston and Minneapolis. Their experiences in 2011 are illustrated in Figure 2.4, with their organic demand levels represented by the blue line (labeled D_{dt}). Naturally, these demand profiles were quite different given the different climates of these two cities. Nevertheless, each city experienced the same national advertising. This national advertising would be based, at least in part, on the nationally forecasted pollen levels (F_t), represented by the red

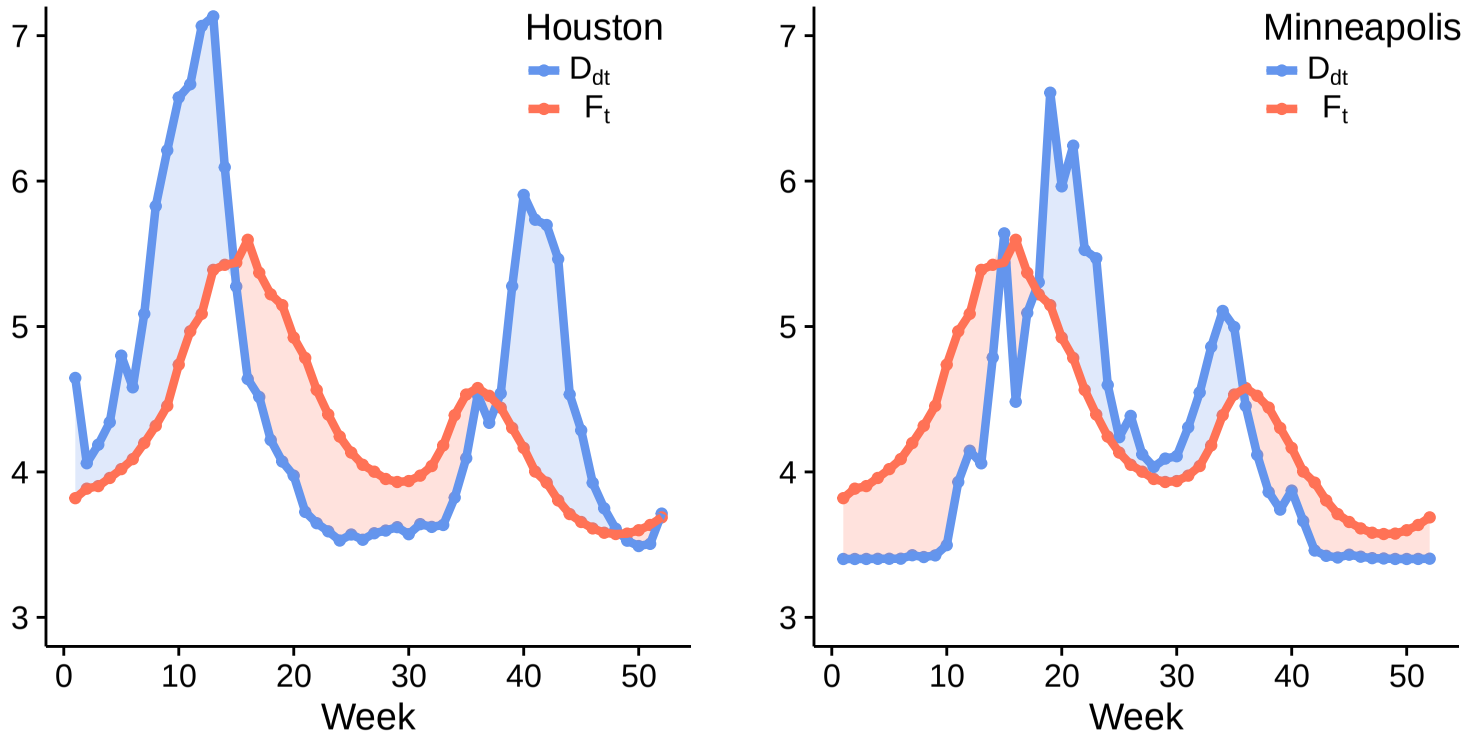
line. As a result, each of these cities is frequently experiencing either too much or too little advertising. For example, around the 10th week of the year, Houston was hit with very high pollen levels, but received relatively modest levels of advertising because many other parts of the United States had little to no demand for antihistamines. By contrast, when Minneapolis was hit with its spring pollen season at about week 20, it received relatively little advertising, because the spring allergy season was coming to an end for most of the country. These types of outcomes are the ones my method treats as quasi-random, and are therefore well suited to identifying the response to advertising. Blue-shaded regions represent observations in which the city is receiving too little advertising, and red-shaded regions represent observations in which the city is receiving too much advertising. Note that at any given time, while one city is receiving too much, the other is receiving too little. Although Houston and Minneapolis represent extreme cases, these types of differing experiences exist across all cities every week of the year. Together, these many differences generate power to estimate the advertising response.

Thus we have seen that even an optimally advertising firm can generate advertising levels that are measurably sub-optimal for some populations at some times. These sub-optimalities are a result of (1) their inability to perfectly forecast future organic demand levels, and (2) their incentives to purchase their advertising at high levels of aggregation, which may be locally sub-optimal. The next section will present a formal model for exploiting these sub-optimalities.

2.4.2 Model

Having overviewed the identification strategy in the last section, I introduce a formal model in this section. Formalizing the model will allow us to exploit both sources of quasi-random variation discussed in the last section. The model takes the form of an IV expression. For

Figure 2.4. IDENTIFICATION EXAMPLE



Notes: The blue line presents the organic demand (D_{dt}) realized in Houston and Minneapolis in 2011. The red line is the nationally forecasted organic demand (F_t).

reference, Table 2.1 describes the key variables used in this paper and their sources, which are introduced below.

First, consider the following first-stage expression for advertising:

$$A_{dt} = \alpha_d^0 + \xi^0 X_t + \gamma_0^0 D_{dt} + \gamma_1^0 F_t + \varepsilon_{dt}^0, \tag{2.1}$$

where A_{dt} is the log of the advertising stock in DMA d during week t. α_d^0 are DMA fixed effects, and X_t contains calendar-year fixed effects. F_t is the forecasted level of organic demand in week t nationally, and serves as the excluded variable. D_{dt} is the measure of the organic demand level realized in DMA d in week t. By including D_{dt} in the expression along with F_t, we effectively capture the difference between these two variables, which measures the degree to which a given DMA in a given week received too much or too little advertising, as was illustrated in Figure 2.4.

We are now ready to write the second-stage expression as

$$y_{dt} = \alpha_d^1 + \xi^1 X_{dt} + \gamma_0^1 D_{dt} + \gamma_1^1 F_t + \beta A_{dt} + \varepsilon_{dt}^1, \tag{2.2}$$

where y_{dt} is the log of sales in DMA d during week t. Each of the covariates from equation (2.1) are included but have new coefficients with a superscript 1. β will now be estimated using only the variation in advertising that is correlated with expected national demand, but not local demand. Specifically, the exclusion restriction requires that $cov\left(\varepsilon_{dt}^1, F_t \mid \alpha_d, X_{dt}, D_{dt}\right) = 0$.

Although this model can be estimated using data aggregated to the DMA-week level, I wish to take advantage of the precision offered by the store-level data. Rewriting equation (2.2) at the store level we have

$$y_{it} = \alpha_i^1 + \xi^1 X_{it} + \gamma_0^1 D_{dt} + \beta A_{dt} + \varepsilon_{it}^1, \tag{2.3}$$

Table 2.1: VARIABLE DESCRIPTIONS

Variable	Description	Source
D_{dt}	Realized organic demand in DMA d during week t.	Antihistamines: $D_{dt} = log(\text{pollen count} + 30)$, where the pollen counts are imputed from station data. See Appendix B.1.1 for imputation details. Others categories: $D_{dt} \in \{\text{average temperature, dew point}\}$
F_{dt}	Forecasted organic demand in DMA d for week t.	$F_{dt} = \frac{1}{N_{year}} \sum_{j=t^-}^{t^+} \mathbf{1}_{f(t)=f(j)\cdot} \cdot D_{dj}$, where $f(t) \to \tau$, $\tau \in \{1,2,...,52\}$ corresponds to a week of the year, t^- is the first week of the first observed year, t^+ is the last week of the last observed year and N_{year} is the number of observed years.
F_t	Forecasted organic demand nationally for week t, weighted by population.	$F_t = \sum_{i=1}^{N_d} F_{dt} \cdot \frac{P_d}{\sum_{j=1}^{N_d} P_j}$, where N_d is then number of DMAs in the sample, and P_d is the population of DMA d.
a_{dt}	Average advertising exposures per person in DMA d during week t.	Aggregated from Nielsen Media data.
λ	Advertising carryover, week-to-week.	Estimated using NLLS. See Appendix B.2.1.
A_{dt}	Advertising stock in DMA d for week t.	$A_{dt} = log\left(\sum_{\tau=T^-}^{0} \lambda^{\tau} a_{d(t+\tau)}\right)$. See Appendix B.2.1.
X_{it}	Control variables: calendar-year fixed effects, the log of average local antihistamine prices, share of antihistamine on feature in store i at time t.	Nielsen RMS data.
y_{it}	Log of total antihistamine sales in store i during week t.	$y_{it} = log(\sum_J c_{ijt})$ where c_{ijt} is the number of tablets of product j sold in store i during week t as reported in Nielsen RMS. See Appendix B.1.2 for market definition.

where y_{it} is the log of sales in store i during week t. α_i are store-level fixed effects and X_{it} now contains store-level estimates of the log of prices, in addition to calendar-year fixed effects. To avoid aggregation bias in the standard errors, they are clustered by DMA (Moulton [1990]).

A common alternative approach to the method introduced above is to use ordinary least squares (OLS) with organic demand as an additional regressor. This approach can be written as:

$$y_{it} = \alpha_i + \xi X_{it} + \gamma D_{dt} + \beta A_{dt} + \varepsilon_{it}, \tag{2.4}$$

where the variables are defined as above. This approach uses more variation in advertising than used by the IV strategy. The identifying assumption in this case is that, conditional on local pollen levels (and other controls), the level of advertising received is essentially random. If correct, this approach will provide more efficient estimates. However, if unobserved factors drove the advertising decision and influence sales, this approach will produce biased estimates.

2.5 APPLICATIONS

2.5.1 Estimates for the Antihistamine Category

To start, I apply the method developed in the last section to the antihistamine category. This analysis indicates that the method can have good power, and standard methods may produce biased estimates.

Estimates from the coarseness strategy described by equations (2.1) and (2.3) are shown in Table 2.2. From the first-stage estimates we can see the instrument is very strong, both in terms of the partial F-statistic for excluded variable, F_t, and its t-statistic. Specification

(a) leaves out the control for expected local demand levels in order to illustrate the bias that results from violating the exclusion restriction, which appears to be quite large. Controlling for local demand levels in specification (b) reduces the estimate by a factor of 4 or more.

Having established that the IV has good power, I now compare its estimates with those produced by other methods in Table 2.3. First, column (a) shows estimates from a simple OLS regression that does not control for organic demand levels. Here we obtain fairly large advertising elasticity estimates. A common approach to controlling for the confounding effect of seasonality would be to include more time fixed effects. Specification (b) does so by adding month fixed effects to the specification in (a). However, this approach does not dramatically change the estimated advertising elasticity. An alternative means of controlling for seasonality would be to include a control for organic demand such as is in column (c). This approach roughly halves the estimating advertising elasticity of the previous methods, which suggests the estimates in (a) and (b) were fairly biased upward. Including both a control for organic demand and month fixed effects in column (d) does not change the estimate significantly. Finally, column (e) presents the estimate from the IV strategy, which is a bit larger, but not statistically different from the previous two estimates.

Together, the estimates in this table provide evidence of the large bias that can arise if the econometrician fails to control for the factors that drove advertising decisions. In this example, our measurement of organic demand reduces the bias to the point where the IV and OLS estimates are indistinguishable. However, we will see in the next section that OLS and IV estimates are not always equivalent, even with a good measure of organic demand.

Table 2.2: INSTRUMENTING FOR ADVERTISING ELASTICITY

	(a)	(b)
First Stage:		
Forecasted Organic Demand, National (Instr.)	0.509	0.513
	(0.001)	(0.033)
Organic Demand		-0.004
		(0.033)
Log of Price	0.076	0.076
	(0.007)	(0.005)
F-statistic for Excluded Variable	59,823	79
R^2	0.491	0.491
Second Stage:		
Advertising Elasticity	0.694	0.171
	(0.031)	(0.026)
Organic Demand		0.266
		(0.013)
Log of Price	-1.038	-1.024
	(0.012)	(0.011)
Year Fixed Effects	X	X
Store Fixed Effects	X	X
Number of Observations	12,036,767	12,036,767
Number of Stores	39,371	39,371
Number of Clusters	204	204
R^2	0.872	0.895

Note: Estimates of the advertising elasticity for the antihistamine industry based on Nielsen data that cover 2006-2013 for each week and store. Robust standard errors, clustered by DMA, are reported in parentheses. Regressions are weighted by the number of units sold in a store-week. "Organic Demand" is a log transformation of the total pollen count. The instrument for the log of advertising is the log of the expected national pollen level for a given week.

Table 2.3: ANTIHISTAMINES CATEGORY ADVERTISING ELASTICITY

	(a)	**(b)**	**(c)**	**(d)**	**(e)**
Advertising Elasticity	0.345	0.288	0.134	0.121	0.171
	(0.021)	(0.025)	(0.012)	(0.016)	(0.026)
Log of Price	-0.995	-0.959	-1.022	-0.986	-1.024
	(0.012)	(0.011)	(0.011)	(0.010)	(0.011)
Organic Demand			0.278	0.270	0.266
			(0.011)	(0.012)	(0.013)
Year Fixed Effects	X		X		X
Month-Year Fixed Effects		X		X	
Store Fixed Effects	X	X	X	X	X
Number of Observations	12,039,907	12,039,907	12,036,767	12,036,767	12,036,767
Number of Stores	39,371	39,371	39,371	39,371	39,371
Number of Clusters	204	204	204	204	204
R^2	0.883	0.891	0.895	0.897	0.895
F-stat (excl. var.)					79
Estimation Method	OLS	OLS	OLS	OLS	IV

Note: Estimates of the advertising elasticity for the antihistamine industry based on Nielsen data that cover 2006-2013 for each week and store. Robust standard errors, clustered by DMA, are reported in parentheses. Regressions are weighted by the number of units sold in a store-week. "Organic Demand" is a log transformation of the total pollen count. The instrument for the log of advertising is the log of the expected national pollen level for a given week.

2.5.2 *Additional Category-level Estimates*

In this section, I expand on the estimates from section 2.5.1, and apply multiple estimation procedures across a number of different product categories. These estimates suggest many of the existing estimation procedures produce biased estimates, whereas the coarseness strategy presented in this paper provides consistent estimates.

The estimation methods used in section 2.5.1 are used again in Figure 2.5. "OLS: naive" refers to an OLS regression as in equation (2.4), but without any control for organic demand, D_{dt}. "OLS: FE" adds month fixed effects to "OLS: naive," while "OLS: OD" runs the specification listed in equation (2.4). "Border" refers to the border strategy presented in Shapiro (2016), which exploits the discontinuity in advertising levels across adjacent counties that belong to different DMAs. Finally, "IV" refers to the coarseness strategy presented in this paper.

Organic demand measures differ across categories. For antihistamines, a transformation of the total pollen count is again used for organic demand. Organic demand is measured with average temperature for sunscreens, whereas the dew point measures organic demand for lip medications and moisturizers. Summary statistics for each of these categories, including their respective measures of organic demand, are presented in Appendix Figure B.6.

A few patterns can be observed in Figure 2.5. First, the "OLS: naive" estimates are often much larger than the others, which is unsurprising given that these estimates are likely to be upward biased by the "unobserved" influence of organic demand. Adding month fixed effects often reduces these estimates, but not always. Controlling for organic demand always reduces the estimates, but in some cases (e.g., lip moisturizers and sunscreens) produces negative estimates. Although advertising in these industries might have a negative effect on sales (e.g., if firms are stuck in a prisoner's dilemma-type equilibrium), most ex-

isting research estimates a positive response to advertising. The border strategy does not produce significant estimates for any category, because it relies on differences in advertising intensity across adjacent DMAs, and if too little advertising is bought at the DMA level, it will not work. Finally, the coarseness strategy typically produces larger confidence intervals than the other estimations because it relies only on the small part of the variation in advertising that is assumed to be quasi-random. However, the point estimates are almost always significant and are always positive and smaller than the "OLS: naive" estimates.

Details for these category-level regressions can be found in Appendix Tables B.3 to B.6.

2.5.3 Estimates for Antihistamine Brands

Brand-level estimates can be more challenging to obtain than category-level estimates. Not only are the factors that drove a given firm's advertising decisions likely to be unobserved, but the same is likely to be true for the firm's competitors. As a result, at least two instruments are required: one for own advertising, and one for competitive advertising. For many firms, this requirement may not be difficult to meet: they may have quantitative measures of many of the competing factors that drive both their – and their competitors – advertising decisions. For demonstration purposes, I will return to the antihistamine market for which many independent measures of organic demand are available.

In previous analyses of this market, I used the total pollen count as the measure of organic demand. However, many different types of pollen are produced by different plants that affect people's allergies differently. For the brand-level analysis, I have selected eight of the most influential pollen genuses available in the NAB data: eupatorium, ligustrum, olea, fagus, typha, artemisia tsuga, and "other grass pollen." Each will serve as a separate measure of organic demand that drives firms' advertising decisions and local demand levels.

The specification I use for brand-level estimates is

$$y_{ijt} = \alpha_{ij} + \xi X_{ijt} + \xi^c X^c_{ijt} + \gamma \vec{D}_{dt} + \beta A_{djt} + \beta^c A^c_{djt} + \varepsilon_{ijt}, \tag{2.5}$$

where the variables are defined as before, but the c superscripts refer to "competitors." Also, note that $\vec{D}_{dt}$ (and $\vec{F}_t$ for the IV approach: see equation [2.1]) are now vectors that contain each of the eight pollen counts. Additionally, the j subscript indexes the brands Allegra, Claritin, Claritin-D, and Zyrtec, which correspond to the antihistamine brands that each contribute more than 5% of the antihistamine television advertising.

Estimates from this specification for β and β^c are presented in Figure 2.6. The methods "OLS: naive" and "Border" are as described previously. "OLS: genus" controls for the pollen measurements for each genus, and "IV: genus" uses each genus as a separate measure of organic demand, D_{dt}, with its own national forecast, F_t. For most of the brands in this figure, the IV estimates indicate a larger response to advertising than the other methods. One explanation for the larger response could be measurement error in the advertising variable, which would attenuate the OLS advertising estimates, especially the border strategy, which includes considerably more fixed-effects coefficients than the other models. Measurement error would not bias the IV estimates, though it could reduce its precision.

Estimates for the effect of Benadryl ads on Benadryl sales do present some puzzling differences, however. For this effect, the border strategy estimates a marginally positive response, IVs (and non-border OLS estimates) estimate negative responses. In general, negative advertising responses present a red flag, though in this case, note that Benadryl discontinued television advertising in 2011 when Allegra, the third offering of the "new generation" antihistamines, was released. The negative estimates may also reflect the limitations of applying the coarseness strategy without internal firm data. As a result of this data limitation, the instrument – forecasted organic demand – is the same across years. In

reality, firms make continuous adjustments to their forecasts, which would increase the accuracy and precision of the coarseness estimates, especially for brand-level estimates for which multiple independent instruments are required.

2.5.4 Google Trends as a Measure of Organic Demand

The coarseness strategy relies crucially on local measurements of organic demand, understanding which measures can be used is worthwhile. One such measure that could potentially be used for a variety of products is Google Trends. However, I find evidence that the Google Trends data contain at least two types of measurement error. The first type is classical measurement error, in which the Google Trends measurement differs from the level of organic demand by some random amount. The second type of error results from the possibility that some Google searches are driven by people having seen advertisements (see Liaukonyte, Teixeira, and Wilbur [2015]). I find that although adjustments can be made for classical measurement error in the Google Trends data, its apparent response to advertising makes it ill suited to the coarseness strategy.

The Google Trends data report the relative search intensity for a given word or phrase. The search intensity for "allergies" at the city-week level was scraped from the Google site for this study. In years before 2013, significant measurement error exists: many of the trends in searches move in a highly random or stepwise manner. For this reason, I restrict analysis to 2013 alone.

Classical measurement error and the influence of advertising on Google searches are likely to bias our advertising estimates in opposite directions. Classical measurement error is well known to attenuate the coefficients of the variable with measurement error, but for variables that are positively correlated with that variable, the estimates will be biased upwards. Intuitively, one could imagine an error-ridden measure of organic demand would do

a poor job controlling for the confounding influence of true organic demand, so advertising estimates would behave more like the condition in which we did not control for organic demand at all.

Alternatively, advertising responses in the Google Trends data will bias the advertising estimates downwards, which can be seen as follows. Assume

$$G = D + \delta A,$$

where G is the observed Google Trends observation, A is advertising intensity, δ is a positive coefficient, and D is the true organic demand. Then, plugging in G for D in equation 2.4, we have

$$y_{it} = \alpha_i + \xi X_{it} + \gamma(D_{dt} + \delta A_{dt}) + \beta A_{dt} + \varepsilon_{it}, \tag{2.6}$$

which causes us to estimate

$$y_{it} = \alpha_i + \xi X_{it} + \gamma D_{dt} + \beta^{'} A_{dt} + \varepsilon_{it}, \tag{2.7}$$

where $\beta^{'} = \beta - \gamma\delta$ and is a downward-biased estimate when $\gamma > 0$ and $\delta > 0$, as is likely the case here.

Estimates consistent with these respective biases can be found in Table 2.4. First consider columns (a) and (b), which examine whether measurement error in the pollen data influence our estimates. Column (a) reports estimates using standard OLS, whereas (b) reports estimates using two independent measures of pollen: each generated from a random selection of different pollen stations. The pollen measurements for these two sets of randomly selected stations were used as training data to predict pollen levels throughout the United States using the method described in Appendix B.1.1. With these two measurements

of pollen, we are able to use one to instrument for the other, and eliminate the influence of measurement error in the original data. Columns (a) and (b) are not statistically different, suggesting the pollen estimates contain limited measurements error.

Column (c) presents a regression using the raw Google Trends data as our control for organic demand. The advertising estimate in (c) is considerably larger than in columns (a) and (b), suggesting these values contain significant classical measurement error. The measurement error is likely to be reduced when running the imputation algorithm described in Appendix B.1.1. Indeed, using the predictions from this algorithm as the measure of organic demand in column (d), the advertising measurements are significantly reduced. Finally, in column (e), we repeat the exercise from column (b) in which we randomly select two sets of pollen stations on which to run the imputation algorithm. Using one prediction as the instrument for the other to remove the effect of the classical measurement error, we obtain even smaller estimates for the influence of advertising. Now that the influence of measurement error has been removed, this final estimate suggests the Google Trends data do contain considerable responses to advertising, causing significant downward bias as suggested by equation 2.7.

This evidence suggests the Google Trends data should be used with caution as a control for advertising confounds, and is not suited for use with the IV strategy described in this paper.

2.6 Implications for the Firm

The strategy presented here has several features that may make it especially appealing to firms. Firms are likely to have measures of the factors that drive their advertising decisions. Those firms that can quantify all of these factors will not need this strategy in order to estimate their own advertising, because they can condition on all the information that went

Table 2.4: MEASUREMENT ERROR CORRECTIONS

	(a)	(b)	(c)	(d)	(e)
Advertising Elasticity	0.155	0.126	0.224	0.123	0.057
	(0.021)	(0.019)	(0.015)	(0.019)	(0.018)
Organic Demand	0.313	0.367	1.360	2.203	2.503
	(0.014)	(0.017)	(0.080)	(0.097)	(0.129)
Log of Price	-1.751	-1.744	-1.771	-1.759	-1.751
	(0.038)	(0.038)	(0.038)	(0.038)	(0.038)
Org. Dem. Measure	Pollen	Pollen	(Raw) GTrend	GTrend	GTrend
Year Fixed Effects	X	X	X	X	X
Store Fixed Effects	X	X	X	X	X
Number of Observations	1,296,961	1,296,961	1,296,961	1,296,961	1,296,961
Number of Stores	29,063	29,063	29,063	29,063	29,063
Number of Clusters	87	87	87	87	87
R^2	0.915	0.915	0.913	0.914	0.913
1st Stage F-statistic (excl. var.)		721			77
Org. Dem. Error Correction with IV		X			X

Note: Estimates of the advertising elasticity using various measures of organic demand with and without correction for measurement error. All measures of organic demand were imputed using the method described in the Appendix, except where indicated "(Raw)," in which case the original values were used. All regressions were run on the subset of data from 2013 for which Google Trends data were available. I corrected for the measurement error in organic demand by using two estimates of organic demand: one measure of organic demand as the instrument for the other.

into their advertising decision. However, in many cases, quantifying all the factors that affected their decision may be difficulty, and in these cases, the coarseness strategy may be useful. Additionally, when considering the effect of their competitors' advertising on their brands, they will only have measures of some of the factors that drove their competitors' decisions, making the coarseness strategy a good fit. Finally, firms may find this approach appealing because it does not require them to alter their advertising strategy, unlike standard experiments, which require firms to trade off short term inefficiencies for long-run gains.

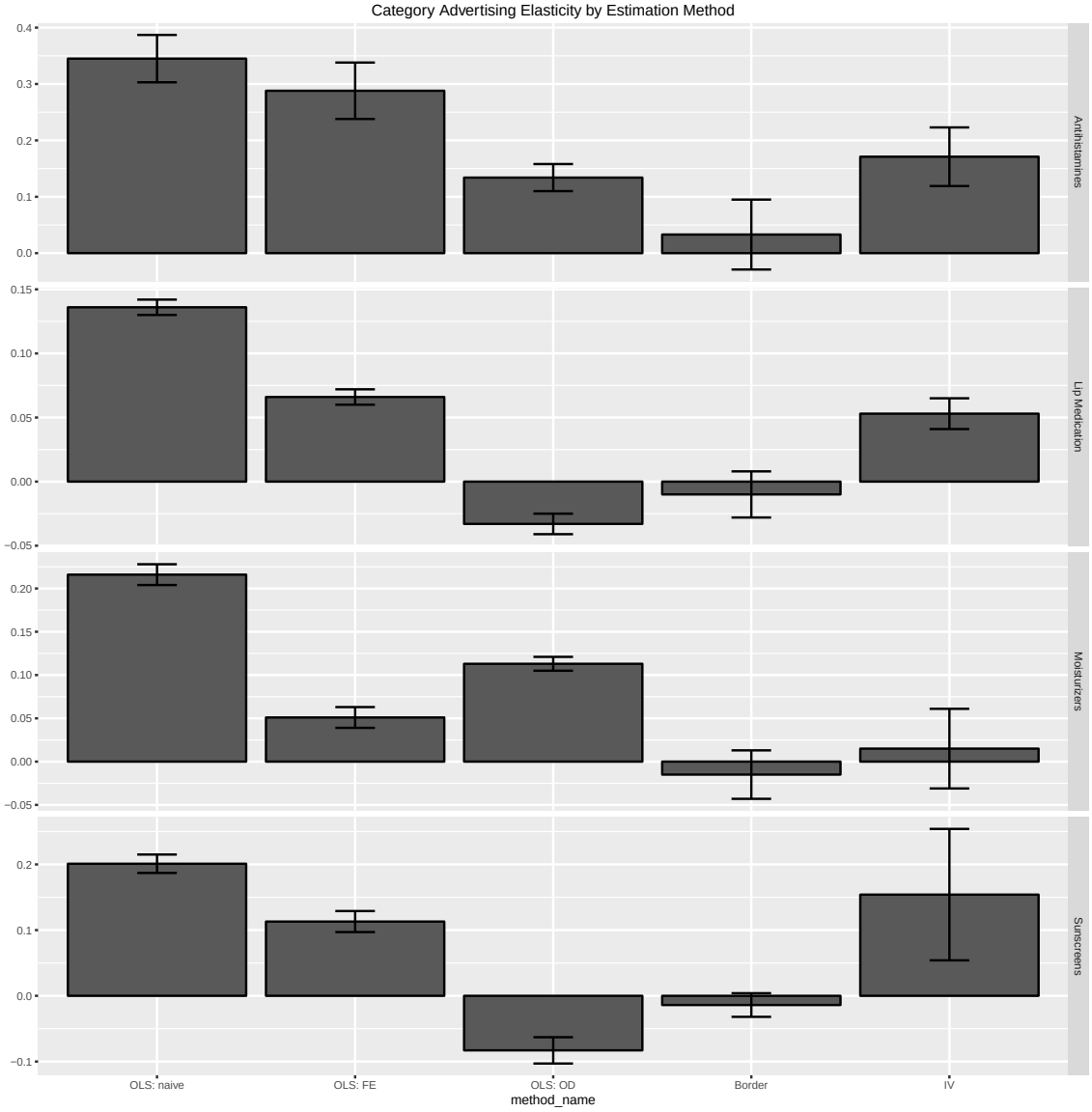
Additionally, the strategy is general enough that it may apply to marketing activity other than advertising. The coarseness strategy may provide useful feedback on the effectiveness of direct-to-consumer marketing such as mailings or emails, which are necessarily "coarse" because each consumer either receives treatment or not, despite having different predicted propensities to respond. Although price elasticties are perhaps less prone to endogeneity bias than advertising, in cases in which endogeneity concerns arise, the coarseness strategy may also prove useful. Often a store will charge all consumers the same price: therefore the strategy discussed in this paper would apply.

2.7 CONCLUSIONS

This paper presents a novel method for estimating the response to advertising. It exploits the common occurrence in which firms implement their advertising (or other marking activity) at high-enough levels of aggregation that measures of organic demand allow the econometrician to approximate the degree to which advertising intensity is not optimal for all who are exposed to it. Such situations are likely to continue to occur due to discounts for purchasing and implementing advertising in bulk and as measures of organic demand become increasingly available. The method provides firms a means to obtain unbiased estimates of the effect of their own and their competitor's marketing activity without having

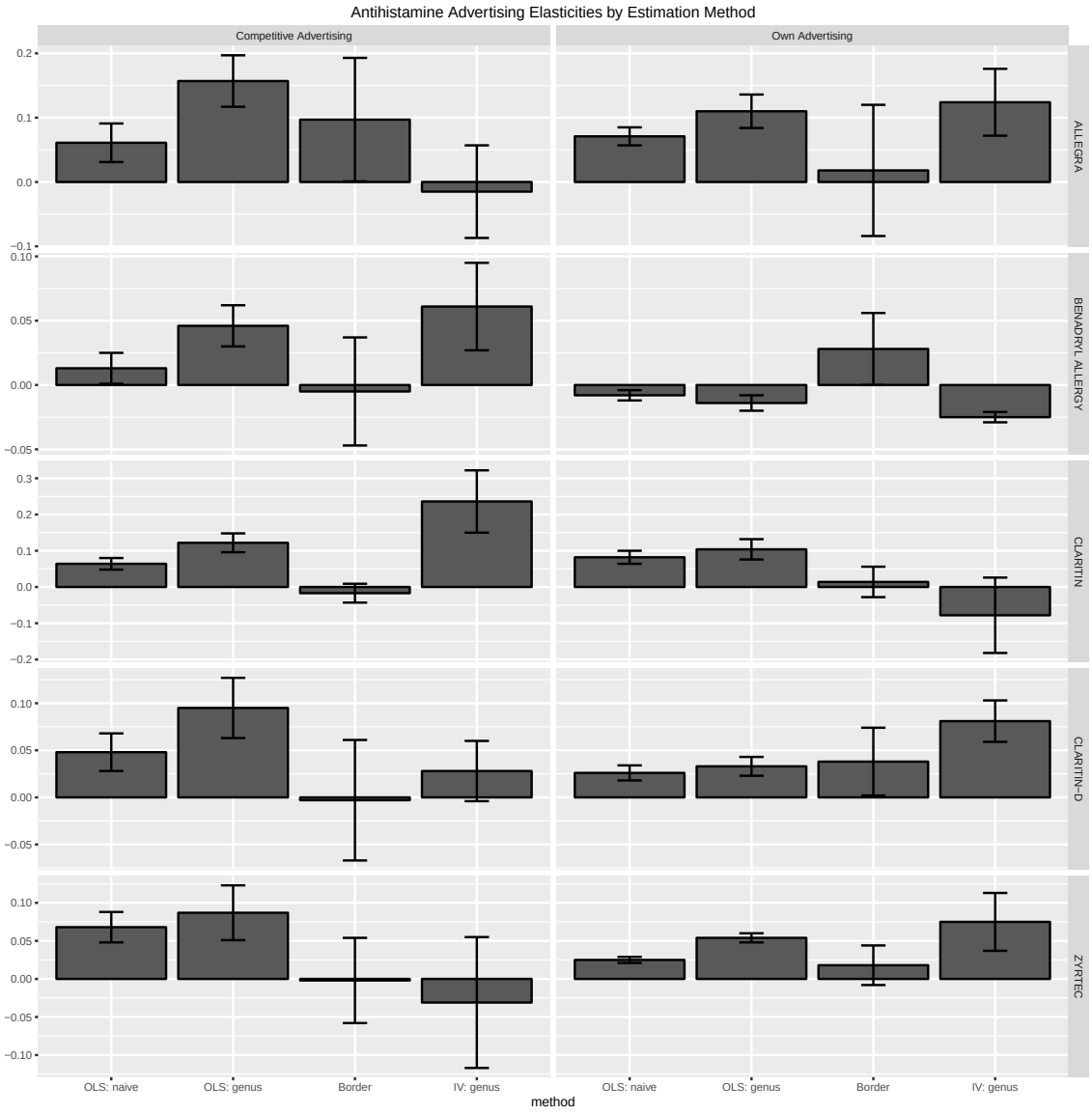
to alter their behavior. Examples of applications presented in the paper suggest standard methods for producing such estimates can produce quite biased estimates.

Figure 2.5. CATEGORY-LEVEL REGRESSIONS



Notes: Estimates for multiple categories using different estimation methods. "OLS: naive" refers to a regression of the log of sales on advertising with controls for DMA fixed effects and average prices, but not organic demand. "OLS: FE" adds month-level fixed effects to the "OLS: naive" specification. "OLS: OD" adds a control for organic demand to the "OLS: naive" specification. "Border" refers to the border strategy presented in Shapiro (2016), which compares the advertising levels of stores from adjacent counties, but different DMAs. "IV" refers to the estimation strategy presented in this paper in equations (2.1) and (2.3). The measures of organic demand for each category are Antihistamines: pollen; Lip Medication: dew point; Moisturizers: dew point; Sunscreens: average temperature. Standard errors (SE) were clustered by DMA, and error bars represent $\pm 2 * SE$.

Figure 2.6. BRAND-LEVEL REGRESSIONS: ANTIHISTAMINES



Notes: Estimates of the advertising coefficients for own-brand advertising and advertising from competing brands. Regressions were run separately for each brand. "OLS: naive" refers to a regression of the log of sales on advertising with controls for DMA fixed effects and average prices, but not organic demand. "OLS: genus" controls for different types of pollen to the "OLS: naive" specification. "Border" refers to the border strategy presented in Shapiro (2016), which compares the advertising levels of stores from adjacent counties, but different DMAs. "IV: genus" refers to the estimation strategy presented in this paper in equations (2.1) and (2.5), with each pollen genus representing a different measure of organic demand. The "competition" for any brand includes the sum of other antihistamine brands that were available in a given year: Claritin, Claritin-D, and Benadryl were both OTC in 2006, Zyrtec became OTC in 2008, and Allegra in 2011. Standard errors (SE) were clustered by DMA, and error bars represent $\pm 2 * SE$.

REFERENCES

Abernethy, Avery M., and Jesse E. Teel. "Advertising regulation's effect upon demand for cigarettes." *Journal of Advertising,* 15, no. 4 (1986): 51-55.

Anand, Bharat N., and Ron Shachar. "Targeted advertising as a signal." *Quantitative Marketing and Economics* 7, no. 3 (2009): 237-266.

Anderson, Simon P., Federico Ciliberto, Jura Liaukonyte, and Regis Renault. "Push-me pull-you: Comparative advertising in the OTC analgesics industry." Available at SSRN 2047106 (2015).

Andrews, Michelle, Xueming Luo, Zheng Fang, and Anindya Ghose. "Mobile ad effectiveness: Hyper-contextual targeting with crowdedness." *Marketing Science* 35, no. 2 (2015): 218-233.

Andrews, Rick L., and George R. Franke. "The Determinants of Cigarette Consumption: A Meta-Analysis." *Journal of Public Policy & Marketing* (1991): 81-100.

Ansari, Asim, and Carl F. Mela. "E-customization." *Journal of marketing research* 40, no. 2 (2003): 131-145.

Assmus, Gert, John U. Farley, and Donald R. Lehmann. "How advertising affects sales: Meta-analysis of econometric results." *Journal of Marketing Research* (1984): 65-74.

Bader, Pearl, David Boisclair, and Roberta Ferrence. "Effects of tobacco taxation and pricing on smoking behavior in high risk populations: a knowledge synthesis." *International journal of environmental research and public health* 8, no. 11 (2011): 4118-4139.

Baltagi, Bad H., and Dan Levin. "Estimating Dynamic Demand for Cigarettes Using Panel Data: The Effects of Bootlegging, Taxation and Advertising Reconsidered." *The Review of Economics and Statistics* (1986): 148-155.

Bass, Frank M., and Darral G. Clarke. "Testing distributed lag models of advertising effect." *Journal of Marketing Research* (1972): 298-308.

Bedard, Kelly, and Olivier Deschênes. "The Long-Term Impact of Military Service on Health: Evidence from World War II and Korean War Veterans." *The American Economic Review* (2006): 176-194.

Besanko, David, Sachin Gupta, and Dipak Jain. "Logit demand estimation under competitive pricing behavior: An equilibrium framework." *Management Science* 44, no. 11-part-1 (1998): 1533-1547.

Beville, H.M, of the National Broadcasting Company, "Radio and Television Audience", Talk at the Pulse Luncheon, October 28, 1949.

Blake, Thomas, Chris Nosko, and Steven Tadelis. "Consumer Heterogeneity and Paid Search Effectiveness: A Large-Scale Field Experiment." *Econometrica* 83, no. 1 (2015): 155-174.

Bitran, Gabriel R., and Susana V. Mondschein. "Mailing decisions in the catalog sales industry." *Management science* 42, no. 9 (1996): 1364-1381.

Blattberg, Robert C., Byung-Do Kim, and Scott A. Neslin. Why Database Marketing?. Springer New York, 2008.

Breiman, Leo. "Random forests." *Machine learning* 45, no. 1 (2001): 5-32. Harvard

Bronnenberg, Bart J., and Vijay Mahajan. "Unobserved retailer behavior in multimarket data: Joint spatial dependence in market shares and promotion variables." *Marketing Science* 20, no. 3 (2001): 284-299.

Bronnenberg, Bart J., Peter E. Rossi, and Naufel J. Vilcassim. "Structural modeling and policy simulation." *Journal of Marketing Research* 42, no. 1 (2005): 22-26.

Burczyk, J., S. P. DiFazio, and W. T. Adams. "Gene flow in forest trees: how far do genes really travel?." *Forest Genetics* 11, no. 3/4 (2004): 179.

Chaloupka, Frank J., and Kenneth E. Warner. "The Economics of Smoking." *Handbook of Health Economics* 1 (2000): 1539-1627.

Ching, Wai-Ki, Michael K. Ng, Ka-Kuen Wong, and Eitan Altman. "Customer lifetime value: stochastic optimization approach." *Journal of the Operational Research Society* 55, no. 8 (2004): 860-868.

Cox, David R. "Regression Models and Life-Tables." *Journal of the Royal Statistical Society. Series B (Methodological)* (1972): 187-220.

Dalton, Madeline A., James D. Sargent, Michael L. Beach, Linda Titus-Ernstoff, Jennifer J. Gibson, M. Bridget Ahrens, Jennifer J. Tickle, and Todd F. Heatherton. "Effect of viewing smoking in movies on adolescent smoking initiation: a cohort study." *The Lancet* 362, no. 9380 (2003): 281-285.

Dalton, Madeline A., Michael L. Beach, Anna M. Adachi-Mejia, Meghan R. Longacre, Aurora L. Matzkin, James D. Sargent, Todd F. Heatherton, and Linda Titus-Ernstoff. "Early exposure to movie smoking predicts established smoking by older teens and young adults." *Pediatrics* 123, no. 4 (2009): e551-e558.

Douglas, Stratford, and Govind Hariharan. "The Hazard of Starting Smoking: Estimates from a Split Population Duration Model." *Journal of Health Economics* 13, no. 2 (1994): 213-230.

Douglas, Stratford. "The Duration of the Smoking Habit." *Economic Inquiry* 36.1 (1998): 49-64.

Dubé, Jean-Pierre, Günter J. Hitsch, and Puneet Manchanda. "An empirical model of advertising dynamics." *Quantitative Marketing and Economics* 3, no. 2 (2005): 107-144.

Elberse, Anita, and Jehoshua Eliashberg. "Demand and supply dynamics for sequentially released products in international markets: The case of motion pictures." *Marketing Science* 22, no. 3 (2003): 329-354.

Erdem, Tülin, and Michael P. Keane. "Decision-making under uncertainty: Capturing dynamic brand choice processes in turbulent consumer goods markets." *Marketing science* 15, no. 1 (1996): 1-20.

Erdem, Tülin, Susumu Imai, and Michael P. Keane. "Brand and quantity choice dynamics under price uncertainty." *Quantitative Marketing and Economics* 1, no. 1 (2003): 5-64.

Feichtinger, Gustav, Richard F. Hartl, and Suresh P. Sethi. "Dynamic optimal control models in advertising: recent developments." *Management Science* 40, no. 2 (1994): 195-226.

Feinberg, Fred M. "On continuous-time optimal advertising under S-shaped response." *Management Science* 47, no. 11 (2001): 1476-1487.

Franses, Philip Hans. "Diagnostics, expectations, and endogeneity." *Journal of Marketing Research* 42, no. 1 (2005): 27-29.

Franses, Philip Hans. "On the use of econometric models for policy simulation in marketing." *Journal of Marketing Research* 42, no. 1 (2005): 4-14.

Freimer, Marshall, and Dan Horsky. "Periodic advertising pulsing in a competitive market." *Marketing Science* 31, no. 4 (2012): 637-648.

The Gallup Organization, *Gallup Poll #379,* September 27-October 2, 1946.

Gentzkow, Matthew. "Television and Voter Turnout." *The Quarterly Journal of Economics* (2006): 931-972.

Gentzkow, Matthew, and Jesse M. Shapiro. "Preschool Television Viewing and Adolescent Test Scores: Historical Evidence from the Coleman Study." *The Quarterly Journal of Economics* (2008): 279-323.

Gönül, Füsun, and Kannan Srinivasan. "Estimating the impact of consumer expectations of coupons on purchase behavior: A dynamic structural model." *Marketing Science* 15, no. 3 (1996): 262-279.

Gönül, Füsun, and Meng Ze Shi. "Optimal mailing of catalogs: A new methodology using estimable structural dynamic programming models." *Management Science* 44, no. 9 (1998): 1249-1262.

Goldfarb, Avi. "What is different about online advertising?." Review of Industrial Organization 44, no. 2 (2014): 115-129.

Gidwani, Pradeep P., Arthur Sobol, William DeJong, James M. Perrin, and Steven L. Gortmaker. "Television Viewing and Initiation of Smoking Among Youth." *Pediatrics* 110, no. 3 (2002): 505-508.

Gutschoven, Klaas, and Jan Van den Bulck. "Television Viewing and Smoking Volume in Adolescent Smokers: A Cross-Sectional Study." *Preventive Medicine* 39, no. 6 (2004): 1093-1098.

Hamilton, James L. "The Demand for Cigarettes: Advertising, the Health Scare, and the Cigarette Advertising Ban." *The Review of Economics and Statistics* (1972): 401-411.

Hamilton, J.L., "The effect of cigarette advertising bans on cigarette consumption." *Proceedings of the Third World Conference on Smoking and Health.* DHEW, Washington, DC. (1975): pp. 829–840.

Hancox, Robert J., Barry J. Milne, and Richie Poulton. "Association between child and adolescent television viewing and adult health: a longitudinal birth cohort study." *The Lancet* 364, no. 9430 (2004): 257-262.

Hartmann, Wesley R., and Daniel Klapper. "Super bowl ads." (2014).

Helmer, Richard M., and Johny K. Johansson. "An exposition of the Box-Jenkins transfer function analysis with an application to the advertising-sales relationship." *Journal of Marketing Research* (1977): 227-239.

Herlitz, Agneta, Lars-Göran Nilsson, and Lars Bäckman. "Gender differences in episodic memory." *Memory & Cognition* 25, no. 6 (1997): 801-811.

Hitsch, Günter J. "An empirical model of optimal dynamic product launch and exit under demand uncertainty." *Marketing Science* 25, no. 1 (2006): 25-50.

Hu, Ye, Leonard M. Lodish, Abba M. Krieger, and Babak Hayati. "An Update of Real-World TV Advertising Tests." *Journal of Advertising Research* 49, no. 2 (2009): 201-206.

Iyer, Ganesh, David Soberman, and J. Miguel Villas-Boas. "The targeting of advertising." *Marketing Science* 24, no. 3 (2005): 461-476.

Johnson, Garrett A., Randall A. Lewis, and David Reiley. "Location, location, location: repetition and proximity increase advertising effectiveness." Available at SSRN 2268215 (2014).

Koyck, Leendert Marinus. "Distributed lags and investment analysis." (1954): 5.

Laugesen, M. and Meads, C. "Tobacco Advertising Restrictions, Price, Income and Tobacco Consumption in OECD Countries, 1960-1986," *British Journal of Addiction*, (1991): 1343-54.

Lewit, E.M., D. Coate and M. Grossman "The Effects of Government Regulation on Teenage Smoking", *Journal of Law and Economics,* 24 (1981): 45-569.

Li, Kai, and Timon C. Du. "Building a targeted mobile advertising system for location-based services." *Decision Support Systems* 54, no. 1 (2012): 1-8.

Liaukonyte, Jura. Is Comparative Advertising an Active Ingredient in the Pain Relief Market. Working Paper, Cornell University, 2011.

Liaukonyte, Jura, Thales Teixeira, and Kenneth C. Wilbur. "Television advertising and online shopping." *Marketing Science* 34, no. 3 (2015): 311-330.

Lieberman, Marvin B. "The learning curve, diffusion, and competitive strategy." *Strategic management journal* 8, no. 5 (1987): 441-452.

Linoff, Gordon S., and Michael JA Berry. Data mining techniques: for marketing, sales, and customer relationship management. John Wiley & Sons, 2011.

Lodish, Leonard M., Magid Abraham, Stuart Kalmenson, Jeanne Livelsberger, Beth Lubetkin, Bruce Richardson, and Mary Ellen Stevens. "How TV Advertising Works: A Meta-Analysis of 389 Real World Split Cable TV Advertising Experiments." *Journal of Marketing Research* (1995): 125-139.

Lodish, Leonard M., Magid M. Abraham, Jeanne Livelsberger, Beth Lubetkin, Bruce Richardson, and Mary Ellen Stevens. "A summary of fifty-five in-market experimental estimates of the long-term effect of TV advertising." *Marketing Science* 14, no. 3_supplement (1995): G133-G140.

Luo, Xueming. "Weather and Mobile Promotion." Working Paper, Temple University, 2016.

MacQueen, James. "Some methods for classification and analysis of multivariate observations." In *Proceedings of the fifth Berkeley symposium on mathematical statistics and probability*, vol. 1, no. 14, pp. 281-297. 1967.

Manchanda, Puneet, Peter E. Rossi, and Pradeep K. Chintagunta. "Response modeling with nonrandom marketing-mix variables." *Journal of Marketing Research* 41, no. 4 (2004): 467-478.

McMaster, Gregory S., and W. W. Wilhelm. "Growing degree-days: one equation, two interpretations." *Agricultural and Forest Meteorology* 87, no. 4 (1997): 291-300.

Moulton, Brent R. "An Illustration of a Pitfall in Estimating the Effects of Aggregate Variables on Micro Units." *The Review of Economics and Statistics* (1990): 334-338.

Murthi, B. P. S., and Sumit Sarkar. "The role of the management sciences in research on personalization." *Management Science* 49, no. 10 (2003): 1344-1362.

Naik, Prasad A., Murali K. Mantrala, and Alan G. Sawyer. "Planning media schedules in the presence of dynamic advertising quality." *Marketing Science* 17, no. 3 (1998): 214-235.

Nerlove, Marc, and Kenneth J. Arrow. "Optimal advertising policy under dynamic conditions." *Economica* (1962): 129-142.

Osinga, Ernst C., Peter SH Leeflang, and Jaap E. Wieringa. "Early marketing matters: a time-varying parameter approach to persistence modeling." *Journal of Marketing Research* 47, no. 1 (2010): 173-185.

Palda, Kristian S. "The measurement of cumulative advertising effects." *The Journal of Business* 38 (1964).

Parsons, Leonard J. "The product life cycle and time-varying advertising elasticities." *Journal of Marketing Research* 12, no. 4 (1975): 476-480.

Pollay, Richard W., Sid Siddarth, Michael Siegel, Anne Haddix, Robert K. Merritt, Gary A. Giovino, and Michael P. Eriksen. "The Last Straw? Cigarette Advertising and Realized Market Shares among Youths and Adults, 1979-1993." *The Journal of Marketing* (1996): 1-16.

Pierce, John P., Lora Lee, and Elizabeth A. Gilpin. "Smoking initiation by adolescent girls, 1944 through 1988: an association with targeted advertising." JAMA 271, no. 8 (1994): 608-611.

Roberts, Mark J., Samuelson, Larry. "An Empirical Analysis of Dynamic, Nonprice Competition in an Oligopolistic Industry." *The RAND Journal of Economics* (1988): 200-220.

Rossi, Peter E. "Even the rich can make themselves poor: A critical examination of IV methods in marketing applications." *Marketing Science* 33, no. 5 (2014): 655-672.

Saffer, Henry, and Frank Chaloupka. "The Effect of Tobacco Advertising Bans on Tobacco Consumption." *Journal of Health Economics* 19, no. 6 (2000): 1117-1137.

Sahni, Navdeep S. "Effect of temporal spacing between advertising exposures: evidence from online field experiments." *Quantitative Marketing and Economics* 13, no. 3 (2015): 203-247.

Sargent, James D., Michael L. Beach, Anna M. Adachi-Mejia, Jennifer J. Gibson, Linda T. Titus-Ernstoff, Charles P. Carusi, Susan D. Swain, Todd F. Heatherton, and Madeline A. Dalton. "Exposure to movie smoking: its relation to smoking initiation among US adolescents." *Pediatrics* 116, no. 5 (2005): 1183-1191.

Schneider, Lynne, Benjamin Klein, and Kevin M. Murphy. "Governmental Regulation of Cigarette Health Information." *Journal of Law and Economics* (1981): 575-612.

Schmalensee, Richard. "The Economics of Advertising." (1972).

Seldon, Barry J., and Khosrow Doroodian. "A Simultaneous Model of Cigarette Advertising: Effects on Demand and Industry Response to Public Policy." *The Review of Economics and Statistics* (1989): 673-677.

Sethuraman, Raj, Gerard J. Tellis, and Richard A. Briesch. "How well does advertising work? Generalizations from meta-analysis of brand advertising elasticities." *Journal of Marketing Research* 48, no. 3 (2011): 457-471.

Sen, Anindya, and Tony Wirjanto. "Estimating the impacts of cigarette taxes on youth smoking participation, initiation, and persistence: empirical evidence from Canada." *Health economics* 19, no. 11 (2010): 1264-1280.

Shapiro, Brad. "Positive Spillovers and Free Riding in Advertising of Prescription Pharmaceuticals: The Case of Antidepressants." *Journal of Political Economy*, forthcoming.

Simmons & Associates, Research Inc. for the National Broadcasting Company. "Television's Daytime Profile: Buying Habits and Characteristics of the Audience", 1954.

Sinkinson, Michael, and Amanda Starc. "Ask your doctor? Direct-to-consumer advertising of pharmaceuticals." *No. w21045. National Bureau of Economic Research*, 2015.

Sismeiro, Catarina, Natalie Mizik, and Randolph E. Bucklin. "Modeling coexisting business scenarios with time-series panel data: A dynamics-based segmentation approach." *International Journal of Research in Marketing* 29, no. 2 (2012): 134-147.

Song, Inseong, and Pradeep K. Chintagunta. "A micromodel of new product adoption with heterogeneous and forward-looking consumers: Application to the digital camera category." *Quantitative Marketing and Economics* 1, no. 4 (2003): 371-407.

Spenkuch, Jörg L., and David Toniatti. "Political Advertising and Election Outcomes." Available at SSRN 2613987 (2015).

Stewart, M.J. "The effect on tobacco consumption of advertising bans in OECD countries." *International Journal of Advertising* (1993): 155–180.

Teixeira, Thales S., Michel Wedel, and Rik Pieters. "Moment-to-moment optimal branding in TV commercials: Preventing avoidance by pulsing." *Marketing Science* 29, no. 5 (2010): 783-804.

Tellis, Gerard J., Rajesh K. Chandy, and Pattana Thaivanich. "Which ad works, when, where, and how often? Modeling the effects of direct television advertising." *Journal of Marketing Research* 37, no. 1 (2000): 32-46.

Telser, Lester G. "Advertising and cigarettes." *The Journal of Political Economy* (1962): 471-499.

Television Magazine, "Television Magazine Audience Research," *Television* (1948-1949).

Tuchman, Anna. "Advertising and Demand for Addictive Goods: The Effects of E-Cigarette Advertising." Working Paper, Stanford University, 2016.

U.S. Department of Health and Human Services. "Preventing Tobacco Use Among Young People: A Report of the Surgeon General." (1994) Atlanta, GA: U.S. Department of Health and Human Services, Public Health Service, Centers for Disease Control and Prevention, Office on Smoking and Health.

Vakratsas, Demetrios, Fred M. Feinberg, Frank M. Bass, and Gurumurthy Kalyanaram. "The shape of advertising response functions revisited: A model of dynamic probabilistic thresholds." *Marketing Science* 23, no. 1 (2004): 109-119.

Van Heerde, Harald J., Marnik G. Dekimpe, and William P. Putsis Jr. "Marketing models and the Lucas critique." *Journal of Marketing Research* 42, no. 1 (2005): 15-21.

Vidale, M. L., and H. B. Wolfe. "An operations-research study of sales response to advertising." *Operations research* 5, no. 3 (1957): 370-381.

Villas-Boas, J. Miguel, and Russell S. Winer. "Endogeneity in brand choice models." *Management Science* 45, no. 10 (1999): 1324-1338.

Wellman, Robert J., David B. Sugarman, Joseph R. DiFranza, and Jonathan P. Winickoff. "The extent to which tobacco marketing and tobacco use in films contribute to children's use of tobacco: a meta-analysis." *Archives of pediatrics & adolescent medicine* 160, no. 12 (2006): 1285-1296.

Wilcox, Gary B., and Barry Vacker. "Cigarette advertising and consumption in the United States: 1961–1990." *International Journal of Advertising.* 11, no. 3 (1992): 269-278.

APPENDIX A
CHAPTER 1 APPENDIX

A.1 ESTIMATES FROM A PROPORTIONAL HAZARD MODEL

A common assumption in the hazard literature is the proportional hazard model as suggested by Cox (1972) who expressed the continuous-time hazard as:

$$\lambda_i(t) = \lambda_0(t) exp\left\{z_i(t)'\beta\right\} \tag{A.1}$$

where $\lambda_0(t)$ is the baseline hazard over time, and the exponential term allows for the influence of covariates on the hazard relative to this baseline.

As discussed in Section 1.6.1 the data on smoking ages are best modeled as discrete periods of time. Also, as we saw in Figure 1.7 the likelihood that a person becomes a smoker can change dramatically from one age to the next which suggests that assuming a smooth, parametric distribution for the baseline hazard – as is common for many applications – would be inappropriate here. The Prentice and Gloeckler (1978) model allows each age to have its own contribution to the hazard, estimated as a separate "fixed effect" in a manner similar to equation (1.4) of this paper. This Prentice and Gloeckler model discretizes equation (A.1) by expressing the probability that i does not become a smoker between age t and $t+1$ as:

$$P[T_i \geq t+1 | T_i > t] = \int_t^{t+1} \lambda_0(u) exp\left\{z_i(t)'\beta\right\} du = exp\left[-h_i(t)\right]$$

$$\text{where } h_i(t) = exp\left[z_i(t)'\beta + \gamma(t)\right]$$

T_i is the age that i became a smoker, covariates $z_i(t)$ are assumed to be constant between t and $t+1$, and $\gamma(t)$ is a constant that captures the effect of the baseline between t and $t+1$, specifically $\gamma(t) = log\left\{\int_t^{t+1} \lambda_0(u)du\right\}$.

This model can be estimated using the likelihood given by:

$$L(\gamma(t),\beta) = \prod_{i=1}^{N}\left[[1-exp\{-h_i(t)\}]^{\delta_i}\prod_{t=0}^{k_i-1} exp\{-h_i(t)\}\right]$$

where N is the number of people in the sample, δ_i equals 1 if person i reported being a smoker and 0 if i reported being a non-smoker. For smokers, k_i is the age at which i reports having become a smoker, for non-smokers k_i is the age at which person i was interviewed. The product on the right side from $t=0$ to k_i-1 corresponds to the joint probability that person i did not become a smoker at any age between 0 and k_i-1. If the person was a smoker, the term on the left side raised to δ_i becomes non-trivial and contributes the probability that i became a smoker at age k_i. See Meyer (1990) for derivations of these expressions and a discussion of the use of non-parametric baselines in a hazard models.

Following the linear expression we used in equation (1.4) , we specify that:

$$h_i(t) = exp\{\gamma_{gt} + \tau_{gy} + \phi_{gy}W_d + \beta\mathbf{TV}\} \tag{A.2}$$

where γ_{gt} are age "fixed effects" associated with each gender, which provide the non-parametric estimate of the baseline hazard. The other terms are the same as described in Section 1.6.3.

Table A.1 presents estimates of the coefficients on $\mathbf{TV}_i$ for specifications analogous to those present for the OLS-based estimates in Table 1.2. Essentially the same pattern holds, with significant coefficients immediately following television entry across specifications

for women. The standard errors are slightly larger relative to the estimates than in Table 1.2 so there are fewer significant coefficients.

A.2 Estimating the Increase in the Share of Smokers with the Hazard

This Appendix derives an estimate for the change in the share of smokers as a result of television using estimates of the change in the hazard of becoming a smoker upon television entry. Start with the definition of the hazard, which is:

$$h(t) = \frac{f(t)}{1 - F(t)}$$

where $h(t)$ is the hazard of person i becoming a smoker at time t, $f(t)$ is the probability density of becoming a smoker at time t and $F(t)$ is the cumulative distribution at time t. We can then define the cumulative hazard function as:

$$H(t) = \int_0^t h(u)du = \int_0^t \frac{f(u)}{1 - F(t)}du = -ln\{1 - F(t)\}$$

It then follows that the probability of being a smoker at time t is $F(t) = 1 - exp\{-H(t)\}$. Because nearly all smokers have started by age 36 we evaluate the cumulative hazard at that age. We can then estimate how the share of smokers changed in presence of television:

$$\Delta S_{TV} \equiv F(t \mid TV) - F(t \mid no\ TV) = exp\{-H(t \mid no\ TV)\} - exp\{-H(t \mid TV)\}$$

The assumption in equation A.2 is that the effect of television on the hazard is additive, so we can write:

Table A.1: HAZARD OF SMOKING UPTAKE WHEN TELEVISION ENTERS: PROPORTIONAL HAZARD MODEL

	(1)	(2)	(3)	(4)	(5)
Year Relative to TV Entry: β					
-5	.067	.024	-.005	-.112	.020
	(.053)	(.061)	(.067)	(.196)	(.091)
-4	.070	.014	.033	-.172	.077
	(.070)	(.078)	(.088)	(.209)	(.110)
-3	.015	-.068	-.028	.088	-.047
	(.076)	(.075)	(.088)	(.182)	(.093)
-2	.100	-.014	.029	.047	.035
	(.069)	(.078)	(.088)	(.217)	(.102)
-1	.017	-.114	-.032	.010	-.029
	(.056)	(.067)	(.085)	(.204)	(.098)
0	**.220**	.120	**.195**	.130	**.223**
	(.056)	(.070)	**(.080)**	(.234)	**(.105)**
1	**.258**	**.170**	**.216**	.051	**.271**
	(.053)	**(.070)**	**(.088)**	(.212)	**(.097)**
2	**.151**	.030	.106	.152	.119
	(.064)	(.078)	(.101)	(.170)	(.119)
3	**.148**	.030	.093	-.138	.168
	(.073)	(.077)	(.089)	(.163)	(.097)
4	**.131**	.043	.095	-.163	.179
	(.058)	(.062)	(.076)	(.190)	(.091)
5	.100	-.028	.034	-.304	.132
	(.080)	(.081)	(.072)	(.160)	(.087)
Age x Gender Fixed Effects Baseline: γ_{gt}	X	X	X	X	X
Calendar-Year x Gender Fixed Effects: τ_{gy}		X	X	X	X
DMA Characteristics x Gender xYear: $\phi_{gy}W_d$			X	X	X
Gender	Pooled	Pooled	Pooled	Male	Female
Number of Observations	345,968	345,968	343,068	59,638	283,430
Number of Individuals	22,468	22,468	22,468	5,193	17,275
Number of DMAs	114	114	112	112	112

Note: Estimates of the contribution to the hazard of starting smoking in the years relative to television entry. The specification is a discrete-time hazard model in age that includes age fixed effects, calendar-year fixed effects, and interactions of DMA characteristics (log population and log median income) with the year fixed effects. Estimates are from responses to the 1970 NHIS question "At what age did you start smoking regularly?". Regressions use sample weights provided by the NHIS. The values in parentheses are standard errors. Coefficients with t-statistics of two or higher are in boldface.

$$\Delta S_{TV} = exp\{-H_0\} - exp\left\{-H_0 - \sum_{j=0}^{4} TV_i^j\right\}$$

where H_0 is the baseline cumulative hazard of becoming a without the presence of television.

APPENDIX B

CHAPTER 2 APPENDIX

B.1 POLLEN DATA

B.1.1 Pollen Imputation

Pollen data missing for DMA-weeks between 2006 and 2013 are imputed using a random forest algorithm. The training data are the log of measured pollen levels from the NAB data plus supplemental data from Weather.com covering 2014-2015 across 55 DMAs. Predictors of pollen levels include local weather records (e.g., high, low, and average temperature, humidity, qualitative weather indicators) lags of these values, local length of day, altitude, and so on, all obtained from the National Oceanic and Atmospheric Administration (NOAA); local land-use descriptors focusing on vegetation classifications extracted from raster files available in the National Land Cover Database, 2006 produced by the U.S. Geological Survey; and DMA latitude and longitude.

For this step, I implemented R's ranger package with about 6,500 predictive variables. The algorithm was run with 2,000 trees, and the number of potential features at each split was limited to one third the total number of features. The out-of-sample R^2 for the log of the pollen count is estimated at 0.79.

A summary of the predictive power for different groups of variables is presented in Appendix Table B.2. I use about 6,500 variables used to predict pollen levels, which I have assigned to 12 different groups for descriptive purposes and reported summary statistics for the top 1,000 most predictive variables. Each variable's importance is calculated by the decrease in the residual sum of squares achieved when that variable is used to split a tree. These values are averaged across all trees. I report the average of this value for each

variable group in the second column. The third column shows the number of variables in each group that are in the top 1,000 predictors.

The most predictive variables appear to be those that report the local length of day, perhaps because they capture information both on the time of the year and the DMA's latitude. The largest number of variables in the top 1,000 are related to current and past temperatures, which also tend to be fairly predictive variables. This group includes estimates of "growing degree days," which is a heuristic used to estimate when plants will bloom (McMaster and Wilhelm [1997]). Perhaps unsurprisingly, variables related to air pressure and wind are among the least predictive variables.

B.1.2 Antihistamine Market Definition

The allergy-product market is defined by those products that have a strong correlation with local pollen levels. Figure B.2 shows how these products are clustered away from others in Nielsen's "Cold Remedies - Adult" category. The vertical axis presents the estimated coefficient from regressions of the log of product sales on the log of the local pollen level, controlling for DMA and week fixed effects. In the upper-right corner of this plot, we can see a cluster of products that have both high correlations with local pollen levels and high sales levels. Based on this cluster pattern, the products I consider in this analysis are Benadryl, Claritin, Claritin-D, Zyrtec and Allegra. Zyrtec-D and Allegra-D each represent less than 5% of the total advertising in this industry and are left out, as are some other minor brands.

The RMS sales data are available from 2006 onward. In 2006, only Claritin and Benadryl were available over-the-counter (OTC). Zyrtec was made OTC in 2008, and Allegra in 2011.

Table B.1: POLLEN DATA CONTRIBUTORS

Allegheny General Hospital	Pittsburgh, PA	David Skoner, MD FAAAAI
Allergic Disease and Asthma Center	Greenville, SC	Neil L Kao MD FAAAAI
Allergy & Asthma Associates	Houston, TX	
Allergy & Asthma Associates of Northeastern PA	Erie, PA	Philip E. Gallagher, MD FAAAAI
Allergy & Asthma Care of Waco	Waco, TX	Pramila K. Daftary, MD FAAAAI
Allergy & Asthma Center	Oklahoma City, OK	Martha Tarpay, MD
Allergy & Asthma Center of Georgetown	Austin, TX	Sheila M. Amar, MD FAAAAI, FACAAI
Allergy & Asthma Research Group	Eugene, OR	Kraig W. Jacobson, MD, FAAAAI
Allergy & Asthma Research Group	Erie, PA	Philip E. Gallagher, MD FAAAAI
Allergy Associates of LaCrosse	Onalaska, WI	David Morris, MD
Allergy Clinic of Tulsa	Tulsa, OK	James Love, Jr MD PhD FAAAAI
Allergy Medical Group of the North Area	San Jose, CA	
Allergy Partners of Western North Carolina	Asheville, NC	David Cypcar, MD FAAAAI
Allergy and Asthma Associates of No. California	San Jose, CA	Alan Goldsobel, MD FAAAAI
Allergy and Asthma Center	Waco, TX	N.J. Amar, MD FAAAAI
Allergy, Asthma & Immunology Associates, PC Lincoln	Lincoln, NE	
Allergy, Asthma & Immunology of Rochester	Rochester, NY	Donald W. Pulver, MD FAAAAI
Allergy, Asthma and Immunology	Knoxville, TN	Michael Miller, MD FAAAAI
Allergy, Immunology and Asthma Medical Group	Stockton, CA	Gregory W. Bensch, MD FAAAAI
Asthma & Allergy Associates, PC	Colorado Springs, CO	Robert A. Nathan, MD FAAAAI & Daniel F. Soteres, MD MPH FAAAAI
Asthma & Allergy of Idaho	Twin Falls, ID	Richard Henry, MD
Atlanta Allergy and Asthma Clinic	Marietta, GA	Stanley M. Fineman, MD FAAAAI
Birmingham-Southern College/Alabama Allergy & Asthma Center	Birmingham, AL	Weilly Soong, MD FAAAAI
Carolina Asthma and Allergy Center	Charlotte, NC	John T. Klimas, MD FAAAAI
Children's Mercy Hospital	Kansas City, MO	Jay Portnoy, MD FAAAAI
City of Houston	Houston, TX	Tony Huynh
Clinical Research Institute	Minneapolis, MN	Harold B. Kaiser, MD FAAAAI
Coastal Allergy & Asthma, P.C.	Savannah, GA	Brad H. Goodman, MD & Bruce D. Finkel, MD
Colorado Allergy & Asthma Centers, PC	Centennial, CO	Leon S. Greos, MD FAAAAI
Division of Air Quality, DNREC, State of Delaware	New Castle, DE	Michael McDowell
Dr. Golden and Dr. Matz, LLC	Baltimore, MD	Jonathon Matz, MD FAAAAI & David Golden, MD
Dr. Joseph Leija	Melrose Park, IL	Joseph G. Leija, MD FAAAAI
Family Allergy & Asthma	Louisville, KY	James L. Sublett, MD FAAAAI
Fordham College at Lincoln Center	New York, NY	Guy Robinson, PhD
Fred Lewis, MD FAAAAI	Olean, NY	Fred Lewis, MD FAAAAI
Hedberg Allergy & Asthma Center	Rogers, AR	Curtis L. Hedberg, MD FAAAAI
Intermountain Allergy & Asthma Clinic	Draper, UT	Duane J. Harris, MD FAAAAI
Little Rock Allergy & Asthma Clinic	Little Rock, AR	Karl V Sitz, MD
Northeast Georgia Research Center, LLC	Gainesville, GA	John A. Yarbrough, MD
Northwest Asthma & Allergy Center	Seattle, WA	Frank Virant, MD FAAAAI
OK Allergy Asthma Clinic, Inc.	Oklahoma City, OK	Warren V. Filley, MD FAAAAI
RAPCA	Dayton, OH	Andy Roth
Scott & White Clinic	College Station, TX	David R. Weldon, MD FAAAAI, FACAAI
Springfield -Greene County Health Department	Springfield, MO	Ms. Rhizza Adams
St. Louis County Health Department	St. Louis, MO	
Sylvana Research Associates	San Antonio, TX	Paul Ratner, MD MBA FAAAAI
The Asthma and Allergy Center, PC	Bellevue, NE	Linda B. Ford, MD FAAAAI
The William Storms Allergy Clinic	Colorado Springs, CO	Mathew S. Bowdish, MD FAAAAI
Theodore J. Chu, M.D.	San Jose, CA	Theodore Chu, MD FAAAAI
UW Medical School	Madison, WI	Robert Bush, MD FAAAAI
University of Tulsa	Tulsa, OK	
Walter Reed Army Medical Ctr.	Washington, DC	
Wilford Hall Ambulatory Surgical Center	San Antonio, TX	Mr. Robert Gomez

Table B.2: POLLEN PREDICTORS

Variable Group	Importance (Avg.)	Number of Variables
Day Length	.00407	87
Time of Year	.00135	2
Longitude	.00093	1
Temperature	.00082	324
Land Cover	.00078	96
Latitude	.00067	1
Altitude	.00031	1
Air Moisture	.00022	95
Precipitation	.00019	119
Air Pressure	.00015	29
Wind	.00015	217
Other	.00015	28

Notes: Summary statistics for categories with pollen predictors that are among the top 1000 most predictive out of about 6,500. Pollen is predicted in order to impute missing pollen data. The "Variable Group" definitions are for descriptive purposes only. The "Importance" for each variable is estimated by the decrease in the residual sum of squares achieved when that variable is used during a split in a given tree. These values are averaged across all trees in the forest for each variable. I have reported the averages across these importance estimates for each "Variable Group." I also report the number of variables from each group in the top 1,000 for each category.

B.1.3 Regional Differences in Pollen Levels

Whereas Figure 2.3 shows important differences in pollen levels by census region, finer differences in pollen levels exist for smaller geographic units. To gain a better understanding of which regions have similar pollen patterns, I have used a k-means algorithm to define groups of DMAs with similar patterns across time (MacQueen [1967]). The algorithm's results with 10 groups are plotted in Figure B.1.3, which shows that to a large extent, pollen patterns differ by latitude.

B.2 ROBUSTNESS TESTS

B.2.1 Advertising Carryover

This section estimates the response to past advertising both non-parametrically and using the standard geometric-decay assumption for the antihistamine industry. This industry is well suited to such estimates because the random mismatch between the national advertising schedule and local pollen seasons provides a means of separating the effect of past advertising from recent changes in demand. Estimates from this analysis suggest advertising can affect sales several weeks after it has aired.

Consider a distributed-lag model expressed as

$$y_{it} = \alpha_i + \xi X_{it} + \gamma D_{dt} + \sum_{\tau=T^-}^{0} b_\tau a_{d(t+\tau)} + \varepsilon_{it}, \tag{B.1}$$

where y_{it} is the log of antihistamine of sales at store i in week t. Store fixed effects are represented by α_i, and X_{it} contains controls for the log of average price across antihistamine products, month fixed effects, year fixed effects, and a dummy for whether an antihistamine product was on feature. D_{dt} represents local levels of demand measured using the log of pollen for DMA d. Each of the $a_{i(t+\tau)}$ terms represent the log of advertising received in

DMA d at time $(t+\tau)$. Accordingly, each b_τ estimates the advertising elasticity contributed by advertising τ weeks in the past as in a distributed lag model (Bass and Clarke [1972]). The identifying assumption for this expression is that after controlling for D_{dt} and other covariates, the variation in current and past advertising is random.

Estimates of b_τ are presented in Figure B.4. Generally, the effect of advertising appears to decline gradually as we move further into the past, as one might expect. However, the effect of same-week advertising is small relative to advertising that occurred one to three weeks ago, which suggests the effect is not immediate. A potential explanation could be that consumers often wait until their next trip to the store before the effect of the advertising they saw is realized as a sale. Excluding the same-period coefficient, the effect of advertising may be consistent with a constant-carryover assumption, but the precision is insufficient to test whether this or some other model of carryover is more consistent with the data. Placebo tests for these estimates, presented in Appendix B.2.2, are consistent with interpreting the b_τ as the true advertising effects. They also illustrate the importance of controlling for local demand levels, D_{dt}.

An expression with fewer parameters can be obtained by imposing the constant-carryover assumption:

$$y_{it} = \alpha_i + \xi X_{it} + \gamma D_{dt} + \gamma_0 \sum_{\tau=T^-}^{0} \lambda^\tau a_{d(t+\tau)} + \varepsilon_{it}, \tag{B.2}$$

where λ is the rate of geometric decay. Based on the estimates from the distributed-lag model in (B.1), I set $T^- = 10$, which produces an estimate of $\lambda = 0.84$ using non-linear least squares. I use this estimate to calculate the advertising stock for each period as $A_{it} = \sum_{\tau=T^-}^{0} \lambda^\tau a_{i(t+\tau)}$, which reduces advertising's impact on sales to a single variable. Note that this model imposes an assumption on the way advertising interacts with organic

demand: only current-period organic demand affects the response to current advertising stock.

B.2.2 Placebo Test for the Distributed-Lag of Advertising

In this section, I generalize the distributed-lag model given in equation (B.1) in order to perform a placebo test. We know future advertising can have no effect on sales, because consumers have not yet been exposed to it. By altering the original model to estimate the influence of future advertising on current sales, we can verify that our estimates are consistent with this knowledge. If the model were to produce estimates suggesting future advertising has similar effects to past advertising, it would suggest our model is incorrect and the "advertising" estimates are in fact being driven by some unobservable. Equation (B.1) is generalized as

$$y_{it} = \alpha_i + \xi X_{it} + \gamma D_{it} + \sum_{\tau=T^-}^{T^+} b_\tau a_{i(t+\tau)} + \varepsilon_{it} \tag{B.3}$$

where the new T^+ term allows us to estimate the effect of future advertising.

Estimates from this model are presented in Appendix Figure B.5 with $T^+ = 10$. For non-positive time values, the estimates are similar to those presented in Appendix Figure B.4. Although we expect the estimates for future advertising to all equal to zero, a number of them show significance. Nevertheless, they are much closer to zero than estimated for weeks -1 to -3.

B.3 ESTIMATION DETAILS

B.3.1 Category-level Estimates

Details from the category-level regressions can be found in Appendix Tables B.3 to B.6.

Table B.3: CATEGORY REGRESSIONS: ANTIHISTAMINES

Advertising Elasticity	0.345	0.288	0.134	0.033	0.171
	(0.021)	(0.025)	(0.012)	(0.031)	(0.026)
Price Coefficient	-0.995	-0.959	-1.022	-0.934	-1.024
	(0.012)	(0.011)	(0.011)	(0.015)	(0.011)
Organic Demand Coef.	.000	.000	0.278	.000	0.266
	(.000)	(.000)	(0.011)	(.000)	(0.013)
Number of Obs	12,039,907	12,039,907	12,036,767	4,342,419	12,036,767
Number of Stores	39,371	39,371	39,371	8,117	39,371
Number of Clusters	204	204	204	198	204
R^2	0.883	0.891	0.895	0.861	0.895
F-stat (excl. var.)					79
Method	OLS: naive	OLS: moFE	OLS: OD	Border	IV: OD

B.3.2 Brand-level Estimates

Details from the brand-level regressions can be found in Appendix Tables B.7 to B.10.

Table B.4: CATEGORY REGRESSIONS: LIP MEDICATION

Advertising Elasticity	0.136	0.066	-0.033	-0.010	0.053
	(0.003)	(0.003)	(0.004)	(0.009)	(0.006)
Price Coefficient	-0.361	-0.296	-0.320	-0.308	-0.323
	(0.012)	(0.011)	(0.013)	(0.027)	(0.012)
Organic Demand Coef.	.000	.000	-0.025	.000	-0.018
	(.000)	(.000)	(0.000)	(.000)	(0.001)
Number of Obs	13,583,519	13,583,519	13,472,781	4,385,315	13,472,781
Number of Stores	39,643	39,643	39,385	8,118	39,385
Number of Clusters	204	204	204	198	204
R^2	0.770	0.815	0.807	0.856	0.802
F-stat (excl. var.)					3,168
Method	OLS: naive	OLS: moFE	OLS: OD	Border	IV: OD

Table B.5: CATEGORY REGRESSIONS: MOISTURIZERS

Advertising Elasticity	0.216	0.051	0.113	-0.015	0.015
	(0.006)	(0.006)	(0.004)	(0.014)	(0.023)
Price Coefficient	-0.690	-0.684	-0.639	-0.651	-0.629
	(0.005)	(0.005)	(0.004)	(0.010)	(0.005)
Organic Demand Coef.	.000	.000	-0.006	.000	-0.007
	(.000)	(.000)	(0.000)	(.000)	(0.000)
Number of Obs	6,960,492	6,960,492	6,901,143	2,528,880	6,901,143
Number of Stores	19,567	19,567	19,432	4,689	19,432
Number of Clusters	204	204	204	198	204
R^2	0.908	0.911	0.909	0.927	0.909
F-stat (excl. var.)					1,398
Method	OLS: naive	OLS: moFE	OLS: OD	Border	IV: OD

Table B.6: CATEGORY REGRESSIONS: SUNSCREEN

Advertising Elasticity	0.201	0.113	-0.083	-0.014	0.154
	(0.007)	(0.008)	(0.010)	(0.009)	(0.050)
Price Coefficient	-1.005	-0.741	-0.948	-0.715	-0.905
	(0.025)	(0.011)	(0.021)	(0.015)	(0.020)
Organic Demand Coef.	.000	.000	0.062	.000	0.033
	(.000)	(.000)	(0.002)	(.000)	(0.006)
Number of Obs	7,640,796	7,640,796	7,595,404	3,255,893	7,595,404
Number of Stores	37,453	37,453	37,413	8,112	37,413
Number of Clusters	204	204	204	198	204
R^2	0.459	0.759	0.588	0.852	0.538
F-stat (excl. var.)					492
Method	OLS: naive	OLS: moFE	OLS: OD	Border	IV: OD

Table B.7: ANTIHISTAMINE BRAND REGRESSIONS: ALLEGRA

Advertising Elasticity	0.071	0.110	0.018	0.124
	(0.007)	(0.013)	(0.051)	(0.026)
Competitive Ad. Elast.	0.061	0.157	0.097	-0.015
	(0.015)	(0.020)	(0.048)	(0.036)
Price Coefficient	-1.852	-1.861	-2.140	-1.849
	(0.037)	(0.037)	(0.055)	(0.036)
Organic Demand Coef.	0.291	0.178	.000	0.035
	(0.011)	(0.085)	(.000)	(0.052)
Number of Obs	2,986,671	3,031,547	1,180,679	3,031,547
Number of Stores	34,525	34,533	7,920	34,533
Number of Clusters	204	204	198	204
R^2	0.711	0.694	0.763	0.712
F-stat (excl. var.)				3,291
Method	OLS	OLS: genus	Border	IV: genus

Table B.8: ANTIHISTAMINE BRAND REGRESSIONS: CLARITIN

Advertising Elasticity	0.082	0.104	0.014	-0.078
	(0.009)	(0.014)	(0.021)	(0.052)
Competitive Ad. Elast.	0.064	0.122	-0.017	0.236
	(0.008)	(0.013)	(0.013)	(0.043)
Price Coefficient	-1.843	-1.882	-2.208	-1.840
	(0.035)	(0.037)	(0.038)	(0.035)
Organic Demand Coef.	0.332	0.295	.000	0.152
	(0.013)	(0.089)	(.000)	(0.052)
Number of Obs	9,055,713	9,147,155	3,745,082	9,147,155
Number of Stores	38,487	38,497	7,920	38,497
Number of Clusters	204	204	198	204
R^2	0.686	0.664	0.750	0.682
F-stat (excl. var.)				1,929
Method	OLS	OLS: genus	Border	IV: genus

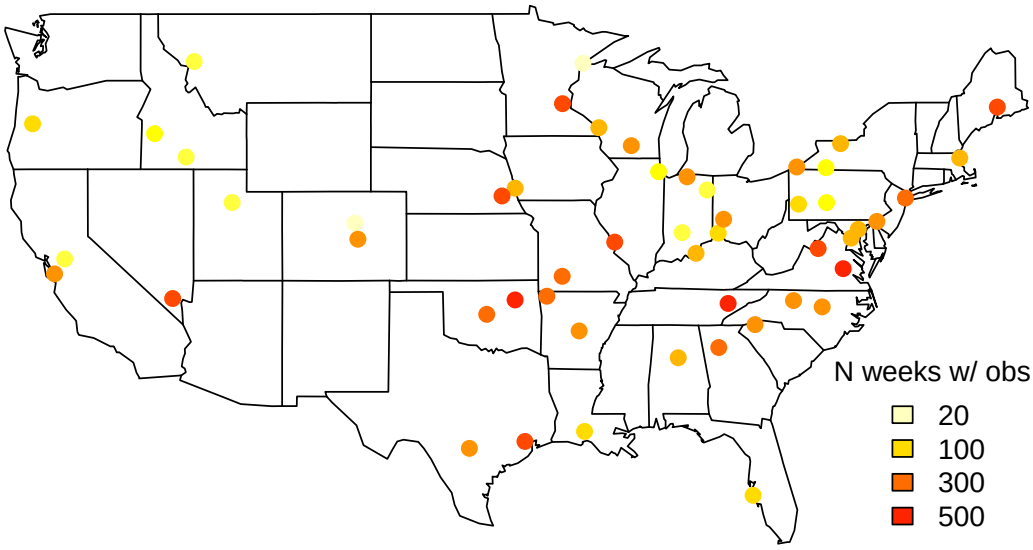
Table B.9: ANTIHISTAMINE BRAND REGRESSIONS: CLARITIN-D

Advertising Elasticity	0.026	0.033	0.038	0.081
	(0.004)	(0.005)	(0.018)	(0.011)
Competitive Ad. Elast.	0.048	0.095	-0.003	0.028
	(0.010)	(0.016)	(0.032)	(0.016)
Price Coefficient	-1.347	-1.359	-1.193	-1.345
	(0.079)	(0.080)	(0.121)	(0.079)
Organic Demand Coef.	0.223	0.181	.000	0.093
	(0.012)	(0.069)	(.000)	(0.041)
Number of Obs	6,865,146	6,933,826	2,471,947	6,933,826
Number of Stores	25,855	25,859	7,920	25,859
Number of Clusters	204	204	196	204
R^2	0.552	0.536	0.627	0.551
F-stat (excl. var.)				26,641
Method	OLS	OLS: genus	Border	IV: genus

Table B.10: ANTIHISTAMINE BRAND REGRESSIONS: ZYRTEC

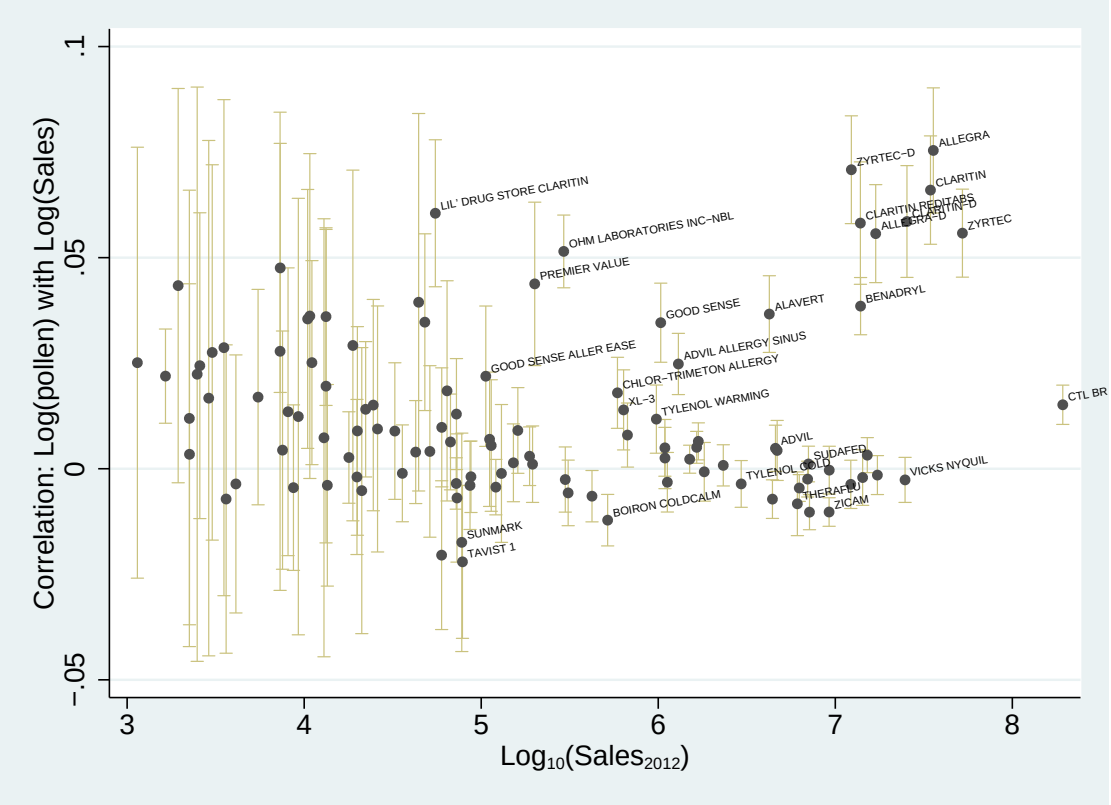
Advertising Elasticity	0.025	0.054	0.018	0.075
	(0.002)	(0.003)	(0.013)	(0.019)
Competitive Ad. Elast.	0.068	0.087	-0.002	-0.031
	(0.010)	(0.018)	(0.028)	(0.043)
Price Coefficient	-1.865	-1.874	-2.202	-1.857
	(0.037)	(0.037)	(0.070)	(0.037)
Organic Demand Coef.	0.298	0.293	.000	0.105
	(0.011)	(0.081)	(.000)	(0.042)
Number of Obs	6,989,902	7,082,627	2,751,969	7,082,627
Number of Stores	37,558	37,567	7,920	37,567
Number of Clusters	204	204	198	204
R^2	0.761	0.745	0.791	0.760
F-stat (excl. var.)				2,088
Method	OLS	OLS: genus	Border	IV: genus

Figure B.1. LOCATIONS OF NAB POLLEN STATIONS



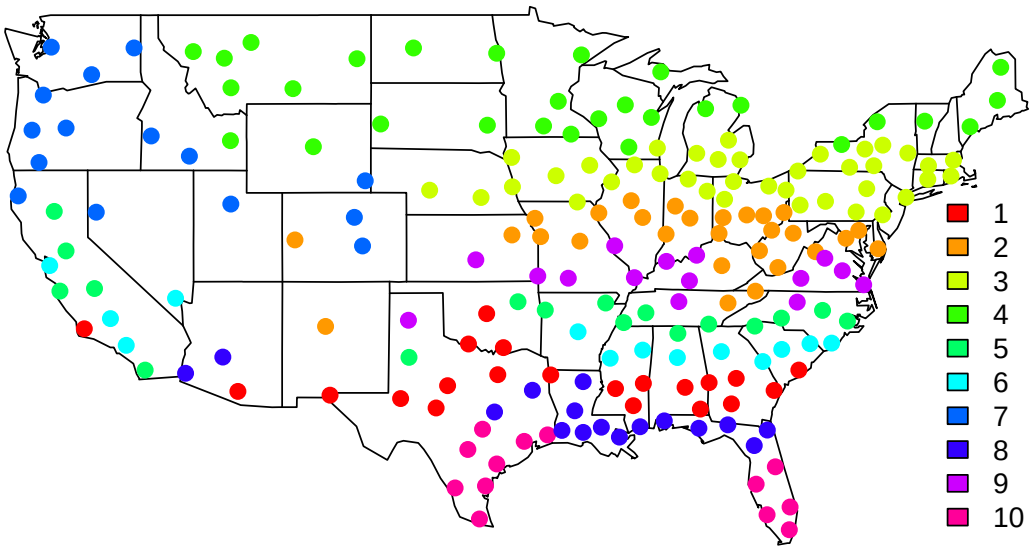
Notes: The plot shows the location of the pollen stations and the colors indicate the number of weeks with at least one observation made for each DMA with a station.

Figure B.2. THE ALLERGY-MARKET PRODUCTS



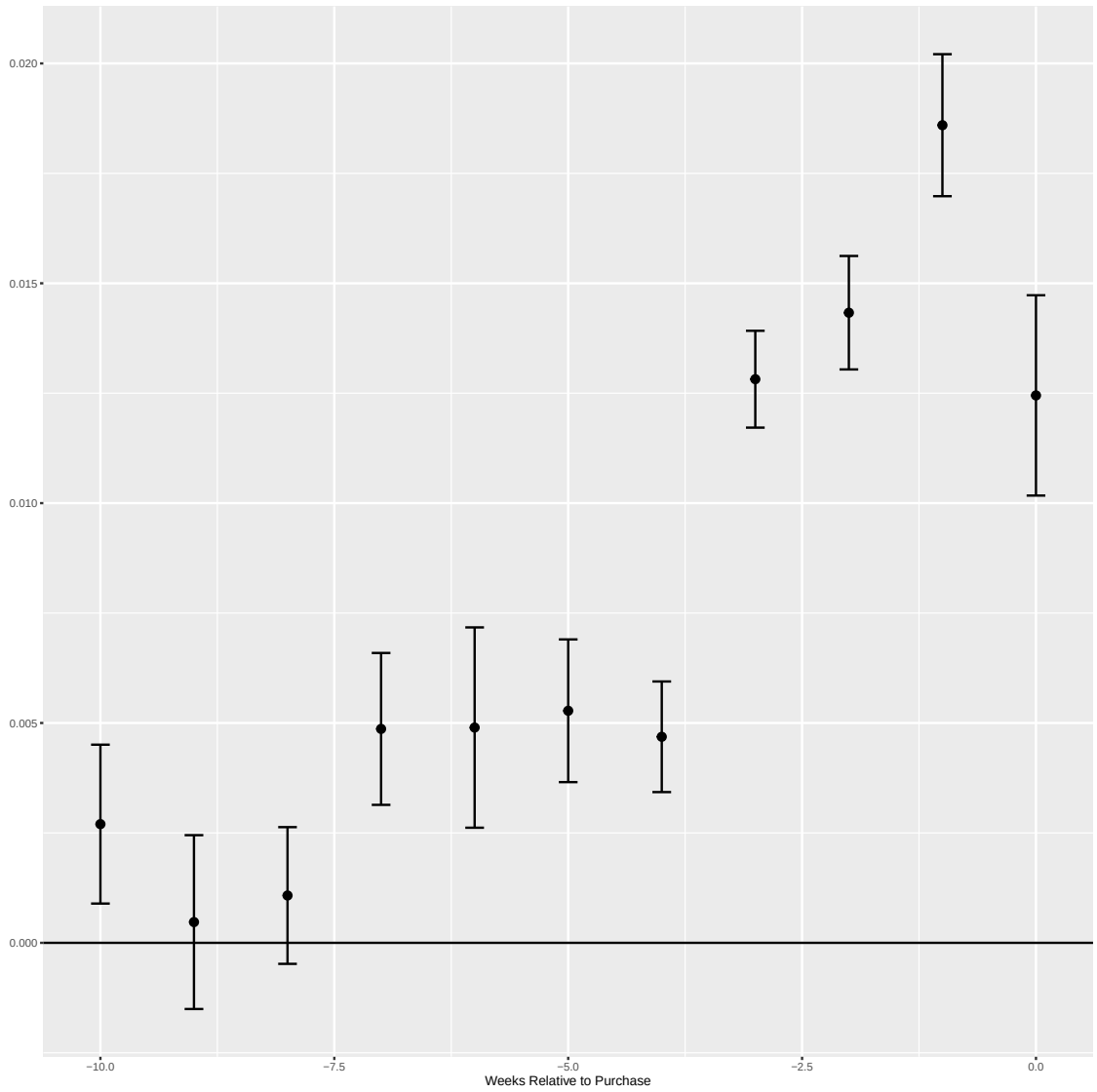
Notes: The figure presents a plot of products from Nielsen's "Cold Remedies - Adult" category. The vertical axis represents the coefficient estimates of product-wise regressions of the log of sales on the log of pollen with fixed effects for DMA and week. Error bars represent ± 2 times the robust standard error estimates. The log of total Nielsen-recorded sales is on the horizontal axis. All data are for the year 2012 only.

Figure B.3. DMAs WITH SIMILAR POLLEN PATTERNS



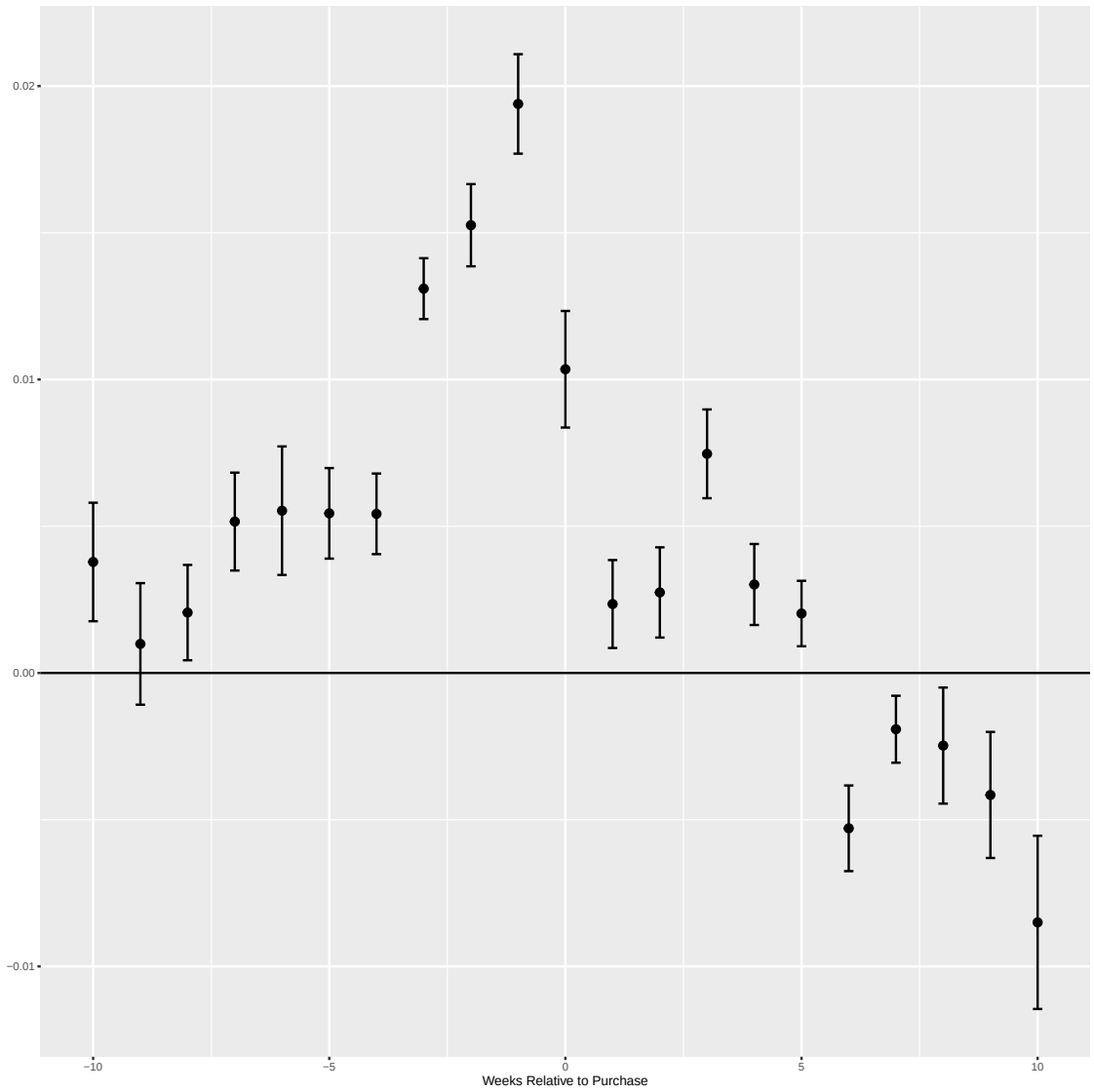
Notes: Groups of DMAs with correlated pollen levels across time are indicated by color. The groups are defined using a k-means clustering algorithm with 10 groups. Each plotted point is a DMA.

Figure B.4. ADVERTISING CARRYOVER



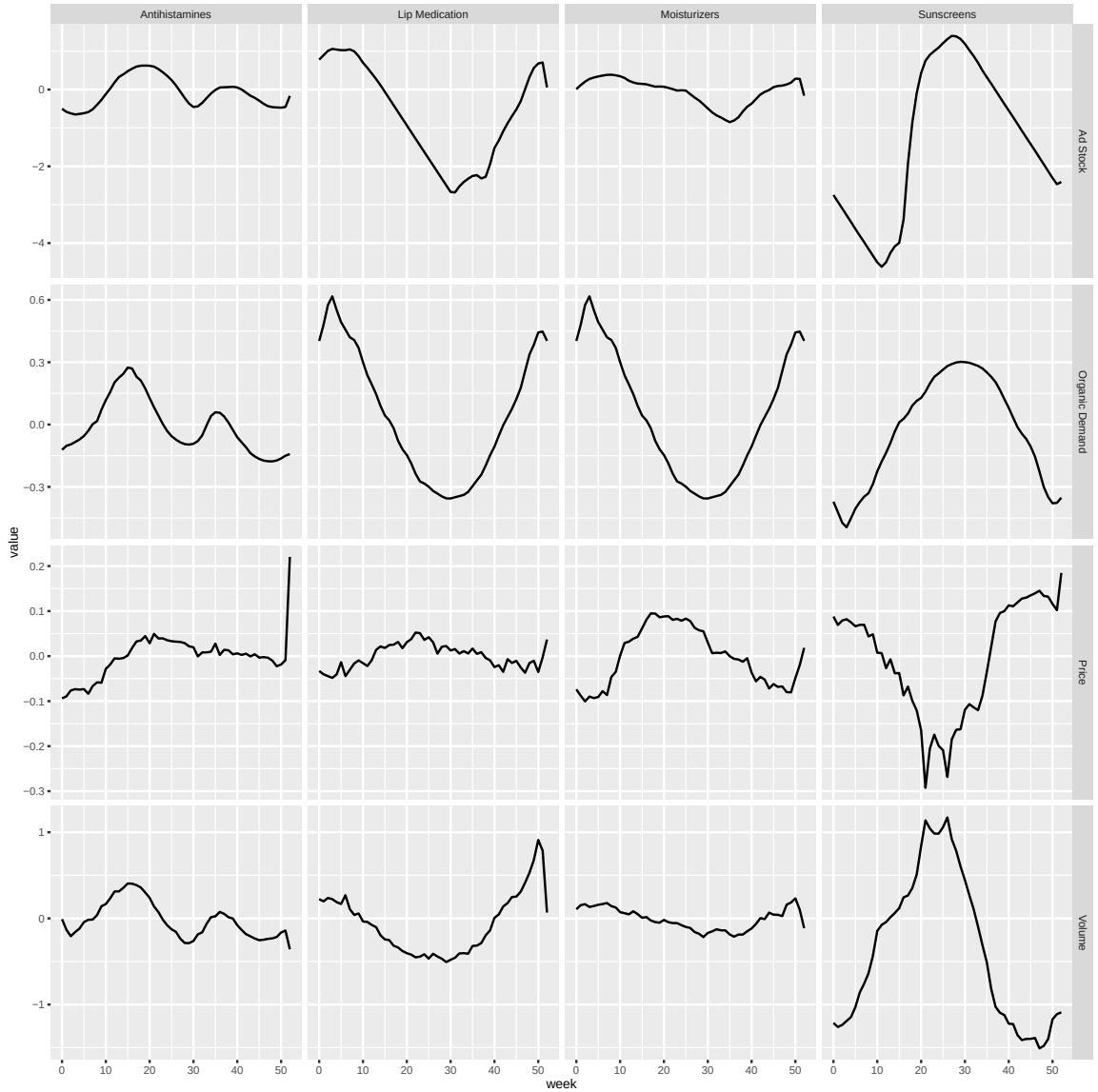
Notes: Estimates from a distributed-lag model of advertising effects. Each coefficient reports the advertising elasticity associated with advertising that occurred the indicated number of weeks in the past. The regression is run at the store-week level and controls for the log of prices, the log of the local pollen count, plus store fixed effects. Standard errors (SE) are clustered by DMA, and the error bars indicate $\pm 2 * SE$ for each estimate.

Figure B.5. ADVERTISING CARRYOVER PLACEBO TEST



Notes: Estimates from a distributed-lag model of advertising effects. Each coefficient reports the advertising elasticity associated with advertising that occurred the indicated number of weeks in the past. The regression is run at the store-week level and controls for the log of prices, the log of the local pollen count, plus store fixed effects. Standard errors (SE) are clustered by DMA, and the error bars indicate $\pm 2 * SE$ for each estimate.

Figure B.6. SUMMARY STATISTICS FOR KEY CATEGORIES



Notes: National averages across each week of the year. All y-axis values are the percent deviations from the annual average. Organic demand for Antihistamines is (the log of) total pollen count; for Lip Medications and Moisturizers, it is (opposite of) the dew point; and for Sunscreens, it is the average daily temperature.