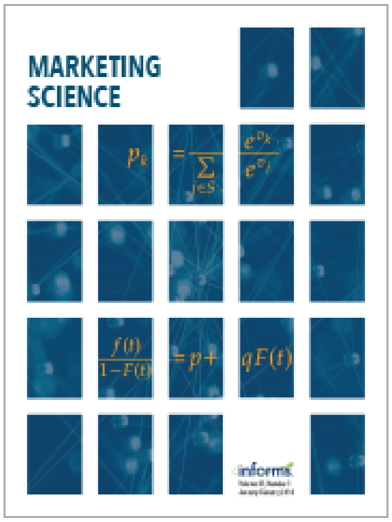


Marketing Science

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

Spillovers from Mass Advertising: An Identification Strategy

Michael Thomas

MARKETING SCIENCE
Vol. 39, No. 4, July–August 2020, pp. 807–826
ISSN 0732-2399 (print), ISSN 1526-548X (online)
http://pubsonline.informs.org/journal/mksc

Spillovers from Mass Advertising: An Identification Strategy

Michael Thomas[a]

[a] Santa Clara University, Santa Clara, California 95053
Contact: mwthomas@scu.edu, https://orcid.org/0000-0002-2229-3778 (MT)

Received: December 25, 2017
Revised: June 17, 2019; October 14, 2019
Accepted: October 29, 2019
Published Online in Articles in Advance: July 15, 2020

https://doi.org/10.1287/mksc.2019.1217

Abstract. Increasingly, firms have the ability to make high-quality, microlevel predictions of demand for their products, which improves their ability to target advertising. In spite of this, firms may choose to target advertising at a higher level of aggregation than their predictions allow to benefit from the significant discounts that often accompany mass advertising purchases. We argue that firms making such a choice generate "advertising spillovers" that are quasi-random and can be used to identify the response to advertising. These advertising spillovers occur when local levels of advertising are higher or lower than locally optimal because of the influence of other markets or individuals on the mass advertising decision. We formalize the supply-side conditions that incentivize firms to generate these spillovers as part of their optimization strategy, present an empirical strategy for exploiting these conditions, and apply the strategy to multiple product categories and brands. Estimates from this "spillover strategy" agree with recent literature that suggests many standard approaches to estimating the response to advertising may produce biased results because of unobservables; our estimates also suggest that some recent empirical strategies, such as the DMA-border strategy, can produce biased estimates for seasonal products.

History: Avi Goldfarb served as the senior editor and Hema Yoganarasimhan served as associate editor for this article.
Supplemental Material: The online appendix and data/replication files are available at https://doi.org/10.1287/mksc.2019.1217.

Keywords: natural experiments • advertising • econometrics • endogeneity • instrumental variable methods • panel data

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 or, in some cases, the firm itself (Nair et al. 2017). To address this difficulty, recent work by Shapiro (2018) and Sinkinson and Starc (2019) 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 illustrate the potential for biased advertising estimates if the researcher is not careful to rely solely on quasi-random variation in advertising levels. However, because these methods rely on local differences in advertising levels, they may be rendered ineffective for product categories that rely heavily on national advertising.

To address this limitation, this paper proposes a method that is more effective for products that rely on mass advertising, such as national television. This strategy for identifying advertising effects is motivated by recent improvements in firms' predictive capacity resulting from the availability of high-quality data, improved prediction algorithms, and increased computational power. Using these resources, firms can predict, for example, how demand for their products will change in response to a variety of factors at very low levels of aggregation. Such predictions of the changes in "organic demand" allow firms to target their advertising at groups that may be especially receptive to their message over specific time periods. However, because mass advertising often receives significant price discounts, firms may not always fully exploit the precision of their demand predictions. Instead, firms may find that advertising at higher levels of aggregation than their predictions allow optimizes their total profit despite knowledge that some identifiable groups of consumers receive too much or too little advertising as a result. Under such conditions, each group of consumers experiences a "spillover" from the advertising needs of the other groups: their advertising experience is partially determined by the organic demand levels of other groups. In this paper, we argue that these spillovers present an 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, such as designated marketing area (DMA) boundaries. We offer a formal presentation of the supply-side incentives that give rise to these advertising spillovers, present a means of exploiting them with an empirical model, and apply the identification strategy to a variety of product categories using a variety of data sources.

To provide an intuitive application of this strategy, we analyze advertising for antihistamines. Antihistamines treat the symptoms of a variety of allergies; chief among them are pollen allergies, which are highly seasonal. The "organic" changes in demand for these medications are well known to pharmaceutical firms, which target their advertising to match the high-demand spring and fall allergy seasons with national advertising. As a side effect of this national advertising, many cities receive more or less advertising than is optimal. 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 determine 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 alter the level of the organic demand for their product. The spillover strategy presented in this paper applies to a product if (1) forecasted changes in organic demand influence the marginal response of advertising and, therefore, firms' advertising decisions; (2) the changes in organic demand are not uniform across all observable units (e.g., markets or individuals); (3) the advertising levels are not perfectly tailored to each unit so that each unit's organic demand levels affect the advertising levels in the others; and (4) the researcher can measure at least some of the local changes in organic demand that influence the mass advertising decision. Opportunities to meet these requirements are improving as large data sets with observations at small levels of aggregation are increasingly available. This paper demonstrates the use of the spillover strategy for the antihistamine, sunscreen, and fire log categories because these categories offer nonproprietary measures of organic demand. However, we believe that data-rich firms with proprietary measures of organic demand are especially well positioned to make use of this strategy.

Like many other studies, we find evidence suggesting researchers may obtain significantly biased estimates in the absence of quasi-random identifying variation. Such evidence for prescription medications is presented by Sinkinson and Starc (2019) and by Shapiro (2018), the latter of which proposes a "border strategy" that exploits differences in advertising levels across DMA boundaries. Also using the border strategy, Shapiro et al. (2019) estimate advertising responses for hundreds of consumer packaged goods and recover a distribution of values smaller than generally reported in the literature; they suggest publication bias as a potential culprit. However, more recent work suggests limitations to the border strategy: Li et al. (2019) find evidence that the advertising responses estimated from border counties may not generalize to urban areas for political advertisements. Additionally, our study finds evidence of biased estimates for "unobserved" seasonal changes for the border strategy, among other strategies. Finally, experimental evidence for online advertisements leaves no doubt that existing observational methods can produce rather biased results (Lewis et al. 2011, Blake et al. 2015, Gordon et al. 2019).

The remainder of this paper is organized as follows. In the next section, we review this paper's contribution to a literature that uses decisions made at high levels of aggregation to isolate quasi-random variation for identification. Next, Section 3 motivates the spillover strategy with a supply-side model that, when optimized, generates opportunities for identification. This section also provides estimating equations, discusses potential instrument definitions, conditions for consistent estimates, and provides a simulation to demonstrate the method's utility. Section 4 applies the method to the antihistamine category at both the category and brand levels. Section 5 applies the method to other product categories and explores the robustness of our supply-side assumptions. Section 6 concludes.

2. Literature Review

This paper contributes to a literature that exploits events taking place at high levels of aggregation for the purpose of identification. This begins with Hausman (1996), who proposes instrumenting for local prices with prices in other markets that serve as a proxy for aggregate supply shocks, an approach also implemented in Nevo (2001). More similar to our study, in an investigation of vertical integration's effect on prices, Hastings (2004) observes that a merger in the retail gasoline industry is likely independent of the concentration of stations in individual markets affected by the merger. This allows her to identify the effect of vertical integration on prices.

Perhaps most conceptually similar to our paper are strategies that exploit how a given consumer's product offerings are affected by the preferences of others. George and Waldfogel (2003) introduce the notion of "preference externalities" whereby the preferences of one group of consumers affect the newspaper content delivered to others, an idea used by Gentzkow and Shapiro (2010) to estimate demand for slant in newspaper content. By analogy, our approach could be viewed as exploiting "advertising externalities." More specifically, Gentzkow and Shapiro (2010) estimate demand for slant in newspaper content by measuring cross-sectional differences between county-level political orientation and (multicounty) newspaper-market averages that determine the profit-maximizing level of slant. Identification comes from assuming that the amount of newspaper slant supplied to each county is exogenously determined by the preferences of other counties in the market. Alternatively, to estimate advertising responses in this paper, we exploit both cross-sectional and time-series variation in the difference between local levels of organic demand and the nationally forecasted level of organic demand that affects the profit-maximizing level of advertising. Identification comes from assuming that forecasted organic demand conditions in other markets exogenously affect the amount of advertising received locally. In each of these instances of "Waldfogel IVs," exposure intensity is constant, but differences in local conditions alter a treatment of interest.

Additionally, Waldfogel IVs can exploit conditions that vary local exposure intensity. For example, Hartmann and Klapper (2018) and Stephens-Davidowitz et al. (2017) observe that preferences for watching the Super Bowl are driven by regional allegiances that are likely independent of preferences for products advertised because the participating teams are not known when advertising is purchased. Such conditions allow them to estimate the effect of Super Bowl advertising on sales.

Numerous other strategies have been employed to estimate advertising effects. Shapiro (2018) makes use of the discontinuity that occurs at DMA boundaries, observing that adjacent counties on either side of the boundary receive slightly different levels of local television advertising. More recently, Li et al. (2019) provide a framework for unifying the border strategy with Waldfogel IVs and provide empirical comparisons. Sinkinson and Starc (2019) propose supply-side instruments that result from state-level political advertising, and Moshary et al. (2019) characterize the empirical conditions for which this strategy is best suited.

Another collection of studies has made use of the exact timing of television advertising for identification. For example, Tellis et al. (2000) obtain estimates from the spike in customer calls that immediately follows an ad for a toll-free referral service. Similarly, Joo et al. (2014) find significant increases in Google searches for branded products following a television advertisement, and Liaukonyte et al. (2015) also find effects on shopping and purchase behavior. Typically these time-based approaches capture advertising responses that occur within minutes or hours of an advertisement but are less suited to measuring in-store purchase responses or requests for prescription medication.

Finally, a handful of papers have grounded the literature with experimental estimates of advertising responses for both television (Lodish et al. 1995) and online (Lewis et al. 2011, Blake et al. 2015, Kalyanam et al. 2018, Gordon et al. 2019).

3. Method

This section begins by presenting a supply-side model to understand the conditions in which firms maximize profits with mass advertising despite knowledge of local inefficiencies. Motivated by the resulting supply-side behavior, we present a basic demand model that allows us to identify advertising effects. We then discuss possible definitions of the instruments that result from advertising spillovers, discuss the market and data conditions under which the method produces consistent estimates of advertising, and demonstrate its use on simulated data. A summary of the key variables used in the models is provided in Table 1.

3.1. Supply-Side Model

We assume a supply-side model in which firms' sales are jointly produced by the goodwill stock that results from advertising and organic demand levels. Organic demand captures shifts in demand that are not under the control of the firm and—essential to the spillover strategy—also shifts the marginal response of advertising. To optimize profits, firms choose between buying national advertising at a discount or paying a premium for local advertising to tailor the advertising intensity to a particular market. Specifically, we assume profits are given by

$$\pi = \sum_{t=1}^{T} \sum_{j=1}^{N} s_j \left[m D_{jt}^{\gamma} G_{jt}^{\theta} - c_1 A_{jt} \right] - c_0 A_{0t}, \tag{1}$$

where A_{jt} is the advertising intensity per person in market j at time t, and A_{0t} is national advertising intensity. Goodwill stock, G_{jt}, is the combination of local and national advertising, each of which decay geometrically over time such that $G_{jt} = 1 + \sum_{\tau=0}^{L} \lambda^{\tau} \times (A_{j(t-\tau)} + A_{0(t-\tau)})$. We assume goodwill stock from local and national advertising are perfect substitutes but come at costs c_1 and c_0, respectively, such that the national advertising is cheaper per person (i.e., $c_1 > c_0$).

Table 1. Key Variables Summary

Variable	Description	Related variables
D_{jt}	Realized organic demand in market j during period t. Organic demand variables affect sales but are outside the control of the firm. $D_{jt} = (\prod_m D_{jt}^m)\xi_{jt}$	D_{jt}^m = observed component m of organic demand, D_{jt}; ξ_{jt} = unobserved component of organic demand, D_{jt}; $\mathbf{d_{jt}}$ = vector of $\log(D_{jt}^m)$ values
F_{jt}	Forecasted organic demand in DMA j for week t, that is, $F_{jt} = E[D_{jt}]$. Forecasted at the time advertising is purchased. F_t is the national forecast for week t.	$F_{jt}^m = E[D_{jt}^m]$ = observed component m of forecasted organic demand; $\mathbf{f_{jt}}$ = vector of $\log(F_{jt}^m)$ values, among other definitions discussed. Provides the excluded variable in the IV regressions.
A_{jt}	Advertising purchased for market j in period t. A_{0t} denotes national advertising.	c_0 = cost of national advertising, A_{0t}; c_1 = cost of local advertising, A_{jt}
G_{jt}	Goodwill stock accumulated in market j at time t. $G_{jt} = 1 + \sum_{\tau=0}^{L} \lambda^{\tau}(A_{j(t-\tau)} + A_{0(t-\tau)})$	$g_{jt} = \log(G_{jt})$; $\mathbf{g_{jt}}$ = vector of g_{jt} values for different brands. λ = advertising carryover across periods
X_{ijt}	Nonadvertising marketing variables in store i in period t. May include prices, display advertising, and feature status, for example.	

Notes. Notational conventions: lowercase variables are the log of capital variables, for example, $d_{jt} = \log(D_{jt})$. Bold print are vectors of variables, for example, $\mathbf{d_{jt}} = (\log(D_{jt}^1)\log(D_{jt}^2)\ldots\log(D_{jt}^M))'$.

The level of organic demand realized in market j in period t is given by D_{jt} and m is the margin received on product sales. Finally, we have normalized the total population to one, and s_j represents the share of the population in market j.

For many applications, including those presented in this paper, firms do not know D_{jt} when advertising is purchased. In such cases, we assume firms rely on forecasts of organic demand, $E[D_{jt}] = F_{jt}$, which may be based on historical averages. Although the quality of these forecasts may influence the firms' optimal level of advertising, we do not explicitly model this (see Section 3.4.3 for further discussion).

3.1.1. Advertising Spillover in Two Markets and One Period.

To illustrate how the model in (1) generates advertising spillovers, we present a simplified version of this model. This exercise provides intuition for the source of advertising spillovers and insight into the market conditions that encourage firms to generate advertising spillovers as part of their optimization strategy.

Consider a model in which the firm sells to two markets, z and y, and there is only one period for which the firm performs the profit optimization. Additionally, for simplicity of notation, we set $\gamma = 1$ and $s_z = s_y = \frac{1}{2}$. Given these simplifications, we can rewrite Equation (1) as

$$\pi = \frac{1}{2}\left(D_z G_z^{\theta} + D_y G_y^{\theta} - c_1 A_z - c_1 A_y\right) - c_0 A_0. \qquad (2)$$

Goodwill stock simplifies to $G_j = 1 + A_j + A_0$.

To solve this optimization problem, the firm must consider two cases to determine which is most profitable. In the first case, the firm only buys national advertising, so $A_z^* = A_y^* = 0$, where "$*$" indicates the optimal solution for the case with only national advertising. Taking the first order condition with respect to A_0 gives the optimal level of national advertising for this case:

$$A_0^* = \left(\frac{\theta(D_z + D_y)}{2c_0}\right)^{\frac{1}{1-\theta}} - 1. \qquad (3)$$

From this expression, it is clear that, for either market, the organic demand level in the other affects the level of advertising received as discussed earlier. Specifically, Equation (3) demonstrates the advertising spillovers that occur because $\frac{\partial G_z^*}{\partial D_y} \neq 0$ and $\frac{\partial G_y^*}{\partial D_z} \neq 0$: the advertising needs of each market influence the amount of advertising received in the other. These nonzero derivatives form the crux of the spillover strategy.

In the second case, the firm buys some national plus local advertising in the market with the highest per capita returns to advertising but not in the other market. The optimal firm would never buy local advertising in the market with the lowest returns because national advertising is cheaper, and returns in the other market are, by definition, higher and receive even greater returns from additional national advertising. To formalize this argument, assume, without loss of generality, that $D_z > D_y$. It follows that

$$\theta D_z G^{\theta-1} - c_1 > \theta D_y G^{\theta-1} - c_1.$$

Additionally, given α, D_z, D_y, and c_0, there exists a G' such that

$$\theta D_z G'^{\theta-1} - c_1 > 0 > \theta D_y G'^{\theta-1} - c_1.$$

Therefore,

$$\left.\frac{\partial\pi}{\partial A_z}\right|_{G'} = \left(\theta D_z G'^{\theta-1} - c_1\right) > 0 > \left.\frac{\partial\pi}{\partial A_y}\right|_{G'} = \left(\theta D_y G'^{\theta-1} - c_1\right). \tag{4}$$

This expression shows that the incentive to buy additional advertising may exist for z but not y because z has the highest returns to advertising. We can make a similar argument regarding the returns to national advertising, finding a G'' such that

$$\left.\frac{\partial\pi}{\partial A_0}\right|_{G''} = \theta G''^{\theta-1}(D_z + D_y) - 2c_0 > 0 > \left.\frac{\partial\pi}{\partial A_y}\right|_{G''} = \left(\theta D_y G''^{\theta-1} - c_1\right). \tag{5}$$

Equations (4) and (5) demonstrate that the incentive to buy local advertising for y would never exist because, as G increases, returns to A_y become negative before returns to A_0 or A_z become negative. Having shown that $A_y^\dagger = 0$, we can now solve for $A_z^\dagger$ and $A_0^\dagger$, where "$\dagger$" indicates the optimal solution for the case with both national and local advertising.

Assuming that $A_z > 0$, the optimal solution for A_0 is found by setting $\frac{\partial\pi}{\partial A_0} = \frac{\partial\pi}{\partial A_z} = 0$, which leads to the solution

$$A_0^\dagger = \left(\frac{\theta D_y}{2c_0 - c_1}\right)^{\frac{1}{1-\theta}} - 1. \tag{6}$$

Local advertising is purchased in market z such that $\frac{\partial\pi}{\partial A_z} = 0$. This leads to the solution

$$A_z^\dagger = \left(\frac{\theta D_z}{c_1}\right)^{\frac{1}{1-\theta}} - \left(\frac{\theta D_y}{2c_0 - c_1}\right)^{\frac{1}{1-\theta}}. \tag{7}$$

Equations (6) and (7) make apparent that, in the second case, the advertising received in each market is driven only by the local organic demand levels. That is, $G_y^\dagger = (\frac{\theta D_y}{2c_0 - c_1})^{\frac{1}{1-\theta}}$, so $\frac{\partial G_y^\dagger}{\partial D_z} = 0$ and $G_z^\dagger = (\frac{\theta D_z}{c_1})^{\frac{1}{1-\theta}}$, so $\frac{\partial G_z^\dagger}{\partial D_y} = 0$. As a result, the spillover strategy will not work if local advertising purchases maximized profits.

Consequently, Equation (7) has important implications for the identification strategy. For example, if $A_z^\dagger$ is negative, we know that, to find a feasible solution, we have to constrain $A_z = 0$. This corresponds to the first case with the solution given in Equation (3), which encourages firm behavior we can exploit using the spillover strategy. Alternatively, if $A_z^\dagger$ is positive, then the second case is optimal, and the spillover strategy will not work. Thus, by inspecting Equation (7) we can see how the market parameters affect the relevance of the spillover strategy. The results are intuitive: the strategy is more likely to work as local prices increase and national prices decrease. Additionally, organic demand levels that are more similar across locations help ensure that the firm avoids local advertising, but at the same time, Equation (3) suggests that statistical power may be improved if they are no more similar than necessary.

3.1.2. Advertising Spillover in *N* markets and One Period. Extending the model to N markets reveals that local advertising may not be as devastating to the identification strategy as the two-market model suggests. With N markets and, for simplicity, $\gamma = 1$ and $s_j = \frac{1}{N}$, the firm wishes to optimize

$$\pi = \frac{1}{N}\sum_{j=1}^{N}\left[D_j G_j^\theta - c_1 A_j\right] - c_0 A_0. \tag{8}$$

For convenience, we order the N markets according to the magnitude of their organic demand such that $D_1 > D_2 > \ldots > D_N$. Now, consider a solution to this N-market problem in which it is optimal to buy local advertising for any market j such that $D_j \geq D_n$ for some integer $n \in [1, N]$. Under these conditions, the optimal amount of national advertising occurs at the point at which $\frac{\partial\pi}{\partial A_0} = \frac{\partial\pi}{\partial A_1} = \ldots = \frac{\partial\pi}{\partial A_n} = 0$. This leads to the solution

$$A_0^* = \left(\frac{\theta\sum_{j=n+1}^{N} D_j}{Nc_0 - nc_1}\right)^{\frac{1}{1-\theta}} - 1. \tag{9}$$

The N potential values of n can be evaluated to find the n that optimizes profits. Importantly for the identification strategy, Equation (9) shows that, even if n markets receive local advertising, $N - n$ markets are left for identification because their shared advertising levels are determined by the others' demand levels. Although the optimal value of n depends on market parameters, only for the special cases $n = N$ and $n = N - 1$ does the spillover strategy break down completely.

3.1.3. Advertising Spillover over Time. Introduction of the time dimension does not meaningfully change the insights from our single-period models. In the absence of advertising carryover, single-period results remain optimal: without dependencies across time periods, each time period can be viewed as a repeated cross-section. In the presence of advertising carryover, dependencies across periods arise that complicate the optimal solution but do not negate the cross-market influence demonstrated by the single-period models.

Although our empirical applications do not offer evidence of spillovers across time, such spillovers could be generated under a couple of conditions. First, spillovers are generated if the firm is required to set a constant advertising level over a given time span though their data reveal that demand changes meaningfully within that span. Second, cross-period spillovers are generated if there exist significant adverting carryover and significant changes in demand across periods. For example, if a firm faces high demand in the current period and low demand in the next, it may optimally purchase less advertising than is optimal for the current period to mitigate the over-supply of advertising stock in the following period. Finally, although not technically spillovers, the time dimension also introduces the potential for firms to experience forecast errors, which also provide quasi-random variation that can be used for identification (Hortacsu and Puller 2008).

3.2. Demand Model

Having formalized the supply-side incentives that give rise to the spillover strategy, we now present a means of exploiting mass advertising for identification. We present a general expression for the demand model that allows for the identification of multiple advertising effects, such as competitive advertising, though simpler implementations are also possible.

We begin by considering the following first-stage expression for advertising:

$$\mathbf{1}'\mathbf{g_{jt}} = \alpha_i^0 + \delta_\tau^0 + \phi^0 X_{ijt} + \left(\boldsymbol{\gamma}^0\right)'\mathbf{d_{jt}} + \left(\boldsymbol{\theta}^0\right)'\mathbf{f_{jt}} + \varepsilon_{ijt}^0, \quad (10)$$

where we have used lower case to indicate logs and bold to indicate vectors. Following this notation, $\mathbf{g_{jt}}$ is a vector containing the log of goodwill stock for market j during time t and may include components such as own- and competitive-brand advertising or different mediums of national advertising. As a result, $\mathbf{1}'\mathbf{g_{jt}} = \sum_b \log(G_{jt}^b)$, where b indexes advertising types of interest. Although the supply model describes advertising decisions at the market level, in practice, in may be helpful to control for store-level effects. As such, α_i^0 represents fixed effects for store i, and δ_τ^0 are time fixed effects for each period τ, which may span more than one t. Additional covariates are represented by X_{ijt}, including marketing activity other than advertising. Because the econometrician may not have a perfect measure of organic demand and because organic demand may have multiple components, we set $D_{jt} = (\prod_m D_{jt}^m)\xi_{jt}$, where D_{jt}^m is the mth observed component of organic demand and ξ_{jt} is the unobserved component of organic demand. The log of each of the observed demand components are represented by $\mathbf{d_{jt}}$, implying $(\boldsymbol{\gamma}^0)'d_{jt} = \sum_m \gamma^{0m} d_{jt}^m$. Finally, we denote forecasted levels of organic demand as $f_{jt}^m = \log(F_{jt}^m) = \log(E[D_{jt}^m])$, which serve as the excluded variable(s) when organic demand is not known with certainty when advertising is purchased. We discuss other possible definitions for $\mathbf{f_{jt}}$ as follows.

We now write the second-stage expression as

$$y_{ijt} = \alpha_i + \delta_\tau + \phi X_{jt} + \left(\boldsymbol{\gamma}\right)'\mathbf{d_{jt}} + (\boldsymbol{\theta})'\mathbf{g_{jt}} + \varepsilon_{ijt}, \quad (11)$$

where y_{jt} is the log of sales in market j during period t. Each of the covariates from Equation (10) are included but without supercripts. Given this model, the exclusion restriction requires that

$$\mathrm{cov}\left(\varepsilon_{ijt}, \mathbf{f_{jt}} \mid \alpha_i, X_{ijt}, \mathbf{d_{jt}}\right) = 0. \quad (12)$$

3.3. Instrument Definitions

The excluded variables, $\mathbf{f_{jt}}$, may be defined a number of different ways depending on the application. Under the simplest definition, $\mathbf{f_{jt}} = f_t$, which is the case in which the econometrician has a single measure of organic demand, and the firm sets advertising using a forecasted national average. In such a case, d_{jt} is also included in the first stage, so the identifying variation comes primarily from $f_t - d_{jt}$. This difference captures the degree to which a given market in a given period received suboptimal advertising. For example, if the difference is positive, nationally expected levels of organic demand are higher than they are locally, and as a result, the market is likely to receive more advertising than they require. The opposite follows for negative values of $f_t - d_{jt}$. Note that there are two reasons a market received locally suboptimal advertising. The first is forecast error: firms predicted some level of organic demand, but a different level was realized. The second source is the advertising spillovers that we explored in Section 3.1. Depending on the application, one of these sources may dominate the other.

Multiple instruments may be obtained with alternate definitions of $\mathbf{f_{jt}}$ depending on the data environment. For example, if the econometrician has measures of multiple components of organic demand, D_{jt}^m, these naturally provide the instruments $\mathbf{f_{jt}}$. Alternatively, with a single measure of organic demand but more than two markets, the forecasted level of organic demand for each market provides a different instrument such that $\mathbf{f_{jt}}$ contains the forecasts for each market. Another possible definition arises in conditions with significant advertising carryover and large variation in demand across periods. In such cases, firms may adjust advertising levels in period t based on forecasted levels of organic demand in both periods $t-1$ and $t+1$, implying $\mathbf{f_{jt}} = (f_{t-1} \quad f_t \quad f_{t+1})'$ provides valid instruments.

3.4. Estimate Robustness

The spillover strategy provides consistent estimates for the effect of advertising in a number of environments but certainly not all. We discuss the consistency in some of the most common data environments. Ultimately, the consistency of the estimates depends on the validity of the exclusion restriction (12). Most types of measurement error would not affect our estimates, including incomplete measurements of organic demand even if the firm has a better measure of organic demand than the econometrician and uses this measure to optimize advertising decisions. However, unobserved demand shocks or marketing activity that are correlated with the instrument violate the exclusion restriction and, thus, bias the estimates.

3.4.1. Unobserved Demand Shocks. An unobserved demand shock, ξ_{jt}, does not bias the advertising estimates as long as it does not correlate with the instrument, $\mathbf{f_{jt}}$. Given this lack of correlation, even if the advertising levels were chosen in part based on ξ_{jt} and are, therefore, correlated with advertising levels, this would not bias the advertising estimates. This is the key advantage of the spillover strategy over ordinary least squares (OLS), which would be biased if ξ_{jt} correlates with the advertising levels even after controlling for $\mathbf{d_{jt}}$. We demonstrate this advantage over OLS in Section 3.5 using simulated data. However, if ξ_{jt} correlates with the instrument, $\mathbf{f_{jt}}$, conditional on observables, the spillover strategy produces biased estimates.

3.4.2. Measurement Error in Organic Demand. The type of measurement error encountered in most applications would not bias the estimates; however, classical measurement error would bias the estimates. We expect that, in most applications, $\mathbf{d_{jt}}$ would be the output of a prediction model that conditions on predictors of organic demand. In such cases, the measurement error associated with observed $\mathbf{d_{jt}}$ does not correlate with the true values of $\mathbf{d_{jt}}$ and, therefore, does not bias the estimates (Hyslop and Imbens 2001). However, if there is classical measurement error, these unobserved values correlate with $\mathbf{d_{jt}}$ and, therefore, bias the advertising estimates upward.

3.4.3. Forecast Error. We consider three ways forecasting error could impact the results; none of these violate the exclusion restriction. First, the simple fact that $\mathbf{f_{jt}}$ is an imperfect predictor of $\mathbf{d_{jt}}$ improves the statistical power of the estimation strategy. These prediction errors, $\mathbf{f_{jt}} - \mathbf{d_{jt}}$, provide a high-quality source of random variation that can be used to identify the advertising effect though they are not, strictly speaking, advertising spillovers.

Second, the forecasts available to the econometrician may contain larger errors than those available to the firm. This scenario corresponds to that presented in Section 3.4.1 in which the firm has knowledge of ξ_{jt} (or high-quality predictions of it), which allows it to make better forecasts of organic demand than the econometrician and uses those higher-quality predictions to set advertising levels. The spillover strategy provides consistent estimates under these conditions.

Finally, we acknowledge that firms' advertising optimization is influenced by the existence of forecasting errors. The supply-side model (1) does not explicitly capture this part of the firm's decision, but forecasting errors would reduce targeting efficacy, thus lowering returns to advertising and, therefore, reduce optimal advertising levels. Such adjustments to the level of advertising overall would not violate the exclusion restriction (12).

3.4.4. Observed Local Advertising. As discussed in Section 3.1, firms may be incentivized to top up advertising in locations with high demand; such behavior does not bias the estimates of the spillover strategy but does limit its statistical power. In the extreme case, in which firms top up local advertising to its optimal level, the only source of identifying variation, $\mathbf{f_{jt}} - \mathbf{d_{jt}}$, results from prediction errors, not from advertising spillovers. If, in addition, firms perfectly predict the demand level in each market, then $\mathbf{f_{jt}} - \mathbf{d_{jt}} = \mathbf{0}$, and the estimation strategy has no power. The lack of power resulting from local advertising would be reflected by the *t*-statistics on the coefficient(s) for $\mathbf{f_{jt}}$ in the first-stage regression. This effect is demonstrated in the simulations (Section 3.5).

3.4.5. Unobserved Marketing Activity. Unobserved marketing activity may describe the most common violation of the exclusion restriction (12). We consider three types of unobserved marketing activity and their consequences.

First, the most benign type of unobserved marketing activity would simply correlate with unobserved demand shocks, ξ_{it}, but not the nationally forecasted level of demand. Because this marketing activity is random, it does not violate the exclusion restriction.

Another form of unobserved marketing activity would correlate with the forecasted demand. For example, the firm might increase the number of coupons for its product nationwide along with its advertising levels in anticipation of elevated demand. If the coupons are not observed, then the econometrician estimates the joint effect of the coupons and the television advertising. Thus, this generates a biased measure of the effect of advertising. However, in some cases, the firm may be interested in the joint effect of its advertising along with its other, correlated, marketing activity in which case the spillover strategy recovers the parameter of interest.

Finally, a type of unobserved marketing activity that has the potential to generate highly biased estimates is adjusted based on nationally purchased advertising levels and locally realized demand. For example, consider a case in which the econometrician only has data on national advertising levels, but the firm purchases a significant amount of local advertising to adjust its total advertising levels in response to locally realized demand shocks. In this case, $\mathbf{f_{jt}} - \mathbf{d_{jt}}$ is highly predictive of the amount of unobserved advertising purchased: more positive values of $\mathbf{f_{jt}} - \mathbf{d_{jt}}$ predict little to no unobserved advertising purchase, and more negative values predict large amounts of unobserved advertising purchases. Clearly, this violates the exclusion restriction and the spillover strategy provides inconsistent estimates.

3.4.6. Alternate Functional Forms. We have motivated the spillover strategy with a supply model (1) in which organic demand has a multiplicative effect on demand, which the firm may further amplify using advertising. The spillover strategy does not require this exact functional form, but does require a variable—similar to our organic demand—which exogenously shifts the marginal response to advertising. Because this variable shifts the marginal response to advertising, the firm optimally adjusts advertising levels based on the variable's expected levels at some aggregate level. In this way, the aggregate-level predictions of this variable can serve as an instrument that explains at least some of the variation in advertising levels. Many functional forms meet this requirement, not just the standard production function assumed in (1). However, not all functional forms meet this requirement. For example, if organic demand has an additive effect on advertising, organic demand does not shift the marginal response to advertising and, therefore, does not support the spillover strategy.

3.4.7. Local Average Treatment Effects. Because the spillover strategy is an IV strategy, it estimates local average treatment effects (LATE) (Imbens and Angrist 1994). In the case of only two markets (Section 3.1.1), the estimates combine the LATE of the market with too much advertising and that with too little advertising. Alternatively, with N markets (Section 3.1.2), the estimates combine the LATEs of the $N - n$ markets that rely solely on national advertising; the n markets for which it is optimal to include local advertising do not contribute to the estimate.

3.5. Spillover Strategy Estimates on Simulated Data

To demonstrate how an econometrician with incomplete information may benefit from the spillover strategy, we simulate data generated by the supply-side model presented in Section 3.1.1 that contains only two markets, $j \in \{y, z\}$. For this exercise, we assume the econometrician observes sales outcomes, Y_j, (where $Y_j = s_j D_j G_j \exp\{\varepsilon_j\}$ and ε_j is an unobserved random shock to sales), and the econometrician calculates goodwill stock from observed advertising expenditures, $G_j = 1 + A_0 + A_j$. We decompose organic demand such that $D_j = D_j^1 \xi_j$ to allow for the possibility that the econometrician does not observe ξ_j, but the firm does. Because there are just two markets, the econometrician instruments for the goodwill stock in market j using the demand level in the other market, $d_{-j} = \log(D_{(-j)})$. Specifically, for these spillover estimates, the estimating equations are:

$$\begin{aligned} \text{1st Stage}: g_{kj} &= \alpha^0 + \gamma^0 d_{kj}^1 + \theta^0 d_{k(-j)} + \varepsilon_{kj}^0 \\ \text{2nd Stage}: y_{kj} &= \alpha + \gamma d_{kj}^1 + \theta g_{kj} + \varepsilon_{kj}, \end{aligned} \tag{13}$$

where k indexes a cross-section of simulated, two-market environments. Lowercase indicates logs of the corresponding variables. To simulate the data, values d_{kj}^1, $\log(\xi_{kj})$, and ε_{kj} are drawn independently from $\sim \mathrm{N}(0, 1)$, and for each set of draws, k, the firm maximizes profits (2) by setting optimal advertising as specified in (3), (6), and (7).

The estimates recovered by the econometrician under different scenarios are presented in Table 2. In the first column, we see that, with complete information on all variables that affect the advertising decision, the econometrician can recover consistent estimates of the advertising elasticity using OLS. However, if data on ξ_j are missing, as in the second column, OLS produces biased estimates. The third column demonstrates the utility of the spillover strategy that produces consistent estimates of the advertising elasticity even without knowledge of ξ_j despite its influence on the advertising decision. Finally, in the last column, we simulate the effect of raising the price of national advertising. This price change gives the firm a greater incentive to incorporate local advertising into its strategy, and indeed, the share of simulated observations with local advertising increases. Additionally, the greater reliance on local advertising reduces the power of the spillover strategy: both the first-stage F- and t-statistics have been reduced and the standard errors on the advertising elasticity estimate increase. However, the increased reliance on local advertising does not affect the consistency of the advertising estimate.

4. Application

In this section, we apply the spillover strategy to the antihistamine category. First, we describe the data and provide summary statistics that provide an intuitive illustration of the source of identification for the strategy. Next, we present category- and brand-level estimates of advertising elasticities and compare these with the estimates produced by other strategies.

Table 2. Spillover Estimates on Simulated Data

		OLS		Spillover	
	True parameter values	Complete information (1)	Unobserved demand shocks (2)	Unobserved demand shocks (3)	More expensive national ads (4)
Advertising elasticity, θ	0.1	0.100 (0.006)	1.032 (0.005)	0.107 (0.019)	0.127 (0.077)
Observed organic demand, γ	1.0	0.999 (0.005)	0.460 (0.005)	0.989 (0.012)	0.977 (0.046)
Unobserved organic demand	1.0	1.003 (0.005)			
First-stage t-statistic on γ^0				97	24
F-statistic for excluded variable, $d_{(-j)}$				3,165	191
N simulated observations		100,000	100,000	100,000	100,000
Local advertising price		0.10	0.10	0.10	0.10
National advertising price		0.06	0.06	0.06	0.08
Share of observations with local ads		0.23	0.23	0.23	0.44

Note. Estimates are recovered from simulated data generated from the two-market, static, supply-side model presented in Section 3.1.1 with optimal advertising decisions made by the firm.

4.1. Data

The data required for the spillover strategy include data on sales, advertising, and organic demand. As discussed earlier, organic demand measures factors that influence demand that are outside the control of the firm. Such data are often available to firms because these are factors that influence their decision to adjust advertising levels. In this paper, however, we demonstrate the use of this methodology with nonproprietary data: pollen counts and weather measurements.

Raw pollen data are provided by the National Allergy Bureau, which is administered by the American Academy of Allergy, Asthma, and Immunology. These data were collected by 51 stations covering 39 DMAs across the United States, 2004–2013. 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. These contributors are listed in Online Appendix Table 1, and their locations are shown in Online Appendix Figure 1.

Pollen values are imputed for each DMA-week used in our analysis. These imputed values are generated by a random forest algorithm trained on the raw pollen values using weather and other variables as pollen predictors. Imputing the pollen values this way offers a few advantages over relying on the raw pollen values. First, the sample size is increased dramatically because the imputation allows us to make use of the more complete sales and advertising data. Second, the imputation helps ensure the measurement error in organic demand does not bias the estimates by reducing the chance the errors are classical in nature but instead are best estimates conditional on an information set (Hyslop and Imbens 2001). Finally, the imputation model generates smaller errors than the raw measurements. Details on the pollen data, imputation, measurement errors, and other summary statistics can be found in Online Appendix A. The imputed pollen counts are used to calculate organic demand as d_{jt} = log(Imputed Total Pollen Count + 30), which correlated well with sales in pretesting. Additionally, nationally forecasted pollen levels for each week, f_t, are calculated as the population-weighted average organic demand level for the years 2004–2011.

Weather data are used both for the imputation of the pollen data and separately as measures of organic demand for some product categories. These data were provided by the National Oceanic and Atmospheric Administration and were collected from 1,301 stations across the United States from 2003–2015. Weather data were interpolated for each county-day using bicubic interpolation.

Sales and advertising data come from Nielsen. Sales data were derived from the Retail Measurement Services, covering stores across the United States for 2010–2015, which are reported by UPC-store-week. Advertising expenditures from Nielsen's AdIntel product are reported for numerous media, are differentiated by national and local purchases, and report estimated spending on each advertisement. Estimates of the share of spending on each medium for the antihistamine category are provided in Table 3 and demonstrate that the antihistamine category overwhelmingly relies on national television advertising. Unless otherwise indicated, we compute advertising stock for each DMA-week as the combination of national advertising from all reported media, plus local television advertising (local advertising in other media are available for a limited number of DMAs). For antihistamines, we estimate that these sources account for 96% of all advertising. To compute advertising stock for antihistamines,

Table 3. Advertising Expenditures: Share of Total

Medium	National	Local
TV	0.823	0.051
Magazine	0.063	0.000
Internet	0.013	0.000
Newspaper	0.007	0.001
Radio	0.006	0.016
FSICoupon	0.000	0.020

Notes. Based on data from Nielsen's AdIntel for the antihistamine category. Estimates are based on average per-person expenditures across DMAs that include data on local expenditures for a given medium.

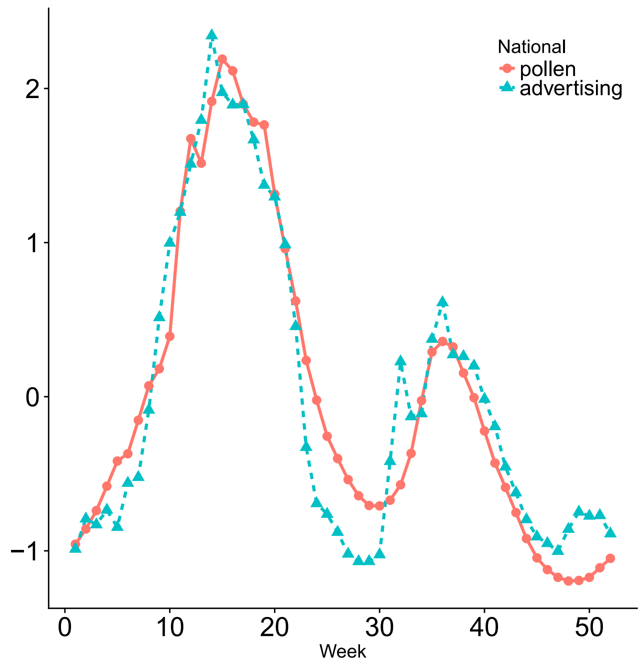
an advertising carryover of 0.69 per week is used, a value that optimizes a border-strategy model fit based on a grid search (Shapiro et al. 2019). Additionally, the Nielsen data provide estimates of the number of viewers of each advertisement. Using these data, we estimate that the median cost of national ads per viewer is less than half that for local ads, which gives firms a clear incentive to rely on national advertising as we assumed in in our supply-side model (Section 3.1).

In this analysis, we focus on the over-the-counter, second-generation, branded antihistamine products: Claritin, Zyrtec, and Allegra. The first-generation antihistamine, Benadryl, appears to have discontinued TV advertising in 2011 and is not included in our analysis. Brand-level summaries for antihistamines are provided in Online Appendix D.

4.2. Advertising Spillovers in the Antihistamine Category

The mechanics of the spillover strategy are especially tangible in the antihistamine market, which offers a clear measure of organic demand: pollen counts. Indeed, pollen drives much of the variation in antihistamine sales, and the typical pollen cycle also appears to drive changes in advertising levels over the course of a year. For evidence of the relationship between the pollen and advertising cycles, see Figure 1, which shows the national weekly averages of pollen and advertising in the United States, averaged across 2011–2015. In this figure, we can see that advertising levels track pollen levels closely, on average, rising steadily from the beginning of the new year and reaching its 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. This suggests that firms are targeting their advertising based on the forecasted level of pollen. Further analysis of how TV ads and other marketing variables are correlated with locally realized and nationally forecasted pollen levels are presented in Online Appendix A.4.

Figure 1. (Color online) Pollen and Advertising Levels



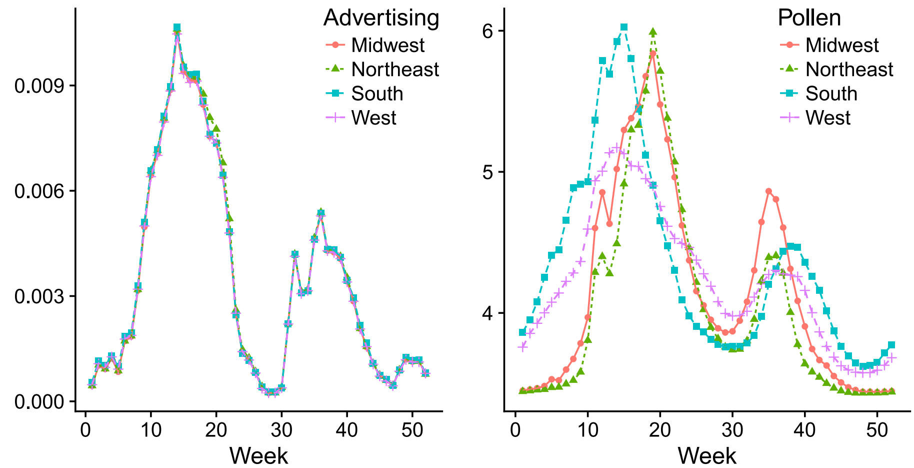
Note. Nationwide advertising and log of pollen levels by week of the year, averaged across 2011–2015.

Given this advertising behavior, one can see how advertising levels may not always be optimal. The first reason for suboptimality is fairly apparent: when firms purchase their advertising for 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 allergy season may occur earlier or later than expected. 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 suboptimal advertising levels may be less apparent but also provides a source of quasi-random variation for many products, including antihistamines. These suboptimal advertising levels result from the fact that firms 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, in many instances, advertising levels are not locally optimal.

For a dramatic illustration of this effect, see Figure 2. This figure shows the same data as Figure 1 but now disaggregated to the census-region level. Figure 2 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

Figure 2. (Color online) Regional Advertising and Pollen Levels

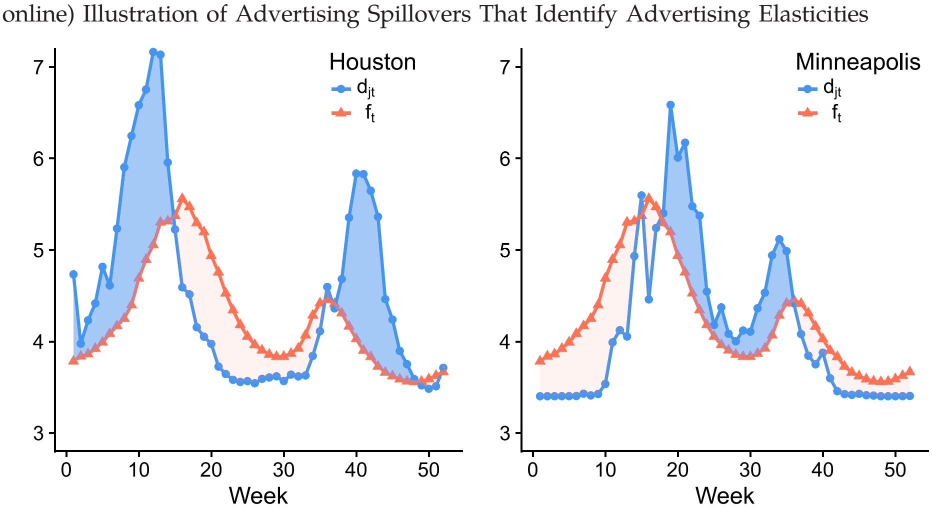


Note. Advertising and log of pollen levels by census region and week of the year, averaged across 2011–2015.

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 areas. Online Appendix A.3 provides a cluster analysis of these regional differences.

To develop a clearer understanding of the identifying variation that is harnessed by the spillover strategy, consider the advertising and pollen experiences of two very different cities: Houston and Minneapolis. Their experiences in 2011 are illustrated in Figure 3 with their organic demand levels labeled d_{jt}. Naturally, these demand profiles were quite different given their respective climates. Nevertheless, each city experienced the same national advertising and nearly the same total advertising. As we have seen, this advertising schedule is largely based on the nationally forecasted pollen levels, f_t. As a result, if local advertising is limited, each of these cities experiences either too much or too little advertising throughout the year. For example, around the eighth 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. At the same time, Minneapolis received the same amount of advertising despite there being essentially no pollen in the air. These types of outcomes are the ones our method treats as quasi-random and are, therefore, well suited to identifying the response to advertising. In Figure 3, dark-shaded regions represent observations in which a city received too little advertising ($f_t - d_{jt} < 0$), and light-shaded regions represent observations in which a city received too much advertising ($f_t - d_{jt} > 0$). Note that, at any given time, although one city is receiving too much, the other is

Figure 3. (Color online) Illustration of Advertising Spillovers That Identify Advertising Elasticities



Notes. d_{jt} = pollen realized in Houston and Minneapolis in 2011. f_t = nationally forecasted pollen. Light-shaded regions suggest too much advertising locally because $f_t - d_{jt} > 0$. Dark-shaded regions suggest too little advertising locally because $f_t - d_{jt} < 0$.

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 are used to estimate the advertising response. Additionally, this analysis suggests that identification comes primarily from observations in which the local demand conditions are at the greatest distance from the national forecast, that is, $|f_t - d_{jt}|$ is large. We test this empirically in Online Appendix B.

4.3. Estimates for the Antihistamine Category: Spillover Strategy

We now use the spillover strategy to estimate the response to advertising in the antihistamine category. For these category-level estimates, we implement estimating Equations (10) and (11) as follows. The dependent variable, y_{ijt}, measures the log of total antihistamine sales in each store. We analyze only the net effect of total category advertising, so $\mathbf{g_{jt}}$ is a scalar g_{jt}, which represents the (log of) total advertising stock across all brands in the market. In addition, to control for the potentially confounding influence of other marketing activity, X_{ijt} includes averages of the log of prices, display, and feature activity across brands in a store-week. Finally, we set $\mathbf{f_{jt}} = f_t$ as the nationally forecasted level of organic demand. This provides a single instrument, which is sufficient for category-level estimates. (For details on how f_t and d_{jt} are calculated, see Section 4.1.) In tables and graphs, we refer to this estimation method as "spillover."

Estimates for the spillover strategy are presented in Table 4. From the first-stage estimates, we can see the instrument is strong, in terms of both the partial F-statistic for the excluded variable, f_t, and its t-statistic. The first column leaves out the control for expected local demand levels to illustrate the bias in the advertising elasticity that results from violating the exclusion restriction (12), which appears large in this case. The correct specification of the spillover strategy in the second column reduces the advertising elasticity estimate more than five times.

Table 4. Spillover Strategy Application: Antihistamines

	(1)	(2)
First stage:		
Forecasted organic demand National (instr.)	1.195***	1.150***
	(0.006)	(0.035)
Organic demand		0.044
		(0.034)
Log of price	0.052***	0.048***
	(0.016)	(0.015)
F-statistic for excluded variable	10,881	220
R^2	0.633	0.633
Second stage:		
Advertising elasticity	0.246***	0.042***
	(0.013)	(0.016)
Organic demand		0.240***
		(0.016)
Log of price	–0.536***	–0.543***
	(0.017)	(0.012)
Number of observations	769,252	769,252
Number of stores	5,318	5,318
Number of DMAs	200	200
R^2	0.853	0.878

Notes. Estimates of the advertising elasticity for the antihistamine industry based on Nielsen data that cover 2011–2015 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 in a given year. "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. Regressions control for average feature and display activity across antihistamine brands in a store. Regressions include store and year fixed effects.
***$p < 0.01$.

4.4. Estimates for the Antihistamine Category: Estimator Comparisons

Although we have established that the spillover strategy has power, we do not know whether the estimates from the spillover strategy are accurate. To confirm the method's accuracy, we need to know the true advertising elasticity, and in the absence of randomized controlled trials to recover these values, we look to other estimation methods for comparison.[1] Estimates from each method are shown in Table 5. Here we discuss the motivation for each method and its potential biases.

4.4.1. Basic OLS (OLS: Naive). This method represents a standard estimation approach in the absence of data on organic demand. This specification includes store fixed effects but no time fixed effects. We expect this method would produce fairly biased estimates of the advertising elasticity because they are confounded by the effect of pollen and lack any attempt to control for the seasonality of the product.

4.4.2. OLS with Month-Year Fixed Effects (OLS: moFE). This method adds month-year fixed effects to OLS: naive. An econometrician concerned about seasonality might decide to include these additional fixed effects. In this application, however, the already upward-biased advertising elasticity estimates are not changed as a result of the additional fixed effects.

4.4.3. OLS with Month-DMA Fixed Effects. An alternative set of fixed effects that controls for different seasonality patterns across regions is month-DMA fixed effects. For identification, this model assumes that firms condition on the typical demand shocks for each DMA each month of the year but do not target their advertising efforts within months. Given the highly correlated changes in pollen and advertising

Table 5. Antihistamines Category Advertising Elasticity

	OLS					Border			Supply IV	Spillover	
	Naive (1)	Month-year fixed effects (2)	Month-DMA fixed effects (3)	Week-DMA fixed effects (4)	Organic demand (5)	Month-year-border fixed effects (6)	Organic demand (7)	Week-year-border fixed effects (8)	Political ads (9)	All observations (10)	Border observations (11)
Advertising elasticity	0.172***	0.171***	0.056***	0.051***	0.047***	0.139***	0.088***	0.041	−0.033	0.042***	0.027
	(0.010)	(0.009)	(0.008)	(0.015)	(0.008)	(0.007)	(0.006)	(0.046)	(0.050)	(0.016)	(0.027)
Log of price	−0.550***	−0.588***	−0.638***	−0.639***	−0.543****	−0.641***	−0.643***	−0.659***	−0.551***	−0.543***	−0.554***
	(0.016)	(0.016)	(0.015)	(0.014)	(0.012)	(0.013)	(0.013)	(0.015)	(0.015)	(0.012)	(0.011)
Organic demand					0.236***		0.169***	0.110***	0.294***	0.240***	0.264***
					(0.012)		(0.008)	(0.023)	(0.036)	(0.016)	(0.016)
Year fixed effects	X				X				X	X	X
Month-year fixed effects		X									
Month-DMA fixed effects			X								
Week-DMA fixed effects				X							
Month-year-border fixed effects						X	X				
Week-year-border fixed effects								X			
Number of observations	769,252	769,252	769,252	769,252	769,252	403,165	392,149	369,954	769,252	769,252	285,381
Number of stores	5,318	5,318	5,318	5,318	5,318	1,847	1,847	1,847	5,318	5,318	1,834
Number of DMAs	200	200	200	200	200	178	178	160	200	200	178
R^2	0.858	0.868	0.877	0.881	0.878	0.893	0.895	0.910	0.875	0.878	0.867
F-stat (excluded variables)									9	220	95

Notes. Estimates of the advertising elasticity for the antihistamine industry based on Nielsen data that cover 2011–2015 for each week and store. Robust standard errors, clustered by DMA, are reported in parentheses. Regressions are weighted by the number of units sold by store in a given year. "Organic demand" is a log transformation of the total pollen count. Prices, feature status, and display status are averaged across the antihistamine products in each store and these average values are included as controls in each regression. All regressions include store fixed effects. The instrument for the log of advertising is the log of the expected national pollen level for a given week.

$^{***}p < 0.01$.

across weeks (Figure 2), the absence of week-level targeting appears inconsistent with the data. However, in this application, the estimates recovered are much smaller than the previous two estimates, statistically significant, and as we will see, similar in magnitude to those recovered by more preferred approaches.

4.4.4. OLS with Week-DMA Fixed Effects. This specification tightens the time fixed effects relative to those in the last specification to control for week-level adjustments in advertising intensity. As a result, it makes use of only the year-to-year variation in advertising levels for each DMA. For example, if firms only conditioned on the demand shocks for each DMA typically experienced to determine their advertising intensities and the only deviations from this policy were random, this would identify the advertising elasticity. However, this approach limits the identifying variation, and the standard error is larger than for many of the other estimates. Nevertheless, its point estimate is similar to other preferred estimates.

4.4.5. OLS with Organic Demand (OD) as a Control (OLS + OD). This approach offers a potentially superior specification to the preceding specifications by making use of additional data. If the measure of organic demand controls for all factors that are correlated with the advertising decision, this provides consistent advertising estimates. Indeed, we can see that the advertising estimates are a fraction of the size estimated by the first two methods, consistent with our suspicion that those methods over estimate the advertising effect.

4.4.6. Border Strategy (Border). This strategy for estimating TV advertising effects comes from Shapiro (2018). The approach makes use of the discontinuity in advertising intensities that occurs across DMA boundaries. If there are sufficient differences in local advertising levels across DMA boundaries, this offers a plausible source of random variation. Specifically, this method analyzes only counties that are adjacent to counties in different DMAs. To control for unobserved shocks that could confound the advertising estimates, the method can include fixed effects for the set of counties lying along a given DMA border for each month-year. These monthly fixed effects absorb the influence of any unobserved demand shocks (organic, marketing, or competitive) that are approximately uniform for a given DMA border in a given month. However, the antihistamine data suggest that firms may target their advertising to accommodate changes in demand within months. Consistent with this concern, this border strategy specification recovers an estimate similar to that reported in the first column of the table, which we suspect overestimates the effect of advertising.

4.4.7. Border Strategy with Organic Demand as a Control (Bord + OD). To test whether the border-time fixed effects used in the standard border strategy are sufficient to prevent bias from short-term, unobserved events, such as changes in the pollen levels, we include an additional specification that controls for organic demand. The addition of organic demand to the regression lowers the border estimates significantly. The change suggests that the border-month fixed effects are insufficient to protect the border estimates from bias by unobserved demand shocks that change rapidly over time, such as changes in pollen conditions.

4.4.8. Border Strategy with Week-Border Fixed Effects. Because the estimates from the standard border strategy changed when we added a control for organic demand, we explore a version of the strategy with tighter time controls. These time controls are tighter than in the original application of the method (Shapiro 2018) but should help to control for local demand shocks that occur over short time periods, such as pollen. Unfortunately, these tighter controls also restrict the identifying variation: only differences in advertising levels in a given week across a DMA boundary are used for identification. As a result, the standard errors are quite large, and the advertising estimate is not significant. However, this result does not necessarily generalize to all applications. In environments with more local advertising usage, which create potentially larger differences in the advertising levels of adjacent DMAs, this strategy could provide precise and accurate estimates of the advertising effect.

4.4.9. Political Advertising Instrument. Sinkinson and Starc (2019) observe that political advertising may provide a useful instrument because it displaces advertising for other products. Furthermore, the intensity of political advertising in the United States varies widely over time with significant cross-sectional differences driven by the idiosyncrasies of the U.S. electoral process. As result, political advertising may provide variation that is uncorrelated with the demand shifts associated with many products. However, for antihistamines, we find that the cross-sectional increases in political advertising correlate with low pollen periods, and therefore, bias advertising estimates of a control for organic demand is not included (Online Appendix C.1). Controlling for organic demand, the instrument recovers an insignificant estimate for antihistamine advertising though the standard error is relatively large and the confidence interval overlaps those of other estimates.

4.4.10. Spillover Strategy. These estimates are similar in magnitude to those obtained by the other specifications we have used, which include controls for organic demand. Formally, we fail to reject a difference between advertising elasticities in columns (5) and (10) ($p = 0.64$) using a difference-in-Sargan test, which assumes that the instrument for the spillover strategy is valid, and tests whether advertising as specified in column (5) is different and, consequently, endogenous (Baum et al. 2003). In this application, the spillover strategy produces standard errors nearly twice the size of many of the OLS-based methods. In general, IV methods are less efficient than OLS strategies, and because the border strategy critically relies on measures of organic demand for identification, measurement error in this value may increase its standard errors. However, unlike for the OLS-based methods, classical measurement error in advertising intensity does not bias the spillover estimates.

4.4.11. Spillover Strategy Estimated on Border Stores Only. To offer a more direct comparison of the estimates from the spillover and border strategies, we report estimates of the spillover strategy using only stores that are used for the border strategy. Relative to estimates on the entire set of stores, the point estimates drop, and the standard errors increase to the point the advertising estimate is no longer significant. However, the confidence interval still overlaps that of the full-sample spillover estimate in column (10) and with the OLS + OD specification in column (5).

4.4.12. Summary. 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 drive advertising decisions. In this example, including organic demand in the OLS estimates appears to reduce bias to the point at which OLS is indistinguishable from the spillover estimates. In this case, the spillover strategy has helped to confirm that, for antihistamines, by controlling for organic demand, OLS estimates are not subject to large endogeneity bias. However, we remain uncertain of the true advertising elasticity; the two specifications that we suspect have the greatest potential to recover the true estimates, columns (5) and (7), do not have overlapping confidence intervals. Nevertheless, the spillover estimate overlaps the confidence interval for column (5). In comparison with existing literature, the magnitude of the spillover estimate is smaller than the average found in a meta-study by Sethuraman et al. (2011) and larger than the average recovered by the border strategy across hundreds of categories (Shapiro et al. 2019).

Antihistamine advertising also provides an opportunity to test whether methods are robust to unobserved demand shocks. In columns (6) and (7), we showed how a standard specification of the border strategy is prone to bias from unobserved demand shocks from pollen. In Online Appendix C.1, we provide such comparisons for other methods and also find evidence of their biases as well.

4.5. Estimates for Antihistamine Brands

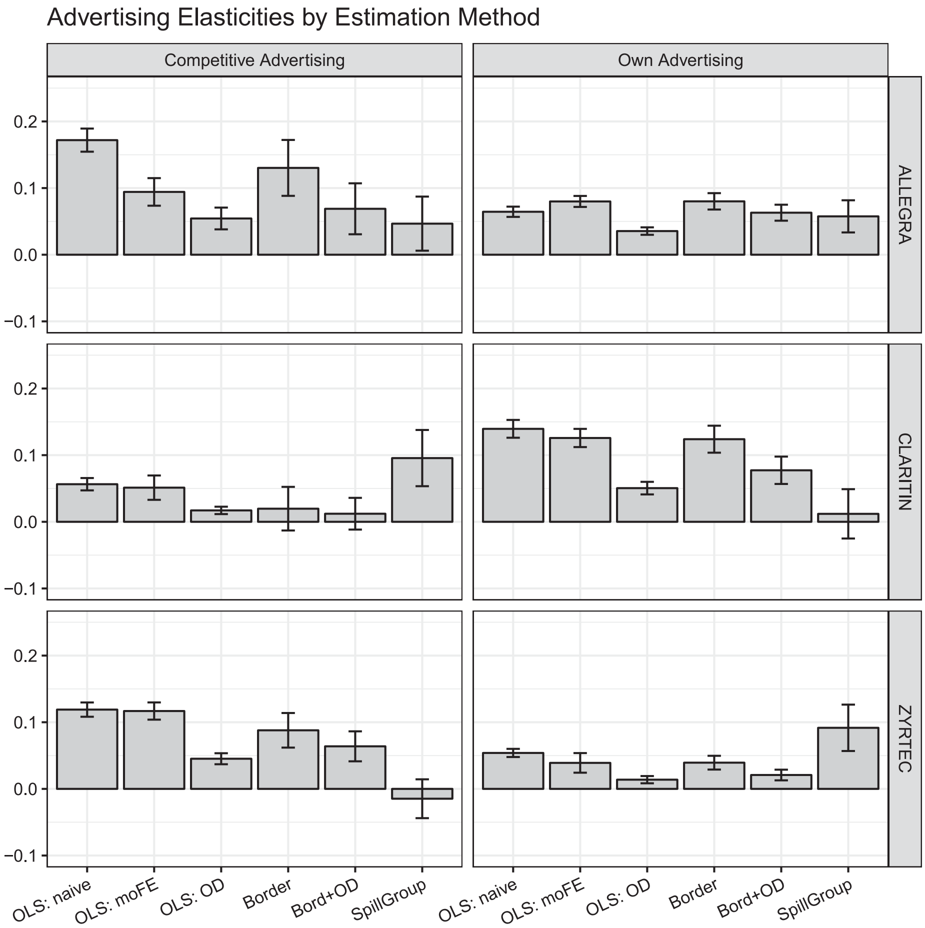
Brand-level estimates can be more challenging to obtain than category-level estimates. Not only are the factors that drove a given firm's advertising decisions likely to be unobserved, but the same is likely for the firm's competitors. As a result, at least two instruments are required: one for the firm's own advertising and one for competitors' advertising.

We discussed a number of potential instruments in Section 3.3, but the relevance of these depends on specific market conditions and data availability. For this application, we make use of the fact that the data contain more than 200 DMAs, each with differing forecasted demand profiles. Although, in principle, this offers a large number of instruments, we limit the number to avoid ensuring they are well powered. We reduce the number of demand forecasts by defining four regions of the United States with common pollen profiles using a clustering algorithm (details in Online Appendix A.3). The demand forecasts for each of these four regions provide our instruments along with the two-way interactions of their forecasts. Estimates using this instrument definition are called "SpillGroup" in figures and tables.

We analyze the advertising responses to the three largest antihistamine brands: Claritin, Zyrtec, and Allegra. We again use (10) and (11) as our estimating equations but with the following definitions. Because the regressions are at the brand level, sales, given by y_{ijt}, now represent the log of sales of a given brand. Additionally, we now consider the influence of two types of advertising stock: competitive and own advertising, $\mathbf{g_{jt}} = (g^c_{jt} \quad g^{own}_{jt})'$, where g^c_{jt} represents all advertising for products that compete with the brand and g^{own}_{jt} represents the advertising of the target brand. We also expand X_{ijt} to control for the average prices of competing brands and their feature and display activity.

Estimates for the advertising elasticities, competitive and own, are presented in Figure 4 (with detailed regression tables in Online Appendix D). Just as we saw for the category-level estimates, OLS: naive tends to estimate much higher responses than OLS: OD for both own and competitive advertising. The border estimates follow a similar pattern to the OLS estimates, producing smaller estimates when controlling for organic demand though the differences are not always statistically significant. Finally, the SpillGroup method estimates large standard errors in this

Figure 4. Brand-Level Regressions: Antihistamines



Notes. Estimates of the advertising elasticities for own-brand advertising and advertising from competing brands. Regressions run separately for each brand and method. Standard errors (SE) are clustered by DMA, and error bars represent 95% confidence intervals.

application, and its confidence intervals overlap with those of many of the other estimators for Allegra but not Claritin or Zyrtec. For Claritin, the SpillGroup estimates insignificant effects for own advertising and highly positive effects for competitive advertising, the opposite of the pattern recovered by other estimators. Similarly, the SpillGroup estimates oppose the other estimators for Zyrtec, estimating large effects for own advertising and insignificant effects for competitive advertising. Given that we do not know the true effects, we cannot determine which set of estimates are most accurate in this application.

5. Discussion

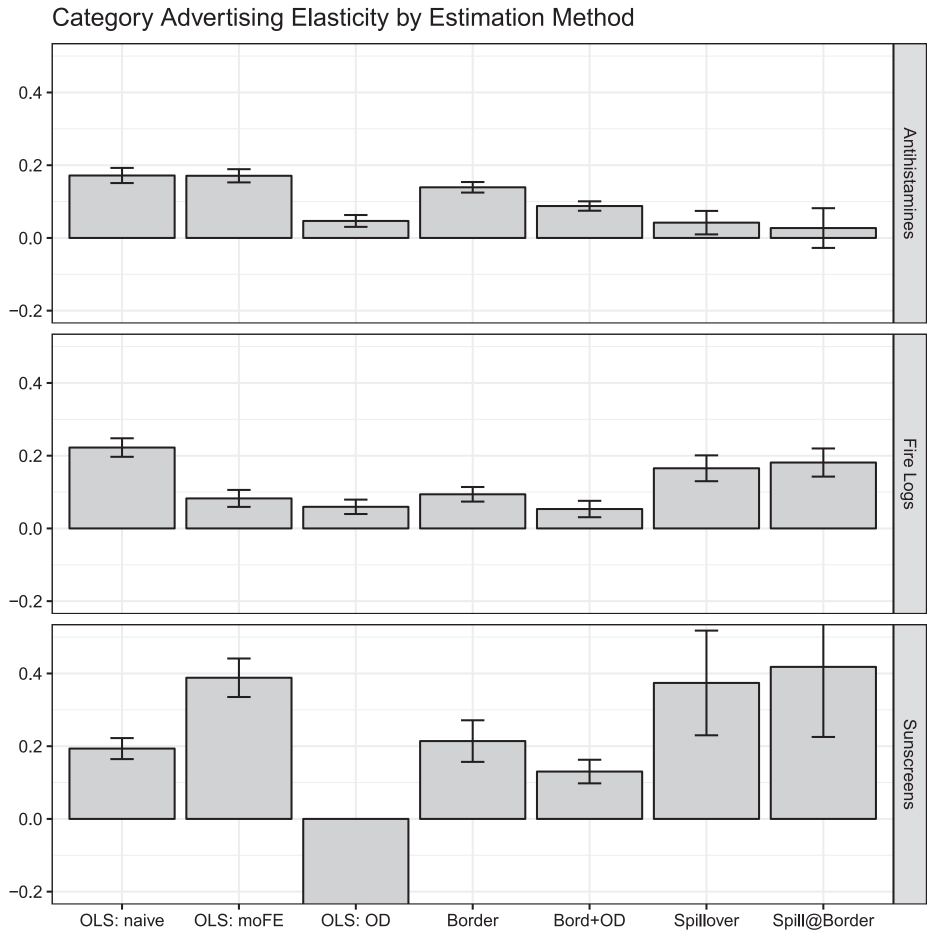
In this section, we briefly present estimates for the fire log and sunscreen categories using the spillover strategy. Additionally, for antihistamines, we test whether our model predicts the same advertising schedule as we observe in the data and explore the robustness of the strategy to counterfactual market conditions.

5.1. Additional Category Estimates

In this section, we use the strategy to produce estimates for two additional seasonal categories that are affected by local weather conditions: sunscreens and fire logs. For each of these categories, we use local average temperature as the measure of organic demand. Results from these estimates are presented in Figure 5 (Online Appendix E provides further summary statistics and regression tables for these categories). The methods applied here are described in Section 4.4.

Estimates for these categories further demonstrate the potential for differences between estimates from the spillover strategy and standard OLS approaches. In contrast to the antihistamine category, the spillover estimates for the new categories are larger than the OLS: OD estimates, instead recovering estimates that are more similar to the OLS: naive. For antihistamines, the difference between spillover and OLS: OD was not significant, but for sunscreens and fire logs,

Figure 5. Category-Level Regressions



Notes. Estimates for multiple categories using different estimation methods. Standard errors (SE) are clustered by DMA, and error bars represent 95% confidence intervals.

our test rejects their equivalence (difference-in-Sargan test with $p < 0.001$ in both cases). Although the spillover strategy recovers larger estimates than the border strategy for sunscreens and fire logs, it recovers similar estimates for border counties and the entire population, unlike Li et al. (2019), which finds the border counties respond differently to political advertising than the rest of the U.S. population. For the sunscreen category, we observe large differences across the estimates recovered by different methods; OLS: OD recovers an estimate of −0.37, a value on the far left tail of the distribution of advertising elasticities estimated by Shapiro et al. (2019). Such large differences across commonly employed estimators suggest the need for continued work to develop reliable methods for estimating advertising effects.

5.2. Tests of the Supply-Side Assumptions

We partially motivate the spillover strategy with its ability to identify advertising effects even under optimal firm advertising behavior. We now investigate whether, in the antihistamine market, firms act optimally by relying heavily on national advertising. We estimate that the market's reliance on national advertising is, in fact, optimal and then explore counterfactual market conditions that could potentially incentivize firms to shift to local advertising. Our supply-side model suggests that the same market conditions that would cause firms to shift from national to local advertising would also cause them to purchase very little advertising overall.

5.2.1. Optimal Advertising Schedule. To solve for the optimal advertising schedule, we optimize the supply-side model (1) subject to market parameters we have recovered for the antihistamine market. These parameters are reported in Table 6 along with their sources. Most of these parameters have been estimated already or are readily obtained. However, no direct measure of the margin earned in the category is available, so we have determined a value of m by finding the value that generates an optimal solution that includes an advertising-to-sales ratio of 0.025, which is broadly consistent with industry reports.

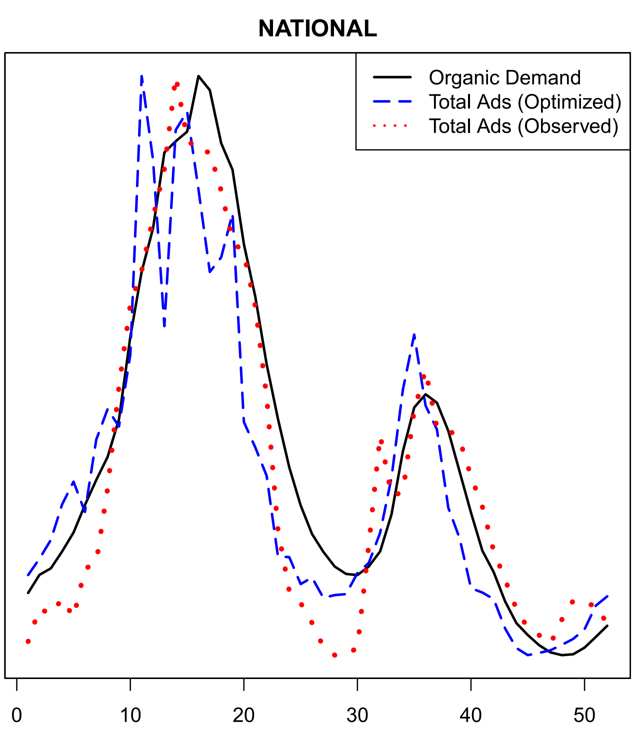
Table 6. Supply-Side Model Parameters

Parameter	Value	Source
Advertising elasticity, θ	0.04	Estimated in Table 4
Organic demand coefficient, γ	0.24	Estimated in Table 4
Advertising carryover, λ	0.69	Grid search estimation
Ad cost ratio, $\frac{c_0}{c_1}$	0.39	Estimated from Nielsen data
Margin, m	5.00	Model calibration
Number of markets, N	197	Nielsen data
Number of weeks, T	52	Assumed
Population shares, s_j	Varies	Share of households in Nielsen data

We assume the firm faces the market conditions represented by the Nielsen data: 197 markets in the country with population shares for each market, s_j, based on the number of households Nielsen reports for each DMA. Additionally, values of d_{jt} are based on average observed values for each week of the year.

Optimizing the local and national advertising schedule using numerical methods, we find that the model recommends a negligible amount of local advertising, which is fairly consistent with our estimates for the market (Table 3). However, the optimized schedule for national advertising, plotted in Figure 6, differs slightly from the schedule we observe in the data. In many weeks, the optimized strategy advertises ahead of forecasted pollen increases and relies on advertising carryover when pollen levels are expected to drop, especially as an allergy season is tapering off. Alternatively, the observed advertising profile follows the forecasted pollen levels closely, apparently failing to account for advertising carryover, which we estimate reduces profits by 0.56%. This difference between observed and optimized advertising should not affect our application of the spillover strategy because we did not rely on the firms' accounting for advertising carryover. However, had the firms accounted for advertising carryover, additional instruments would have been available from the leads and lags of forecasted organic demand as discussed in Section 3.3.

Figure 6. (Color online) Optimized Advertising



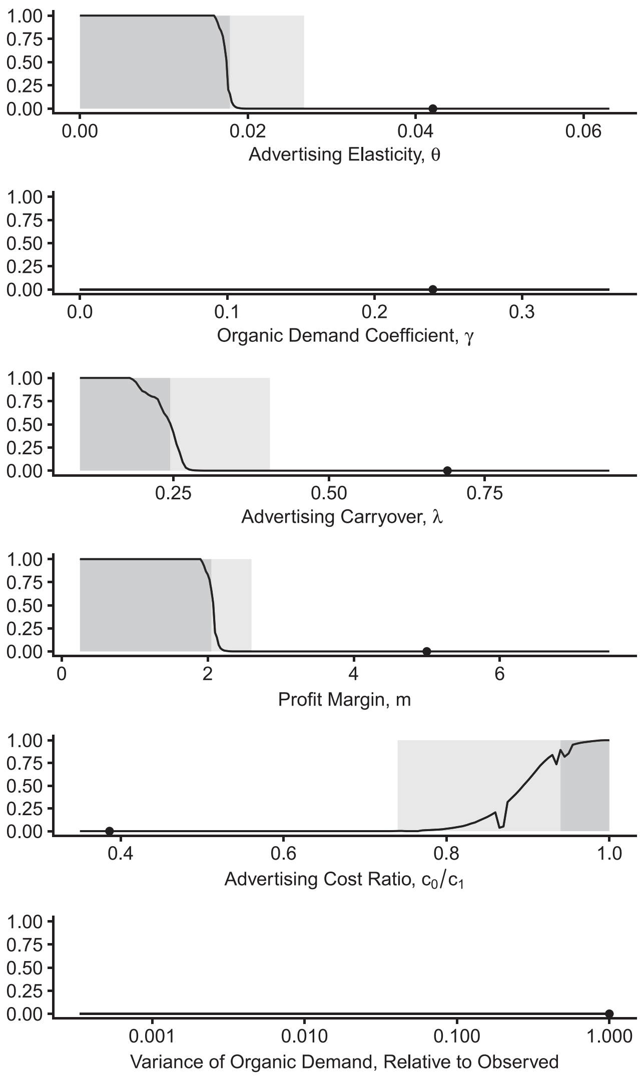
Notes. Optimal advertising levels over the course of the average year. All values have been normalized to be shown on the same y-axis.

5.2.2. The Robustness of the Spillover Strategy to Market Conditions.

Although the antihistamine market's reliance on national advertising appears optimal, the possibility remains that small changes in the market conditions could reverse this result and cause firms to rely more heavily on local advertising, which would reduce the relevance of the spillover strategy. To investigate this sensitivity, we introduce counterfactual market parameters to the supply-side model, solve for optimal advertising under each counterfactual, and observe the resulting changes in the reliance on national versus local advertising. For this exercise, we define a measure of local advertising intensity as the share of DMA-weeks in which the recommended share of local advertising exceeds 0.1%. For the counterfactual settings, we change one market parameter at a time, setting each to discreet points along a range of values surrounding our estimates of the true values.

Plots of the optimal local advertising intensities as a function of market conditions are presented in Figure 7. For reference, the estimated true market values from Table 6 are denoted by large dots. Additionally, the figure reports the advertising-to-sales ratio associated with each solution: light-shaded regions indicate ratios between 0.01 and 0.001, and dark-shaded regions indicate ratios below 0.001; by construction, the dots report points at which the ratio is 0.025.

Figure 7. Local Advertising Intensity: Sensitivity to Market Conditions



Notes. Each plot reports the local advertising intensity on the y-axis for the antihistamine market under different market-parameter settings. Local advertising intensity is the share of DMA-weeks in which more than 0.1% of the advertising is local according to the simulated optimal solution. In each graph, only the parameter indicated on the x-axis is varied; other parameters are set at their estimated/assumed levels. The estimated/assumed level for each parameter in the antihistamine market is indicated with a dot on each graph. Regions in which the optimal solutions have low advertising–sales ratios are indicated with shaded regions: the lighter regions correspond to ratios of 0.01 to 0.001, and the darker regions correspond to ratios smaller than 0.001.

Some of the results in Figure 7 are intuitive. For example, reductions in both the advertising elasticity and the advertising carryover encourage firms to purchase less advertising (i.e., move toward darker shaded regions). Similarly, raising the cost of national advertising eventually shifts firms toward using more local advertising. However, the plots also suggest a less obvious result: reductions in overall advertising levels are accompanied by increases in local advertising intensity. This prediction supports the relevance of the spillover strategy by suggesting that when firms are in market conditions that encourage them to buy a meaningful amount of advertising, they also tend to find themselves in conditions that encourage them to rely more heavily on national advertising.

6. Conclusion

This paper presents a method for estimating the response to mass advertising. It exploits events in which firms implement their advertising or other marketing activity at high-enough levels of aggregation so that measures of organic demand allow the econometrician to approximate the degree to which advertising is too high or low for some markets or individuals. Such conditions are likely to proliferate as long as measures of organic demand are increasingly available and discounts remain for purchasing advertising in bulk. The method provides firms a means of obtaining 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 that both standard methods and the recently popularized border strategy can produce biased estimates.

Acknowledgments

The author thanks his PhD advisers Günter Hitsch and Pradeep Chintagunta for their guidance on this project. The author thanks Bradley Shapiro, Sanjog Misra, and Matthew Gentzkow for their input on this project. Additionally, the author thanks Wesley Hartmann, Aviv Nevo, Sarah Moshary, and other participants at the 2018 Quantitative Marketing and Economics conference for their helpful comments and suggestions.

Endnote

[1] Additional comparisons are provided in Online Appendix C. There we show that "instrument-free" methods do not perform well on the data simulated for Section 3.5 and that advertising cost instruments produce first-stage results that are not consistent with the law of demand.

References

Baum CF, Schaffer ME, Stillman S (2003) Instrumental variables and GMM: Estimation and testing. *Stata J.* 3(1):1–31.

Blake T, Nosko C, Tadelis S (2015) Consumer heterogeneity and paid search effectiveness: A large-scale field experiment. *Econometrica* 83(1):155–174.

Gentzkow M, Shapiro JM (2010) What drives media slant? Evidence from U.S. daily newspapers. *Econometrica* 78(1):35–71.

George L, Waldfogel J (2003) Who affects whom in daily newspaper markets? *J. Political Econom.* 111(4):765–784.

Gordon BR, Zettelmeyer F, Bhargava N, Chapsky D (2019) A comparison of approaches to advertising measurement: Evidence from big field experiments at Facebook. *Marketing Sci.* 38(2):193–225.

Hartmann WR, Klapper D (2018) Super Bowl ads. *Marketing Sci.* 37(1):78–96.

Hastings JS (2004) Vertical relationships and competition in retail gasoline markets: Empirical evidence from contract changes in Southern California. *Amer. Econom. Rev.* 94(1):317–328.

Hausman JA (1996) Valuation of new goods under perfect and imperfect competition. Bresnahan TF, Gordon RJ, eds. *The Economics of New Goods* (University of Chicago Press, Chicago), 207–248.

Hortacsu A, Puller SL (2008) Understanding strategic bidding in multi-unit auctions: A case study of the Texas electricity spot market. *RAND J. Econom.* 39(1):86–114.

Hyslop DR, Imbens GW (2001) Bias from classical and other forms of measurement error. *J. Bus. Econom. Statist.* 19(4):475–481.

Imbens GW, Angrist JD (1994) Identification and estimation of local average treatment effects. *Econometrica* 62(2):467–475.

Joo M, Wilbur KC, Cowgill B, Zhu Y (2014) Television advertising and online search. *Management Sci.* 60(1):56–73.

Kalyanam K, McAteer J, Marek J, Hodges J, Lin L (2018) Cross channel effects of search engine advertising on brick & mortar retail sales: Meta analysis of large scale field experiments on Google.com. *Quant. Marketing Econom.* 16(1):1–42.

Lewis RA, Rao JM, Reiley DH (2011) Here, there, and everywhere: Correlated online behaviors can lead to overestimates of the effects of advertising. Sadagopan S, Ramamritham K, eds. *Proc. 20th Internat. Conf. World Wide Web* (ACM, New York), 157–166.

Li X, Hartmann WR, Amano T (2019) Identification using border approaches and IVs. Preprint, submitted June 19, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3402187.

Liaukonyte J, Teixeira T, Wilbur KC (2015) Television advertising and online shopping. *Marketing Sci.* 34(3):311–330.

Lodish LM, Abraham M, Kalmenson S, Livelsberger J, Lubetkin B, Richardson B, Stevens ME (1995) How TV advertising works: A meta-analysis of 389 real world split cable TV advertising experiments. *J. Marketing Res.* 32(2):125–139.

Moshary S, Shapiro B, Song J (2019) How and when to use the political cycle to identify advertising effects. Preprint, submitted August 7, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3404343.

Nair HS, Misra S, Hornbuckle WJ, Mishra R, Acharya A (2017) Big data and marketing analytics in gaming: Combining empirical models and field experimentation. *Marketing Sci.* 36(5):699–725.

Nevo A (2001) Measuring market power in the ready-to-eat cereal industry. *Econometrica* 69(2):307–342.

Tellis GJ, Chandy RK, Thaivanich P (2000) Which ad works, when, where, and how often? Modeling the effects of direct television advertising. *J. Marketing Res.* 37(1):32–46.

Sethuraman R, Tellis GJ, Briesch RA (2011) How well does advertising work? Generalizations from meta-analysis of brand advertising elasticities. *J. Marketing Res.* 48(3):457–471.

Shapiro BT (2018) Positive spillovers and free riding in advertising of prescription pharmaceuticals: The case of antidepressants. *J. Political Econom.* 126(1):381–437.

Shapiro B, Hitsch GJ, Tuchman A (2019) Generalizable and robust tv advertising effects. Working paper, University of Chicago, Chicago.

Sinkinson M, Starc A (2019) Ask your doctor? Direct-to-consumer advertising of pharmaceuticals. *Rev. Econom. Stud.* 86(2):836–881.

Stephens-Davidowitz S, Varian H, Smith MD (2017) Super returns to Super Bowl ads? *Quant. Marketing Econom.* 15(1):1–28.