BLAKELEY B. MCSHANE, ERIC T. BRADLOW, and JONAH BERGER*

New car purchases are among the largest and most expensive purchases consumers ever make. While functional and economic concerns are important, the authors examine whether visual influence also plays a role. Using a hierarchical Bayesian probability model and data on 1.6 million new cars sold over nine years, they examine how visual influence affects purchase volume, focusing on three questions: Are people more likely to buy a new car if others around them have recently done so? Are these effects moderated by visibility, the ease of seeing others' behavior? Do they vary according to the identity (e.g., gender) of prior purchasers and the identity relevance of vehicle type? The authors perform an extensive set of tests to rule out alternatives to visual influence and find that visual effects are (1) present (one additional purchase for approximately every seven prior purchases), (2) larger in areas where others' behavior should be more visible (i.e., more people commute in car-visible ways), (3) stronger for prior purchases by men than by women in male-oriented vehicle types, (4) extant only for cars of similar price tiers, and (5) subject to saturation effects.

Keywords: social identity, visual influence, automobiles, cars

Visual Influence and Social Groups

New car purchases are among the largest and most expensive purchases consumers make over the course of their lives. Not surprisingly, then, this purchase is heavily shaped by both functional and economic concerns. People may consider the condition of their existing car, budget constraints, or whether a local dealership is having a sale. Commuters may place greater value on fuel efficiency, while families may place greater value on safety. Indeed, existing research shows that factors such as Internet sales (Scott-Morton, Zettelmeyer, and Silva-Risso 2001), dealership locations (Bucklin, Siddarth, and Silva-Risso 2008), and car make, transaction type (purchase or lease), and pricing incentives (Busse, Simester, and Zettelmeyer 2010; Silva-Risso and Ionova 2008) all shape new car purchases.

However, might visual influence, that is, the observation of others' behavior, also play a role in auto purchases? Might people be more likely to buy a new car if they have seen others around them do so recently (Grinblatt, Keloharju, and Ikaheimo 2008)? Furthermore, might these effects vary based on the social identity of the prior adopters (Berger and Heath 2008; Simmel 1908)? For example, might men's behavior vary depending on whether prior buyers were women versus men (Shang, Reed, and Croson 2008; White and Dahl 2006)? Might these effects also depend on the identity relevance of the vehicle type (Berger and Heath 2007)? While prior research has demonstrated social influence for a particular make and model (e.g., the Toyota Prius; Narayanan and Nair 2011), might similar effects hold for new cars more generally? If so, to what degree are they driven by visual influence?

Using data on approximately 1.6 million new cars sold over a nine-year period, this research examines how new car purchase volume is influenced by the social identity and proximity of other new car buyers. We make several contributions. First, we investigate whether, in addition to being shaped by various functional and economic concerns, new car purchases are also shaped by visual influence. In doing so, we consider two tests of how "visibility," or the degree to which people observe others' cars, has an impact on these effects by (1) varying the distance of prior purchases from current purchases to make prior purchases more or less visi-

*Blakeley B. McShane is Assistant Professor of Marketing, Kellogg School of Management, Northwestern University (e-mail: b-mcshane@kellogg.northwestern.edu). Eric T. Bradlow is K.P. Chao Professor of Marketing, Statistics, and Education (e-mail: ebradlow@wharton.upenn.edu), and Jonah Berger is James G. Campbell Jr. Memorial Assistant Professor of Marketing (e-mail: jberger@wharton.upenn.edu), The Wharton School, University of Pennsylvania. The authors are grateful to Jorge Silva-Risso for helping arrange access to the PIN data and to Florian Zettelmeyer for many helpful conversations about this article. All correspondence should be sent to the first author. Eric Bradlow's research is supported in part by the Wharton Customer Analytics Initiative. Ron Shachar served as associate editor for this article.

ble and (2) utilizing the varying levels of car visibility observed empirically across geographic regions. Second, we investigate how visual effects are moderated by the social identity of the current and prior adopters. Does visual influence vary depending on whether current and prior purchasers are from similar or different social groups? Does this depend on the identity relevance of the vehicle type or the vehicle's price tier? Finally, we also investigate the functional form of visual influence, testing whether it is an absolute or relative phenomenon and whether it is subject to saturation effects (e.g., diminishing marginal returns).

Influence is notoriously difficult to estimate. Consequently, we use several strategies to rule out challenges such as homophily and correlated unobservables (Bell and Song 2007; Manski 1993). First, we employ quasi-experimental techniques, including treatment conditions, control conditions, and varying dosages. We implement these techniques through a matching process that utilizes both geographic distance and distance in covariate space, thus serving to make our observation units homophilous but distant (and thus invisible). Second, we include model-based controls such as heterogeneous intercepts, seasonality, time trends (parametric and nonparametric), and overdispersion, which can mitigate these effects (Van den Bulte and Lilien 2001). Finally, we perform extensive out-of-sample tests on our model. The tests outlined previously (and described more fully in the subsection "Is It Really Visual Influence?") are among the richest in the literature to date and serve to rule out several plausible alternatives, thus bolstering our claim of visual influence.

We organize the article as follows. Next, we build on the literature to lay out hypotheses about visual influence effects in the domain of new cars. Then, we provide a detailed description of our data and describe our quasi-experimental design, our model, and how we rule out effects other than visual influence. Because each local region has differing effects and characteristics, we develop a hierarchical Bayesian probability model (Rossi, Allenby, and McCulloch 2003) that governs purchase volumes and enables us to test the various moderators of visual effects mentioned previously in one coherent model. Finally, we discuss our results and their importance.

HYPOTHESES

Decades of research across psychology, marketing, economics, and sociology has examined how social influence shapes decision making. In general, this work finds that people tend to conform to the behavior of others. They evaluate coffee more favorably when others like it (Burnkrant and Cousineau 1975) and judge lines to be of similar length to what others around them suggest (Asch 1956). Similarly, giving consumers information about what songs or dinner entrées prior consumers preferred increased the choice of those items (Cai, Chen, and Fang 2009). Economic models of information cascades and bandwagon effects (Banerjee 1992; Bikhchandani, Hirshleifer, and Welch 1992) take a similar approach, suggesting that people are more likely to do something if others around them have (or have not) done so.

But while prior work has mostly examined somewhat low-cost decisions (e.g., what entrée to buy), we suggest that influence may also shape "big-ticket" decisions such as new car purchases. People are usually uncertain about whether they really need a new car: sure, the old car's interior is a bit ratty and the brakes could use fixing, but one could always patch things up rather than get something new. Seeing others driving new cars, however, should encourage people to do the same and buy one themselves. Others' behavior can act as a kind of social proof, in which people assume the actions of others reflect the correct thing to do in a given situation (Cialdini 2008). In this instance, seeing others driving a new car should make the idea of getting a new car more top of mind and make it seem like a more normative or correct thing to do (Deutsch and Gerard 1955).

As an initial exploration of this phenomenon, consider Figure 1, which plots the monthly sales of automobiles in a given zip code to each gender against the prior month's sales in that zip code to females (left panel) and males (right panel). There are several noteworthy features. First, there is a positive relationship between prior and current sales (i.e., the smooth curves generally slope upward), suggesting a potential for visual influence. Second, this relationship appears to vary by the "sending" (i.e., compare genders across panels) and "receiving" (i.e., compare genders within panels) groups. That is, the four plotted curves are not identical. Third, the relationship appears subject to saturation effects (i.e., the curves tend to flatten out as prior sales increase)—a phenomenon well-studied in marketing and in particular in advertising (Dubé, Hitsch, and Manchanda 2005; Johansson 1973; Laurent, Kapferer, and Roussel 1995).

Before examining these issues, we first lay out our hypotheses beginning with the principal one:

> H_1: Consumers purchase more new cars when they have *seen* others around them do so recently.

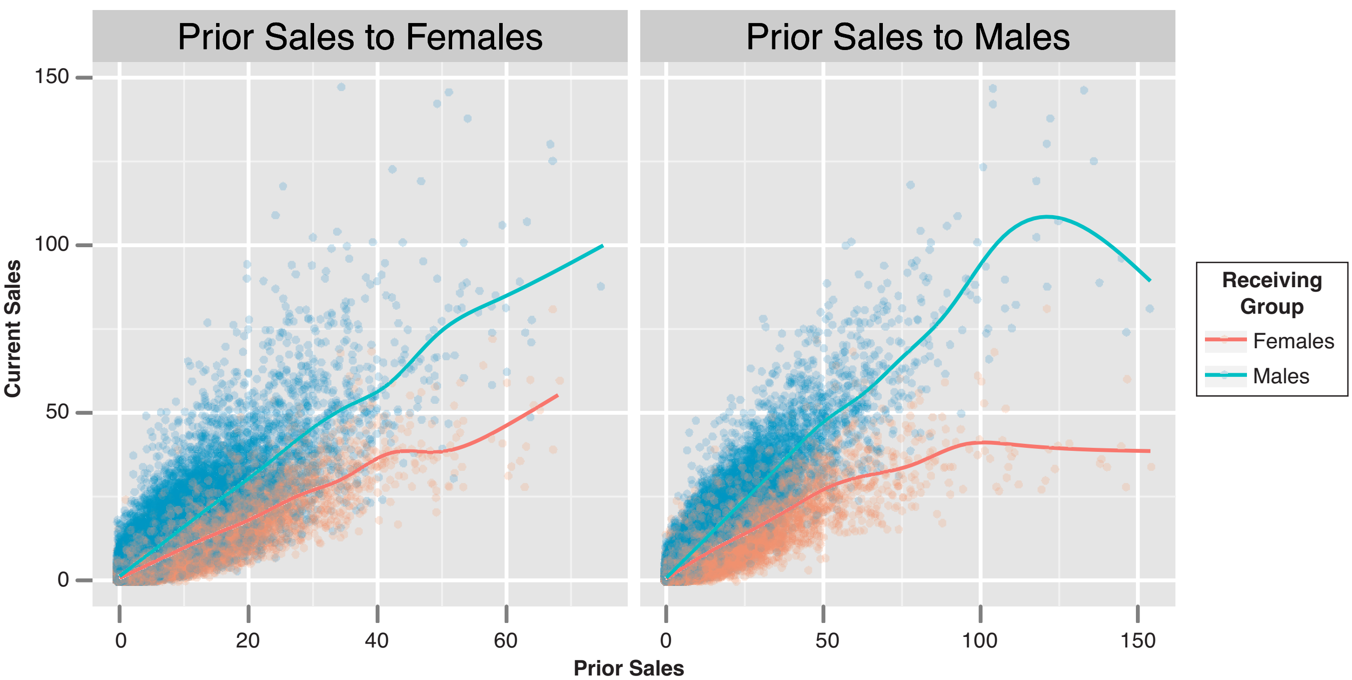
When people see others around them driving new cars, they should be more likely to purchase new cars themselves. Thus, recent prior purchases that are local, and therefore visible, should have an impact on current purchases. That said, while people often see what their neighbors or other people who live nearby are driving, they less frequently see the cars driven by people who live farther away. Because they cannot see these "out of eyeshot" cars, they should not be affected by them:

> H_{1a}: Consumers do not purchase more new cars when those out of eyeshot have done so recently.

While H_{1a} can be viewed as an additional test of visibility (i.e., one cannot be influenced by something one cannot see), it can also be viewed as a robustness check on H_1; thus, we denote it "auxiliary" (H_{1a}). If, by analogy to the experimentation literature, H_1 yields the hypothesized positive "treatment" effect of visible prior purchases, H_{1a} yields the hypothesized null effect for the "control" or "placebo" condition of prior purchases that are far away and thus invisible. Therefore, finding a null effect for H_{1a} (i.e., no evidence of an effect for purchases that are out of eyeshot) is as important for our theory as finding a positive effect for H_1 (i.e., evidence of an effect for purchases that are within eyeshot) because it provides additional evidence that any observed relationship between people's purchases is driven by *visual* influence.

As a further theoretical contribution and to more strongly demonstrate that any observed relationships are driven by

Figure 1
CURRENT SALES VERSUS PRIOR SALES BY GENDER



Notes: Each point gives the monthly sales of automobiles in a given zip code to each gender plotted against the prior month's sales in that zip code to females (left panel) and males (right panel). The smooth curves are fit using a generalized additive model with the degree of smoothness estimated from the data. The data show a positive relationship between prior and current sales and that this relationship varies by gender.

visual influence, we also examine whether local influence effects are moderated by differences in visibility observed across geographic regions. Not only should visual influence be present for visible, within-eyeshot purchases (H_1) and absent for invisible, out-of-eyeshot purchases (H_{1a}), but it should also be stronger in geographic locations where it is easier to see what others are driving. In places where more people tend to be on the road driving, for example, there should be larger effects because these people have a greater opportunity to observe whether others are driving new cars. Similarly, the weather conditions in different geographic regions may make it easier or more difficult to see what other people are driving. Thus, we expect that the size our visual influence estimates should relate to local-level variables (e.g., the number of people who drive to work, weather conditions):

> H_2: In places where it is easier to see what others are driving, visual influence on new car purchases is larger.

To extend the treatment analogy further, we conceptualize this hypothesis as one involving how the treatment effect works as the "dosage" is varied: in places where new car purchases are more visible (i.e., higher dosage), visual influence should be stronger.

In addition to examining whether the local influence effects described in H_1 are moderated by visibility, we also examine whether they are moderated by social identity. While consumers often behave similarly to those around them, in other cases they shun behaviors linked to social groups they want to avoid looking like (Berger and Heath 2007, 2008; Simmel 1908). For example, students are less likely to choose junk food at the dining hall when that behavior is associated with avid video game players (Berger and Rand 2008), and men are less likely to choose a "ladies' cut steak" (White and Dahl 2006). Furthermore, these effects need not necessarily be reciprocal (Van den Bulte and Joshi 2007): geeks may imitate jocks, while jocks may either fail to imitate geeks or even avoid things chosen by geeks.

In this research, we let gender be our marker of social identity. Gender is a good choice because it is an established dimension of social identity (Bem 1981; Deaux et al. 1995) and has been shown to affect behavior in other research (Shang, Reed, and Croson 2008; White and Dahl 2007). Furthermore, the situation of a car speeding by requires the viewer to rapidly classify its driver, and therefore a highly visible and apparent marker of social identity that applies to and is easily identified by all people is desirable (Cialdini 2008). Gender fulfills this criterion, whereas other, more granular social groupings (e.g., preppies) may not be realistic for this setting.

We examine both the magnitude and direction of cross-group influences. For example, are males more likely to buy new cars when males versus females have bought recently? We further examine how these group-on-group differences vary for identity-relevant versus identity-neutral vehicle types (Berger and Heath 2007). As we show subsequently, certain types of cars (e.g., pickup trucks) are more strongly associated with particular social identities than others (e.g., conventional cars such as sedans). Consequently, building on prior work demonstrating that people are more likely to avoid products associated with other social groups in identity-relevant domains, positive cross-gender influence effects should be weaker in these more identity-relevant car categories. More specifically, we hypothesize variation in influence effects by social group:

H_3: Visual influence varies by sending/receiving groups. In particular, the (fe)male sending group has a greater effect for (fe)male-oriented cars.

Comparing these relative effects both deepens our understanding of visual influence and has important managerial implications for targeting effects (Joshi, Reibstein, and Zhang 2009).

We also test the functional form of visual effects. First, social effects are typically assumed to be linear and absolute. However, Figure 1 suggests that saturation effects, commonly observed in other areas of marketing, may exist for visual influence. Consequently, we test a logistic specification in addition to a linear one. Second, because each geographic region has a different volume of sales, we also test whether the effects should be relative (i.e., whether absolute prior sales should be used as the covariate or whether a transformation that accounts for the size relative to the volume of new cars sold in that locale should be used). Finally, we also perform various robustness tests of our model, including a theoretically motivated one pertaining to the price tier of the vehicle.

DATA

Power Information Network Data

Our data on automobile purchases come from the J.D. Power and Associates Power Information Network (PIN). The PIN division was founded in 1993 with the objective of collecting car sales transaction data from a large sample of dealerships representative of the U.S. market; currently, approximately one-third of U.S. dealers are enrolled in the network. Each night, these dealers transmit their daily transactions to J.D. Power and Associates, which then makes the data available to academic researchers (for more elaborate descriptions of this network, see Bucklin, Siddarth, and Silva-Risso 2008; Busse, Simester, and Zettelmeyer 2010; Dasgupta, Siddarth, and Silva-Risso 2007; Scott-Morton, Zettelmeyer, and Silva-Risso 2001, 2003; Silva-Risso and Ionova 2008; Srinivasan et al. 2004).

Our particular data set contains the number of automobiles sold to each gender in 905 U.S. zip codes in each month from January 1999 through March 2008.[1] All 905 zip codes are contained within C = 40 randomly sampled large U.S. counties, and up to 25 zip codes were randomly sampled from each county (we have $905 < 40 \times 25 = 1000$ zip codes because some counties have fewer than 25 zip codes). Note that our zip codes are those of the purchasers as opposed to the dealers.

Zip code–level data are frequently used when studying automobile decisions (Shriver 2010), particularly at the monthly level (Narayanan and Nair 2011). Such data not only are more commonly available but also facilitate the detection of visual influence effects beyond the noise likely to exist in individual-level transactions. Our data also enable us to easily relate the observed effects to visual and other data (through aggregate data sources such as the U.S. Census) available only at the zip code and county levels. Furthermore, given the heavy computation associated with the Bayesian approach taken here, these data provide the ability to observe greater geographic variation (conditional on a fixed data set size). We discuss additional studies using individual-level data in our concluding section, but the main trade-offs are between customer lifetime value issues associated with individual-level repeat purchases data (not available from PIN data) on the one hand and the distance/visibility issues associated with aggregate-level geographic variation (of primary interest in the current study) on the other hand.

In addition to sales volume, we have extensive covariates pertaining to the cars sold, which act as important controls in our model. In particular, we have the price, manufacturer's rebate, annual percentage rate (APR) of interest of the loan, term (length) of loan, monthly payment, amount financed, manufacturer's suggested retail price (MSRP), residual value, and residual percentage of the cars themselves. Table 1 presents a set of summaries of our data set.

Supplemental Data

We supplement the PIN data with county-level and zip code–level covariates from the 2000 U.S. Census. We use the former to test our hypotheses on identity diffusion and its moderation by visibility (H_2) and use the latter for matching (see the "Matching Design" section).

Unfortunately, there are no direct measures of visibility at the local level (Heffetz 2011). Consequently, we constructed two surrogate measures to test H_2. While these measures are by no means perfect, they are a reasonable reflection of the underlying notion of visibility that we hope to capture. First, we examine the number of people in the county who drive to work. This is a direct measure of car visibility in that it captures how many individuals in the county are frequently exposed to other cars. Second, we tap a more literal measure of visibility by considering the weather. If it is rainy outside, visibility is low, but, the sun-

Table 1

SUMMARY OF PIN VARIABLES

Variable	*25%*	*Median*	*Mean*	*75%*
Price ($)	19,800	24,900	27,500	32,000
Rebate ($)	1000	2000	2170	3000
APR (%)	4.19	6.33	6.77	8.64
Term (months)	48	60	55.8	66
Monthly payment ($)	341	433	474	555
Amount financed ($)	18,300	23,700	26,000	30,900
MSRP ($)	21,800	28,100	31,100	36,300
Residual ($)	12,000	16,000	18,400	21,900
Residual percentage	49.0	54.0	52.5	59.0

[1]The counties selected listed in alphabetical order by state name were Jefferson, AL; Pima, AZ; Contra Costa, CA; Los Angeles, CA; Riverside, CA; San Francisco, CA; Denver, CO; New Castle, DE; District of Columbia, DC; Dade, FL; Escambia, FL; Lee, FL; Orange, FL; Pasco, FL; DeKalb, GA; Cook, IL; Polk, IA; Baltimore City, MD; Middlesex, MA; Kent, MI; Macomb, MI; St. Louis, MO; Washoe, NV; Hudson, NJ; Bronx, NY; Erie, NY; Kings, NY; Westchester, NY; Wake, NC; Tulsa, OK; Bucks, PA; Erie, PA; Philadelphia, PA; Richland, SC; Davidson, TN; Bexar, TX; Harris, TX; Salt Lake, UT; King, WA; and Milwaukee, WI

nier it is, the easier it is to see what people next to you are driving. Consequently, we examine the average number of days per year that are sunny in each county.

To support our claim that visual influence should be moderated by local indicators of visibility, we control for variables that are also likely to affect the magnitude of visual influence—even when the effect of these other variables is not of direct interest. This strengthens our claim that it is local visibility and not other factors that moderates the strength of visual influence.

One important potential moderator of visual influence is population size because the number of people is directly related to the number of potential sources for influence. People in certain regions may also care more about cars, so we examine the average number of vehicles per household. Income is likely to affect the size of visual influence because households require the means to "keep up with the Joneses." Similarly, labor participation is likely related to the presence of working-age citizens and dual-income households and thus the need for both transportation and income. Finally, urban versus nonurban environments are associated with different needs for and means of transportation. Table 2 provides summary statistics for these county-level data.

In addition to capturing county-level variables, we also make use of data at the zip code level. For each zip code, we track the population and the fraction of the population that lives in urban areas, is white, is male, and is a child. We also track the number of households, the fraction of commuters who take a car to work, the fraction who take public transportation, the average time to work, the median household income, the per capita income, the average age, and the average education level in years. These covariates are central to the matching design we present subsequently. Table 3 presents summary statistics.

MODEL

Matching Design

In this section, we describe our model for $y_{z,g,t}$, the total number of new cars sold in zip code z to social group g during month t. We begin by discussing a critical feature of our model, choosing "partner" zip codes p(z) for each focal zip code z. Partner zip codes help overcome the threats to validity discussed previously (Manski 1993) and enable us to test H_1 and H_{1a}.

For each focal zip code, we seek a set of partner zip codes that have varying degrees of visibility from the perspective of the focal zip code. We achieve this variation in visibility by using geographic distance, and thus we match each focal zip code with four partner zip codes: (1) a contiguous "neighbor" partner zip code; (2) a "near" partner zip code, which is 10–30 miles away; (3) a "moderate" partner zip code, which is 30–60 miles away; and (4) a "far" partner zip code, which is more than 100 miles away. The neighbor and near partner zip codes should have fairly high visibility from the perspective of the focal zip code, the moderate partner zip code should have moderate visibility, and the far partner zip code should have near zero visibility (of course, the focal zip code itself should have maximum visibility).

These partner zip codes enable us to test H_1 and H_{1a}. Current focal zip code sales $y_{z,g,t}$ should be affected (1) most strongly by prior sales in the focal (and thus highly visible) zip code and (2) less strongly by prior sales in the geographically near (and thus relatively visible) partner zip codes (H_1). In contrast, prior sales in the geographically far (and thus relatively invisible) partner zip codes should have a null effect on $y_{z,g,t}$ (H_{1a}).

While picking partner zip codes based solely on the basis of distance is sufficient to demonstrate H_1 and H_{1a}, we impose an even stronger test. In addition to imposing the geographic constraints discussed previously, we also choose partner zip codes that are homophilous to the focal zip code. We do so by selecting as the partner zip code the zip code nearest in covariate space to the focal zip code among all potential partner zip codes that meet the geographic constraint (where distance in covariate space is defined as the Mahalanobis distance computed from the variables in Table 3).[2] By matching on covariate distance as well as geographic distance, we select partner zip codes that are not only distant from the focal zip code but also similar or homophilous to it; consequently, finding a null effect for H_{1a} rules out several plausible explanations other than visual influence and therefore bolsters the claim of H_1.

[2]Within each of our four geographic distance bands, the correlation between geographic distance and distance in covariate space is either statistically indistinguishable from zero (contiguous, near, and far) or low (–.15 for moderate). Across the distance bands, the correlation is extremely weak (.05).

Table 2
SUMMARY OF COUNTY VARIABLES

Variable	*25%*	*Median*	*Mean*	*75%*
Car commuters	158,000	248,000	37,000	344,000
Sunny days	92.2	102.0	108.0	112.0
Population	568,000	784,000	1,270,000	1,350,000
Vehicles per household	1.48	1.60	1.51	1.70
Average household income ($)	50,900	56,300	58,700	63,400
Labor force participation rate (%)	58.8	63.1	62.3	65.6
Urban population percentage	92.1	96.7	95.4	99.6

Table 3
SUMMARY OF ZIP CODE VARIABLES

Variable	*25%*	*Median*	*Mean*	*75%*
Population	13,900	24,900	2780	36,800
Urban percentage	97.8	100.00	92.2	100.00
White percentage	52.1	76.2	68.1	88.5
Male percentage	47.4	48.8	48.9	49.9
Child percentage	27.7	32.4	31.4	36.7
Number of households	5250	9310	10,200	13,900
Car to work percentage	80.1	89.2	82.6	93.0
Public transport to work percentage	1.14	3.61	9.02	11.0
Time to work (minutes)	23.4	27.0	27.4	30.6
Median household income ($)	353	46,000	50,100	60,000
Per capita income ($)	16,400	21,800	25,500	29,700
Average age (years)	32.8	35.5	35.9	38.3
Average education (years)	13.1	14.0	14.0	15.1

We note that while matching on all possible covariates would be the strongest test of all, matching on any covariates whatsoever provides a stronger test than matching on none (i.e., other than distance). Furthermore, the covariates we do match on are broad and include population, urbanization, race, household composition, age, commuting patterns, income, and education variables.

Model Specification

Our principal specification for the count nature of the car purchase volumes $y_{z,g,t}$ uses a heterogeneous Poisson[3] probability model, letting $y_{z,g,t} \sim \text{Poisson}(\lambda_{z,g,t})$. The specification for $\lambda_{z,g,t}$ is the heart of our model, and to test our hypotheses, we try alternative specifications (detailed next in the "Visual Effects Specifications" subsection), which allow us to formally decide between models with and without various forms of visual influence effects. We begin by discussing our principal model for $\lambda_{z,g,t}$, which is given by

(1) $\log(\lambda_{z,g,t}) =$

$\alpha_{z,g}$	Zip Code-Group Intercepts
$+ \alpha_{c(z),m(t)}$	County-Month Effects
$+ \beta_{c(z),g,1}t + \beta_{c(z),g,2}t^2 + \beta_{c(z),g,3}t^3$	Time Trends
$+ \epsilon_{z,g,t}$	Error / Overdispersion
$+ \vec{\gamma}_{c(z),g} \times \vec{u}_{z,g,t}$	Car Covariates
$+ v_{focal} + v_{partner}$	Focal and Partner Visual Effects

$$\epsilon_{z,g,t} \sim N(0, \sigma_\epsilon^2). \tag{2}$$

We step through this equation line by line, focusing most of our discussion on the last line, which contains the terms that are most central to our hypotheses.[4]

In the first line, the heterogeneous zip-group intercept parameters $\alpha_{z,g}$ allow each social group in each zip code to buy more or fewer cars per month than the overall average. The $\alpha_{c(z),m(t)}$ parameters in the second line provide each county with its own pattern of seasonality at the monthly level (c(z) is the county in which zip code z is located, and m(t) refers to the calendar month of time t). Next, there is a cubic time trend parameterized by $\vec{\beta}_{c(z),g} = (\beta_{c(z),g,1}, \beta_{c(z),g,2}, \beta_{c(z),g,3})$, which allows for long, secular trends in each county and captures the effects of missing covariates that vary with time.[5] Overdispersion is represented by $\epsilon_{z,g,t}$ in the fourth line and serves to dampen the effect of covariates by widening the standard errors when the data require it. In combination, the zip-group intercepts, seasonality, trends, heterogeneity, and overdispersion of the first four lines also help mitigate the potential effects of missing covariates (Van den Bulte and Lilien 2001).

The fifth line takes account of various variables pertaining to the local car market. In particular, the automobile covariate vector $\vec{u}_{z,g,t}$ contains the level of each of the variables listed in Table 1 averaged across the $y_{z,g,t}$ automobiles sold to group g in zip code z during month t. (We recognize the potential endogeneity of $\vec{u}_{z,g,t}$; for further discussion, see the "Model Evaluation and Robustness" section and Appendix B.) The parameters $\vec{\gamma}_{c(z),g}$ allow each group in each county to react differently to these automobile covariates $\vec{u}_{z,g,t}$. This line of the equation thus adjusts for macro-level variations in salient automobile industry–specific variables (e.g., prices, interest rates) that are not of primary interest in this research.

Visual Effect Specifications

The sixth and final line of Equation 1 is the one of primary interest, and we consider a model design with three levels, resulting in a total of 20 specifications (plus a 21st "null" specification). Before introducing the design, we provide an initial "base" specification, which is given by

$$v_{focal} = \sum_{g'=1}^{G} \delta^{focal}_{c(z),g,g'} y_{z,g',t-1}. \tag{3}$$

This specification posits that current sales in the focal zip code are affected in an *absolute* sense by the previous month's sales in the focal zip code, that the effect is linear, and that there are G^2 effects $\delta_{c(z),g,g'}$ (i.e., the visual effect size depends on the receiving group g *and* the sending group g′).[6] We would find evidence for H_1 when $\delta^{focal}_{c(z),g,g'} \geq 0$.

For this and all other specifications, $v_{partner}$ has an identical form to v_{focal} but with coefficients $\delta^{partner}_{c(z),g,g'}$ in place of $\delta^{focal}_{c(z),g,g'}$ and prior sales in the partner zip code $y_{p(z),g',t-1}$ in place of prior sales in the focal zip code $y_{z,g',t-1}$. Thus, the base specification for $v_{partner}$ is given by $v_{partner} = \Sigma^{G}_{g'=1} \delta^{partner}_{c(z),g,g'} y_{p(z),g',t-1}$. We would find support for H_{1a} when $\delta^{partner}_{c(z),g,g'}$ is statistically no different from zero for far partner zip codes.

The first level of our design recognizes that visual effects may operate on a relative level rather than an absolute level (as in Equation 3). For example, one additional new car on the road may mean something very different in a place where many versus few new cars are sold each month. While our δ parameters have a subscript c(z) and thus allow for variation in the size of visual effects across counties, there is an additional way to allow for variation: rather than using absolute lag sales in our model, we use a relative version, which is standardized at the zip code level. Namely, we set $v_{focal} = \Sigma^{G}_{g'=1} \delta^{focal}_{c(z),g,g'} \tilde{y}_{z,g',t-1}$, where $\tilde{y}_{z,g',t-1} = (y_{z,g',t-1} - m_{z,g})/s_{z,g}$ and $m_{z,g}$ and $s_{z,g}$ are, respectively, the sample mean and standard deviation of sales to group g in zip code z. Again, we use the same relative specification for $v_{partner}$.

[3]While a Poisson model is appropriate for data supported on the nonnegative integers (i.e., count data) such as ours, we also confirmed that our results did not change if we instead employed a Gaussian likelihood (though we modeled $\sqrt{y_{z,g,t} + 1/4}$ rather than $y_{z,g,t}$ in the Gaussian case; Brown, Cai, and DasGupta 2006; DasGupta 2008).

[4]We discuss and lay out the priors in detail in Appendix A. Simply put, we use the standard ones for Bayesian hierarchical models.

[5]Analyses showed that cubics were sufficiently flexible; as a robustness check, we replaced $\alpha_{z,g} + \alpha_{c(z),m(t)} + \beta_{c(z),g,1}t + \beta_{c(z),g,2}t^2 + \beta_{c(z),g,3}t^3$ by zip code/group/time-specific parameters $\alpha_{z,g,s(t)}$, where we set s(t) successively at the annual, semiannual, and quarterly levels. All results remained qualitatively similar.

[6]We note that our hypotheses are about "recent" sales and not the previous month's sales in particular. Although we operationalize recent sales here as the previous month's sales, we tested the robustness of this definition by using the previous quarter's sales as well as the previous six months' sales in place of the previous month's sales, and our results remained qualitatively similar.

The second level of our design recalls the findings of Figure 1—that is, that visual effects may not be linear and may actually be subject to the saturation effects commonly seen in marketing (Dubé, Hitsch, and Manchanda 2005; Johansson 1973; Laurent, Kapferer, and Roussel 1995). Thus, we also consider a logistic specification in addition to our linear one. In this case, we set $v_{focal} = \Sigma_{g'=1}^{G}\delta_{c(z),g,g'}^{focal} \times 1/[1 + \exp(-y_{z,g',t-1})]$ in the absolute case and similarly for $v_{partner}$. For the relative case in the logistic specification, there are two possibilities: a zip code–relative case that, as in the linear case, uses $\tilde{y}_{z,g,t-1}$ in place of $y_{z,g,t-1}$ and a county-relative case that uses $\check{y}_{z,g,t-1} = (y_{z,g,t-1} - m_{c(z),g})/s_{c(z),g}$ in place of $y_{z,g,t-1}$ (where $m_{c(z),g}$ and $s_{c(z),g}$ are, respectively, the sample mean and standard deviation of sales to group g in county c(z); we omit this county-relative case from the linear specification because, in that specification, it coincides with the absolute case).

The final level of our design recognizes that social identity may moderate visual influence effects. Thus far, we have considered the fully parameterized G^2 set of visual effects: The visual effect size is a function of the receiving group g and the sending group g′. We also consider three reduced specifications. In the first case, we restrict visual effects so that they vary only by the receiving group, and thus $\delta_{c(z),g,g'}^{focal} = \delta_{c(z),g}^{focal}$ for all g′ (and similarly here and below for $\delta_{c(z),g,g'}^{partner}$). In the second case, we restrict visual effects so that they vary only by the sending group, and thus $\delta_{c(z),g,g'}^{focal} = \delta_{c(z),g'}^{focal}$ for all g. Finally, we also restrict visual effects so they do not vary by group, and thus $\delta_{c(z),g,g'}^{focal} = \delta_{c(z)}^{focal}$ for all g and g′. We note that rejecting $\delta_{c(z),g,g'}^{focal} = \delta_{c(z)}^{focal}$ for all g and g′ provides evidence for H_3.

The three levels of our design (i.e., absolute vs. relative, linear vs. logistic, G^2 vs. G receiving vs. G sending vs. One visual effects) implies 20 variations of our model. We also consider a 21st null specification, which fixes all $\delta_{c(z),g,g'}^{focal} = \delta_{c(z),g,g'}^{partner} = 0$.

In all specifications, we parameterize the focal zip code visual effects in terms of our zip code–level covariates (i.e., those in Table 2) as is standard in hierarchical Bayesian models with covariates (Lenk et al. 1996). In particular, we set

$$\delta_{c(z),g,g'}^{focal} = \delta_{0,c(z),g,g'}^{focal} + \vec{b} \cdot \vec{x}_c, \quad (4)$$

where x_c is the vector of county covariates for county c. We would find support for H_2 if the b coefficients for the number of people who drive to work and/or number of sunny days are positive.

Is It Really Visual Influence?

While many valid objections can be raised against claims of visual influence, our data and model contain several factors that eliminate several alternative hypotheses. Ruling out alternatives makes visual influence an even more plausible driver of our results. In this section, we discuss the various factors that help eliminate alternatives, including (1) explicit hypotheses and quasi-experimental design, (2) model-based controls, (3) covariate-based controls, and (4) out-of-sample testing.

First, our hypotheses help build the case for visual influence. Suppose we find visual influence from prior sales in nearby, visible zip codes (H_1) and zero effect from our homophilous (i.e., covariate-matched), faraway partner zip codes (H_{1a}). Suppose further that these visual influence effects are stronger in more car-visible places (H_2) and that they vary by sending group, receiving group, and vehicle type in a manner that is consistent with the literature on identity signaling (H_3). It is highly unlikely that our estimates, if they were capturing something other than visual influence, would have these unique features. Rather, for example, a homophily effect would violate H_{1a} and should have no consistent relationship to local indicators of car visibility, social groupings, or vehicles, thus violating H_2 and H_3.

Second, our model-based controls (i.e., the terms in Equation 1 represented by α, β, and ϵ) serve to rule out the effect of correlated unobservables, which might cause one to conclude that visual influence exists even where it does not (Van den Bulte and Lilien 2001). For example, some zip codes or social groups naturally buy more cars (perhaps because they are wealthier or are the targets of greater advertising) and some buy fewer (perhaps they are located in urban areas where people are more likely to use mass transit). Such effects, which can be broadly attributed to missing variables, are controlled by our $\alpha_{z,g}$ terms. The $\alpha_{c(z),m(t)}$ terms account for seasonality—which can look like, but is not, influence—at the local county level.

In addition, the natural flow of population and macroeconomic activity—and whatever else these factors bring with it (e.g., changes in promotional activities)—could cause some counties to be ascendant and others to be in decline over our nine-year period. While such phenomena cannot be captured by our intercept-like α terms (and, for that matter, may not be captured by our covariates), they are captured by our $\vec{\beta}_{c(z),t}$ terms, which function as broad-based catchalls for things happening in a given county.[7]

Our final model-based control is given by $\epsilon_{z,g,t}$, which provides additional heterogeneity and overdispersion. These terms widen standard errors of parameter estimates and thereby diminish the risk of false positives. In addition, some zip code/group/month combinations may be different for unexplained and potentially unexplainable reasons, and this additional heterogeneity helps accommodate this fact.

Third, our covariate-based controls consist of the broad swath of auto industry–specific variables $\vec{u}_{z,g,t}$ listed in Table 1, measured at the zip code level, and parameterized by $\vec{\gamma}_{c(z),t}$. Unlike many retailers that have only a single price variable to manipulate, car dealers have a vast array of price-related variables at their disposal: the price itself, the interest rate on the loan, the length of the loan, the residual value of the car, the trade-in price, various rebates and discounts, and the initial financing amount. All these tools can be used to get the monthly payment to a level that satisfies the consumer. Without adjusting for these variables, which can vary considerably by location and over time, one could indeed "find" a spuriously significant visual effect when one is not present (again, for concerns about the potential endogeneity of $\vec{u}_{z,g,t}$, refer to the "Model Evaluation and Robustness" subsection and Appendix B).

[7]The $\alpha_{z,g,s(t)}$ terms discussed in footnote 5 play an identical role, ruling out heterogeneous zip codes, seasonality, and other unobserved covariates. Because we estimate these down to the zip-group-quarter level, any unobserved variable would have to operate at a very local spatiotemporal level to confound our model.

Although there may be several relevant variables (e.g., local inventory levels) for which we cannot explicitly control, such variables are fortunately very likely to be highly correlated with variables we do control for. For example, if inventories are high, dealers are likely to offer price incentives, thus leading to a correlation between inventory levels and price. Such correlations allow the variables we have to at least partially adjust for these omitted ones. Although this is often undesirable in research (in which, e.g., a researcher is interested in the coefficient on price but is lacking inventory information), in our setting, these correlated variables serve as controls (i.e., we are not interested in their coefficient estimates), and therefore there is much less concern.

Fourth and finally, we use out-of-sample and robustness tests to evaluate our model. If our observed effects are spurious, out-of-sample predictions are likely to degrade substantially. Furthermore, as detailed in "Model Evaluation and Robustness," we provide several additional tests to demonstrate the strength of our model and the robustness of its predictions.

RESULTS

Model Specification Selection

Before examining coefficient estimates and testing our explicit hypotheses, we first choose among the various visual effect specifications of the three-level design discussed in the "Visual Effect Specifications" subsection. Each specification has different implications for the nature of visual influence, and we chose among them using the deviance information criterion (DIC; Spiegelhalter et al. 2002). Table 4 presents DIC values.

As Table 4 indicates, the logistic model with prior sales measured in a county-relative sense and a full ensemble of G^2 visual effect parameters $\delta^{focal}_{c(z),g,g'}$ and $\delta^{partner}_{c(z),g,g'}$ is the "winning" model in terms of DIC. This means that visual effects are subject to saturation, that prior sales should be considered relative to the local area, and that social identity plays a role in effect sizes (e.g., the effect of males on males may be different than the effect of females on males). This last finding provides support for H_3, and we devote further attention to it in the subsection "Social Identity and Vehicle Types (H_3)." Because DIC indicates that the county-relative, logistic, G^2 model performs the best, we focus the remainder of the current study on this model and turn our attention in greater detail to H_1, H_{1a}, H_2, and H_3.

Visual Influence by Distance (H_1)

In Figure 2, we test our visibility hypotheses (H_1 and H_{1a}) by examining our estimates of the hyperparameters $\bar{\delta}^{i}_{g,g'}$

Table 4

DEVIANCE INFORMATION CRITERION

	Linear		*Logistic*		
Visual Effects	*Absolute*	*Zip Relative*	*Absolute*	*Zip Relative*	*County Relative*
G^2	208,435	208,290	208,667	208,257	208,038
G receive	208,435	208307	208,782	208,266	208,126
G send	208,360	208,324	208,616	208,231	208,071
One	208,305	208,351	208,792	208,247	208,121
No visual effects			210,144		

(i.e., the mean of the $\delta^{i}_{c(z),g,g'}$, where i = focal or partner) by distance. The plot has several notable features. First and foremost, the effect of prior sales in the focal zip code has a far larger effect on current sales in the focal zip code than on that in any partner zip code.[8] Second, prior sales in the faraway partner zip code have no effect on current sales in the focal zip code; that is, invisible sales have a null effect. Third, there is a striking monotone decay in the size of the effects: effect size estimates decay as distance increases (and thus visibility decreases). In summary, these results provide support for H_1 and H_{1a}.

Before exploring the effect of visibility more directly (as per Equation 4), we note that our Poisson model with logistic visual effects has nonconstant elasticities (and thus "nonconstant implications"). In particular,

$$(5)\ e^{focal}_{c(z),g,g'} = \frac{\partial \mathbb{E}(y_{z,g,t})}{\partial y_{z,g',t-1}} \times \frac{y_{z,g',t-1}}{\mathbb{E}(y_{z,g,t})} = \frac{\partial \lambda_{z,g,t}}{\partial y_{z,g',t-1}} \times \frac{y_{z,g',t-1}}{\lambda_{z,g,t}}$$

$$= \lambda_{z,g,t} \frac{\partial}{\partial y_{z,g',t-1}} \left\{ \frac{\delta^{focal}_{c(z),g,g'}}{1 + \exp\left[-(y_{z,g,t-1} - m_{c(z),g}) / s_{c(z),g}\right]} \right\} \times \frac{y_{z,g',t-1}}{\lambda_{z,g,t}}$$

$$= \frac{\delta^{focal}_{c(z),g,g'} y_{z,g,t-1}}{s_{c(z),g}} \times \frac{\exp\left[-(y_{z,g,t-1} - m_{c(z),g})/s_{c(z),g}\right]}{\left\{1 + \exp\left[-(y_{z,g,t-1} - m_{c(z),g})/s_{c(z),g}\right]\right\}^2}$$

gives the elasticity of current sales to group g in zip code z in county c with respect to prior sales in zip code z to group g′. (An analogous formula can be derived for the elasticities with respect to prior sales in the partner zip code.)

By plugging the posterior medians illustrated in Figure 2 into Equation 5, we can compute these elasticities for the entire range of lag sales $y_{z,g,t-1}$. We present these results in Figure 3 and note several features. The S-shaped logistic function has a saturation effect, and thus the elasticities show that the incremental visual effect of one additional prior sale is zero for very low and very high values of prior sales. In contrast, the effect seems to be largest at approximately 25 cars sold prior. Another noteworthy feature is that there is much greater variation by sending group than by receiving group: The two elasticity curves that give the effect of female lag sales are barely distinguishable except at the peak; there is a similar pattern for the two that give the effect of male lag sales.

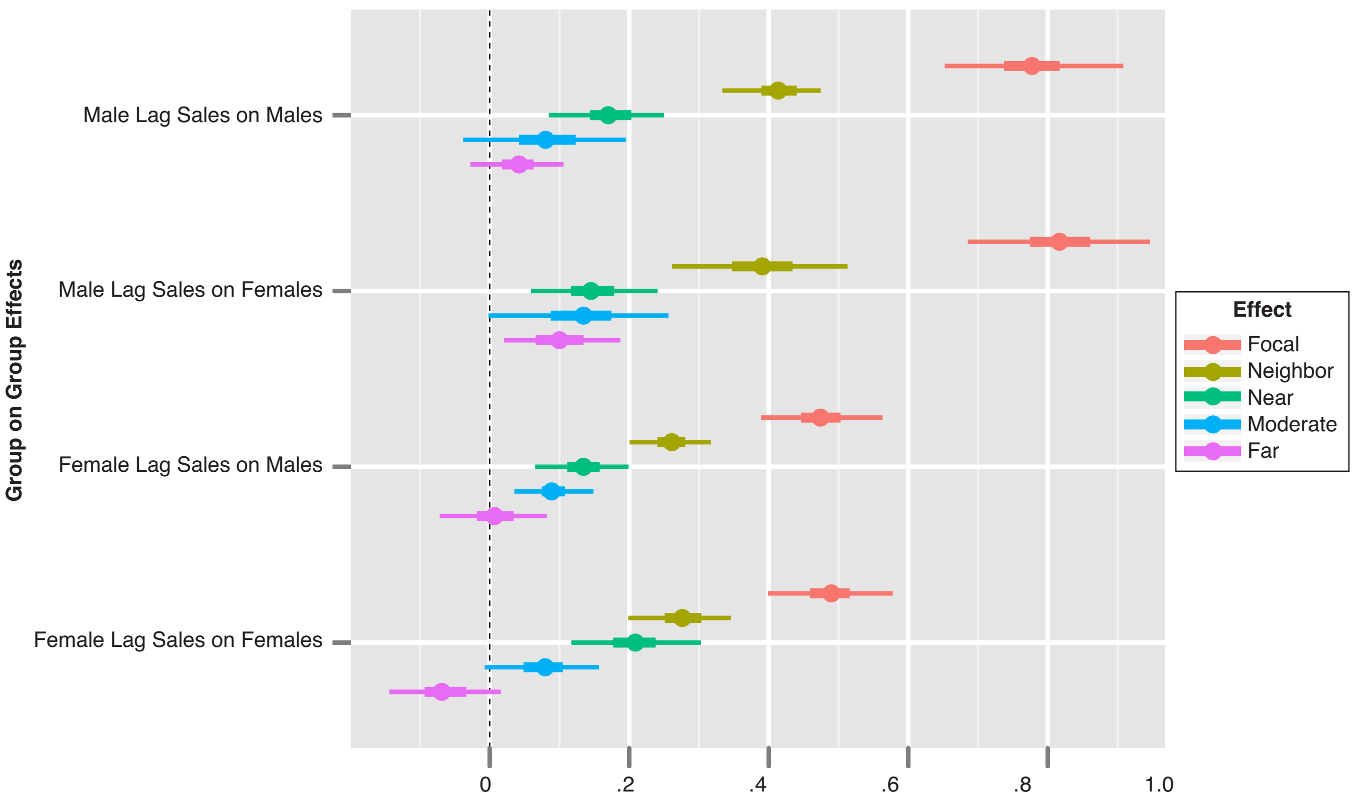
Dosage of Visibility (H_2)

In the previous section, we showed that there is influence from visible cars (i.e., the effect of recent prior sales in the focal and nearby zip codes is positive), but there is no influence from invisible cars (i.e., the effect of recent prior

[8]Although this finding is not of primary interest to us, it does have implications, if empirically general, for other research. In particular, it suggests that distance-based decay functions for visual influence or spatial effects should perhaps either be very rapid or incorporate a nugget-like effect (Banerjee, Carlin, and Gelfand 2004) for zero versus nonzero distances; otherwise, an observed "smooth" process of correlational decay may be somewhat of a falsehood.

Figure 2

HYPERPARAMETERS $\bar{\delta}$g,g' FOR FOCAL AND PARTNER ZIP CODES BY RECEIVING AND SENDING GROUPS



Notes: The posterior median is given by the dot, and posterior 50% and 95% intervals are given by the thick and thin lines, respectively. There is a striking monotone decay in the size of the effects.

sales in the faraway zip codes is null). We now investigate whether the extent of this visual influence can also be tied to the degree of visibility, the subject of H_2.

A benefit of our 40-county data set is that the wide geographic dispersion among counties leads to large variations in their characteristics. Our model is able to account for this variation by estimating a separate effect $\delta^{focal}_{c(z),g,g'}$ for each county c. We present these estimates in Figure 4 and show that there is considerable variation in visual effect sizes across counties. Because we have parameterized these $\delta^{focal}_{c(z),g,g'}$ as a function of the covariates x_c presented in Table 2 (i.e., $\delta^{focal}_{c(z),g,g'} = \delta^{focal}_{0,c(z),g,g'} + \vec{b} \bullet \vec{x}_c$), we can shed light on the effect of these covariates on visual influence by examining the posterior distribution of $\vec{b}$.

In Table 5, we present the posterior mean of $\vec{b}$ along with *p*-values.[9] Beginning with car commuters, a direct measure of car visibility that captures how many people in the county are frequently exposed to other cars, we observe that there is a highly significant and large positive effect. The number of sunny days per year, a measure of a more literal kind of visibility, is also highly statistically significant and positive. As for our control variables, the only ones that are statistically significant are average household income and urban population percentage. Although we have no prediction about the direction of these variables per se, a positive sign for average household income is not particularly surprising because greater income implies a greater ability to respond to visual influence, and a negative sign for urban population percentage is not particularly surprising because increased urbanity is associated with increases in mass transit and thus a lower likelihood of responding to visual influence in the domain of cars.

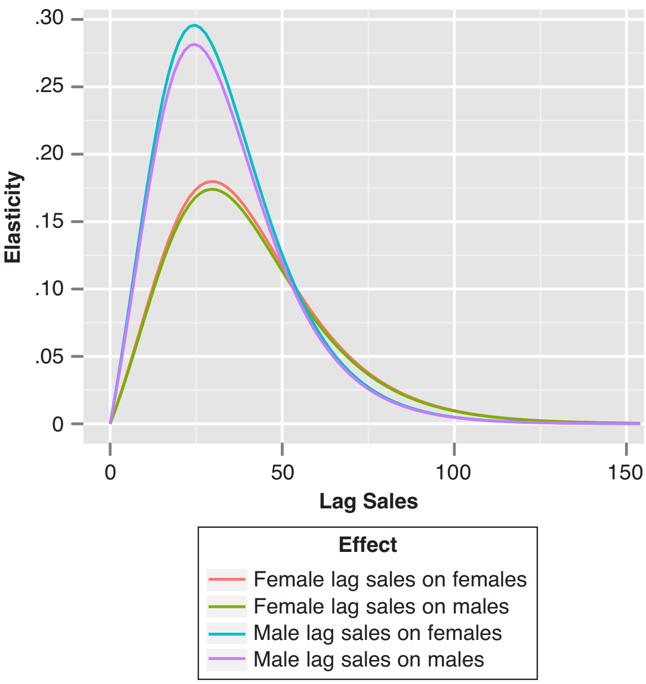
We note that we assessed the robustness of these results to the exact choice of visibility variables. In particular, we reran our model with a several mixes of visibility-related variables (e.g., average temperature rather than number of sunny days, number of mass transit commuters rather than number of car commuters). The robustness of the results is confirmed.

Social Identity and Vehicle Types (H_3)

A considerable amount of research has focused on how social influence is moderated by social identity (Berger and Heath 2008; Simmel 1908; Van den Bulte and Joshi 2007; White and Dahl 2006), though the focus has been on more hedonic or small-stakes settings. Consequently, we sought

[9]Table 5 gives standardized coefficient values. To obtain the raw effect, divide the values in the table by the standard deviations, which are 4.68e + 05, 32.6, 1.65e + 06, .351, 1.26e + 04, .0624, and .0453, respectively.

Figure 3
ELASTICITIES BY RECEIVING AND SENDING GROUPS FOR VARIOUS VALUES OF LAG SALES



Notes: The elasticities are largest at prior sales of approximately 25 cars, and there is greater variation by sending group than by receiving group.

Table 5
POSTERIOR MEAN OF $\vec{b}$

Variable	*Estimate*
Car commuters	.219*
Sunny days	.124*
Population	–.111
Vehicles per household	–.132
Average household income	.117*
Labor force participation rate	.037
Urban population percentage	–.169

*Notes: An asterisk (*) denotes a posterior *p*-value less than .01. All others were greater than .05.

to uncover whether such differences exist in automobile purchases. To examine this, we fit our model separately by the vehicle types (provided in the PIN data), choosing only those types that individually account for more than 10% of aggregate sales (i.e., conventional cars, crossover utility vehicles (CUVs), pickups, and utilities).[10]

To assess the gender identity of each car type, we performed a small survey. Participants (n = 38) were asked how strongly they associated each vehicle type with females versus males. The results appear on the left side of Figure 5. Pickups were very strongly associated with males, and sport utility vehicles were moderately associated with males. In contrast, CUVs had a very weak association with females, while conventional cars were identity neutral.

To test whether (fe)male-oriented vehicles are more affected by prior purchases by males versus females, we computed the difference between the hyperparameters $\overline{\delta}^{\text{focal}}_{g,g'}$ for male versus female sending groups (i.e., we computed $\overline{\delta}^{\text{focal}}_{g,g'=\text{Male}} - \overline{\delta}^{\text{focal}}_{g,g'=\text{Female}}$ separately by vehicle type for each receiving group g). For male-oriented vehicles, this difference should be statistically significantly positive, and, for female-oriented vehicles, it should be statistically significantly negative.

We plot these differences in the right panel of Figure 5; in this figure, vehicles lying to the right of the dashed line at zero are those for which sales to men have a larger effect. As Figure 5 shows, for the male-oriented vehicles type (utilities and pickups), prior sales to men have a statistically significantly greater effect than prior sales to females ($p < .01$). Especially in the male-oriented domain of pickups, the fact that sales to women have a weaker effect shows that people are relatively less likely to buy a male-oriented car when they see a woman driving one. Similarly, for the slightly female-oriented CUV, we observe that prior sales to females have a greater effect than prior sales to males when males are the receiving group ($p < .05$).

As a further test, we can employ a differences-in-differences test using the neutral conventional car as our baseline. That is, we can subtract the conventional car distribution for female receiving groups from the other three vehicle type distributions for female receiving groups and perform the same calculation for males. When conventional cars are used as the baseline, we observe that the only statistically significant difference in difference is for the relative effect of males on males for pickups, the overall most identity-relevant (and male-oriented) vehicle type ($p < .01$).

Thus, our data provide evidence that, not only does visual influence vary by receiving and sending group, it varies in a manner consistent with the nature of social identity. In particular, the visual influence of (fe)males is greater for (fe)male-oriented vehicle types, and there is some evidence that this effect holds more strongly for male receiving groups than for female ones.

Model Evaluation and Robustness

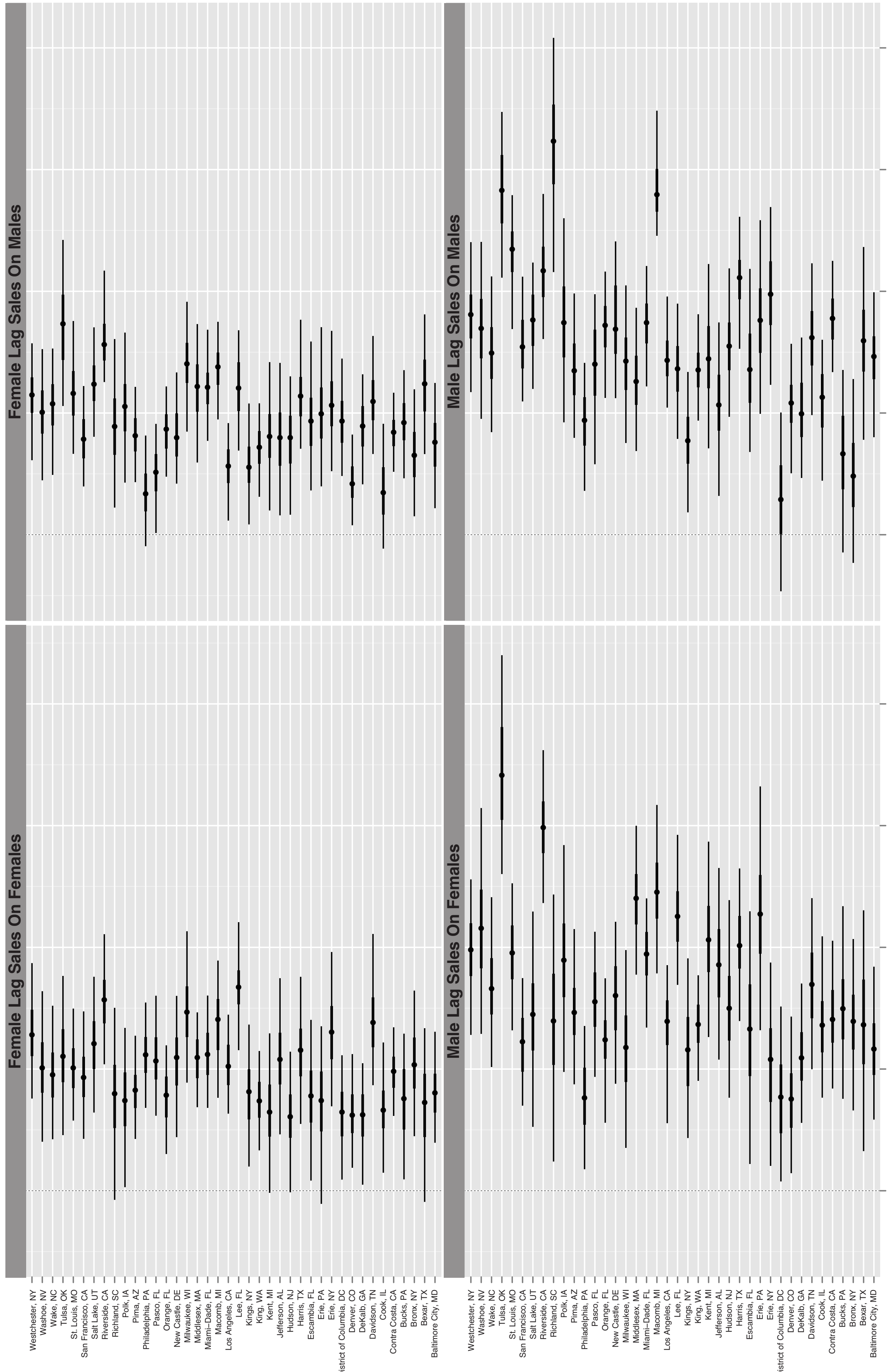
To evaluate our model and demonstrate the robustness of our results, we examined its out-of-sample performance and implemented two separate robustness tests. For our out-of-sample evaluation, we impose a strict test by refitting our model using only the first 92 rather than the full 111 months of data. We hold out 1000 random observations from these first 92 months; consequently, the model is "calibrated" to these time periods, but the observations themselves are out of sample. We also evaluate our model's performance on 1000 random data points from the "uncalibrated" months 93–111 (which are, by definition, out of sample) and 1000 random data points that are entirely in sample.

As Figure 6 indicates, the model appears to fit well with no obvious problems such as departures from linearity regardless of whether the data points come from the in-sample, calibrated out-of-sample, or uncalibrated out-of-sample

[10]On average across all time periods and zip codes, females purchase 3.0 new conventional cars per month and males purchase 3.8; females purchase .8 new CUVs per month and males purchase 1.1; females purchase .4 new pickups per month and males purchase 1.6; and females purchase .9 new CUVs per month and males purchase 1.8. These differences in the baseline rates by group are accounted for by our $\alpha_{z,g}$ terms fit separately for each car type.

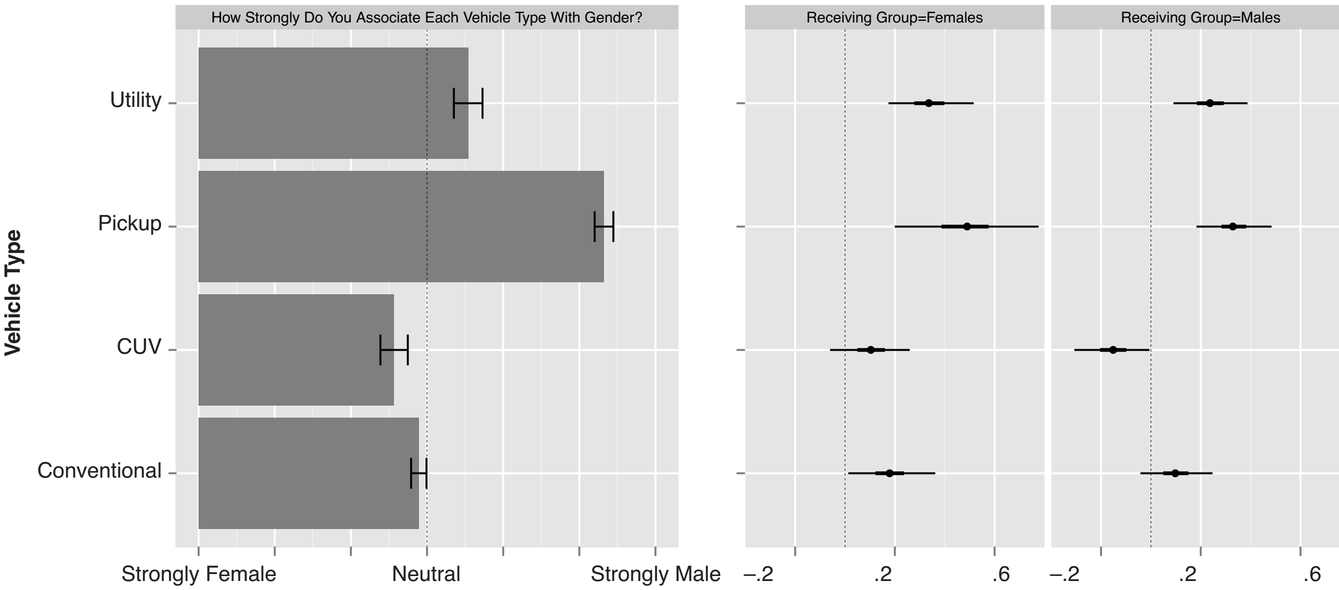
Figure 4

$\delta^{\text{focal}}_{C(z),\, g,\, g'}$ BY RECEIVING AND SENDING GROUPS



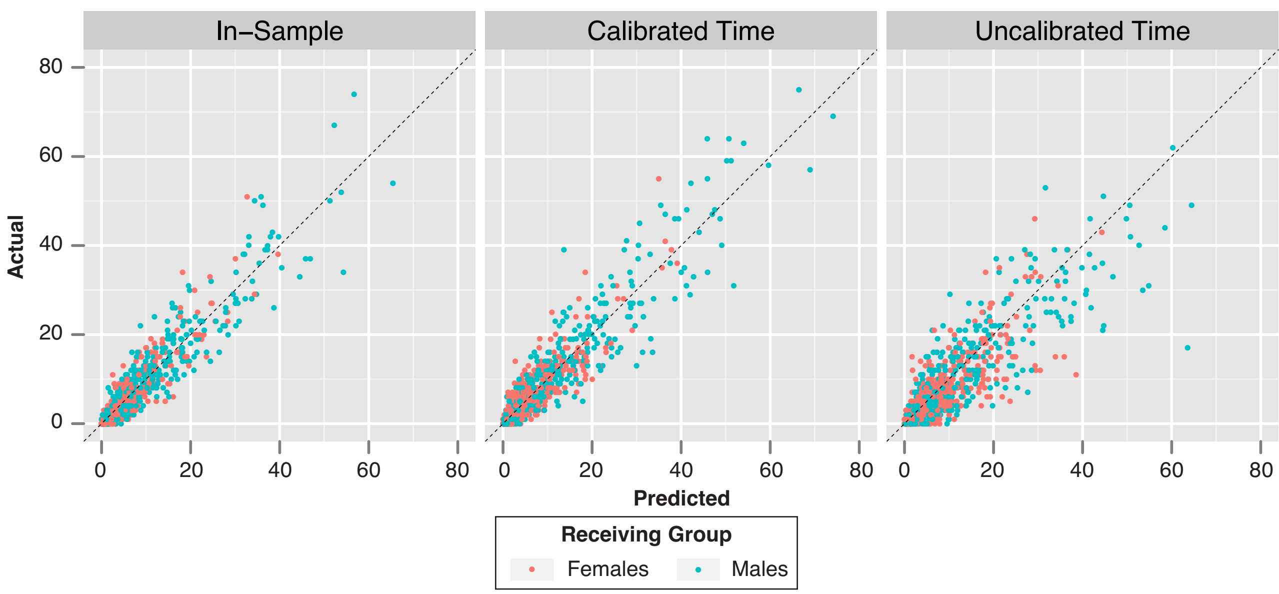
Notes: The posterior median is given by the dot, and posterior 50% and 95% intervals are given by the thick and thin lines, respectively. There is considerable variation in visual effects across counties.

Figure 5
VEHICLE TYPE ANALYSIS



Notes: The left panel indicates how much our survey respondents associated each vehicle type with gender; the length of the bar gives the mean, and the error bar gives ±1 standard error. The right panel shows how much greater prior sales to males versus females affect current sales for each receiving gender (i.e., the plot gives $\overline{\delta}^{\text{focal}}_{g,g'=\text{Male}} - \overline{\delta}^{\text{focal}}_{g,g'=\text{Female}}$ separately by vehicle type); the posterior median is given by the dot, and posterior 50% and 95% intervals are given by the thick and thin lines, respectively. The visual influence of (fe)males is greater for (fe)male-oriented vehicle types.

Figure 6
ACTUAL VERSUS PREDICTED VALUES



Notes: Each point plots y versus $\hat{y}$. The model appears to fit well in all three settings.

periods. Table 6 quantifies fit statistics, and, while performance degrades somewhat in the more difficult uncalibrated time period, the results are still strong. The size of the root mean square error relative to the median absolute error suggests that our model is making large errors on a small number of data points but fits extremely well for the majority of data points.

In addition to the out-of-sample validation, we performed two additional robustness tests. Our first test addresses the possibility that some of our automobile covariates $\vec{u}_{z,g,t}$ (e.g., price, APR) are endogenous. If this is the case, the results we have established (i.e., increased sales associated with increased visible prior sales) could come from a variety of shifts in the underlying demand and supply curves.

Table 6
MODEL FIT STATISTICS

			80% Posterior Interval	
Time	*Root Mean Square Error*	*Median Absolute Error*	*Coverage Percentage*	*Average Width*
In sample	3.28	1	87.5	5.92
Calibrated time	3.83	1	87.7	7.43
Uncalibrated time	5.56	2	80.0	9.49

Notes: The units are given in number of cars.

However, our theory of sales increase is fundamentally a demand-side argument. Therefore, we must show that any result is driven (at least in part) by an outward shift in the demand curve. Our strategy is as follows: short of estimating the parameters of a demand curve, we estimate changes in equilibrium prices and equilibrium quantities associated with changes in prior sales using a true reduced form. If we show that equilibrium quantities *increase* with visibility and that equilibrium prices do *not decrease* as a result of increased prior sales, we can conclude that the demand curve must have shifted out. For a detailed analysis, see Appendix B; to summarize, it is evident that the results indicate an outward shift in the demand curve, thereby providing additional robustness to potential endogeneity.

As a second robustness check, we examined how our model results compared with several well-known results from consumer psychology. In particular, prior research has shown that social influence depends on the relevance or similarity of the other to one's own behavior (Abrams et al. 1990; Berger and Heath 2008; Cialdini 2008; Sherif and Hovland 1961). Similar others are viewed as more diagnostic about what one should do in a given situation, and consequently people conform more to similar than to dissimilar others (Burger et al. 2001). This is particularly true given uncertainty about what the correct behavior is in a given situation.

Applied to our context, we expect that car purchases are more influenced by sales of similarly priced cars than by differently priced cars. Sales of high-priced cars, for example, should be influenced by recent, local sales of high-priced cars but also by recent, local sales of mid-priced cars. They should not be influenced by recent, local sales of low-priced cars. Similarly, for low-priced cars, sales of recent, local low-priced and mid-priced cars should encourage people to buy, but high-priced cars should have less of an effect (in part because the people do not consider high-priced cars a relevant point of influence given their dissimilarity from what they are considering). Because sales of mid-priced cars are close to both high- and low-priced cars, they should be influenced most by themselves, but also by high- and low-priced cars.

To test whether our results conform to this theory, we reran our principal model, but rather than using gender as the social group, we used the price tier of the purchased new cars. Specifically, we considered three price tiers: budget cars costing less than \$15,000, typical cars costing between \$15,000 and \$40,000, and luxury cars costing more than \$40,000 (\$15,000 and \$40,000 correspond to the 10th and 90th percentiles of empirical car prices). We find the same results discussed previously: sales of high-priced cars are influenced by local sales of both high-priced and mid-priced cars but not low-priced cars; sales of mid-priced cars are influenced most strongly by local sales of mid-priced cars but also, to a lesser extent, sales of both high-priced and low-priced cars; and sales of low-priced cars are influenced by local sales of both low-priced and mid-priced cars (all $p <$ .05). Furthermore, relative to the mid-priced cars that serve as the baseline (i.e., 80% of cars), high-priced cars are more strongly influenced by high-priced cars and less strongly influence by mid-priced cars; similarly, low-priced cars are more strongly influenced by low-priced cars (all $p < .05$).

As a final consideration, we note that for spatiotemporal data such ours, an i.i.d. error assumption may not be tenable and that a spatiotemporal covariance structure may be more appropriate. Rather than estimating our ensemble of models with a more general spatiotemporal covariance structure, we instead tested the residuals from our model for evidence of temporal, spatial, and spatiotemporal correlation and found no significant effects.

Managerial Implications

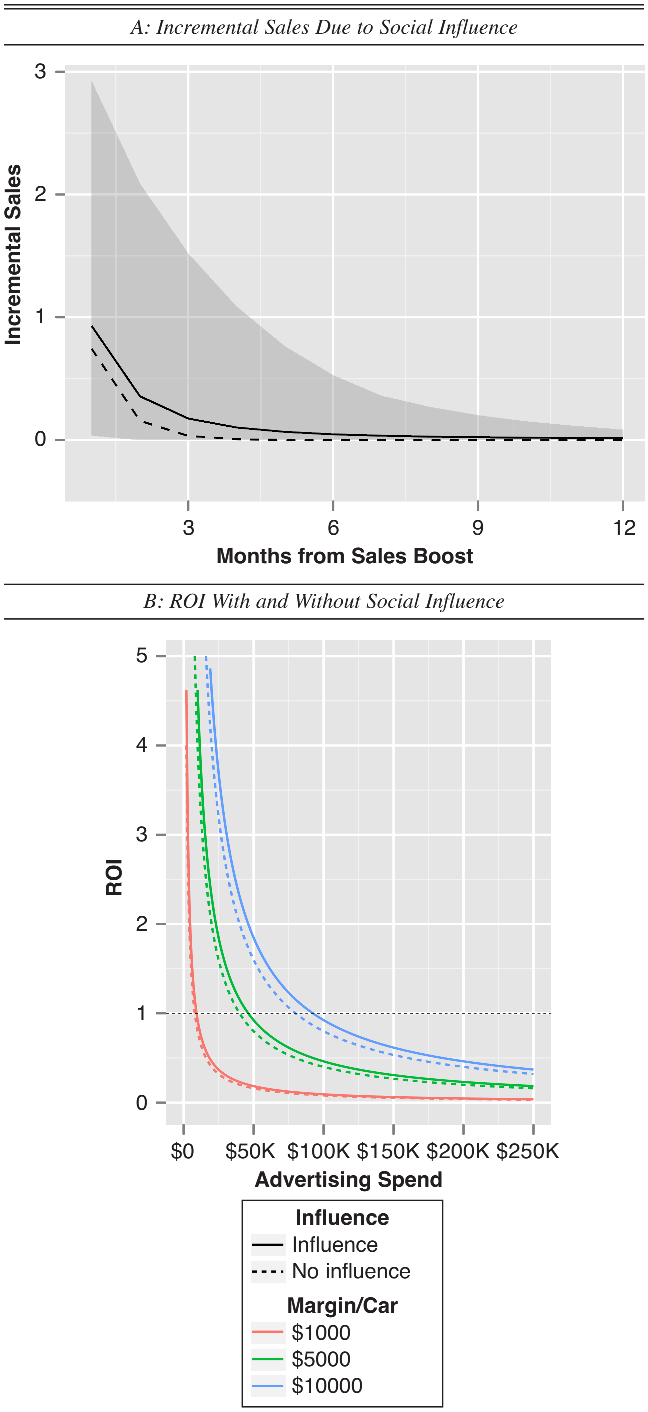
As a final demonstration of the utility of our model, we present some of the managerial implications of our work.[11] Suppose that at month t = 0, a firm engages in additional advertising that leads to four additional women and four additional men purchasing automobiles. If visual influence affects automobile purchases, we would expect these additional purchases at month t = 0 to generate further purchases in months t = 1, 2, ...; our model can estimate this visual influence impulse response function, which we provide in Panel A of Figure 7. Visual influence drives new cars sales up throughout the subsequent year, with the largest effects occurring in the first two months. On average, a total of Δ = 1.24 additional new cars are eventually sold, meaning that visual influence is responsible for roughly one additional purchase for every seven new purchases.

In Panel B of Figure 7, we consider the return on investment (ROI) due to advertising spending in the presence and absence of visual influence. The ROI is equal to $[(c_0 + \Delta) \times m]/a$, where c_0 is the number of new cars sold at time t = 0, m is the margin on each car sold, and a is the increase in advertising required to sell c_0 new cars. We plug in $\Delta = 0$ (no visual influence) and $\Delta = 1.24$ (visual influence), assume three values of m, and let a vary over the x-axis. For reasonable values of m and a, the incremental ROI due to visual influence (i.e., the vertical difference between the "influence" and "no influence" curves of a given margin) is substantial. Moreover, there are nontrivial levels of advertising spending for which marketing ROI is greater than one with visual influence but is less than one without it (i.e., the horizontal difference between where the "influence" and "no influence" curves of a given margin cross y = 1). Thus, if visual influence effects are not built into marketing forecasts, the ROI will be underestimated, and marketing activity may be set suboptimally low.

[11]We note that these effects are at the industry level and are averaged across groups. We also note, on the basis of informal discussions and literature review, that the average amount spent on advertising per automobile sold is on the order of \$1,000–\$1,500 (Busse, Simester, and Zettelmeyer 2010), with margins ranging from less than \$1,000 per car up to several hundred thousand per car for high-end automobiles.

Figure 7

INCREMENTAL SALES AND ROI DUE TO VISUAL INFLUENCE



Notes: In Panel A, we plot the incremental sales due to visual influence in months t = 1, 2, 3... when four additional cars each are sold to women and men at month t = 0; the posterior mean is given by the solid line, the posterior median by the dashed line, and 95% posterior intervals are given in gray. In Panel B, we plot the return on marketing investment with and without visual influence. Visual influence is responsible for roughly one additional purchase for every seven new purchases and thus has a substantial impact on ROI. For supporting details, see the main text.

DISCUSSION AND FURTHER RESEARCH

This research examines the effects of visual influence on the volume of automobile purchases. In particular, we investigate not only whether new car purchase volume is affected by geographically proximate recent purchases but also whether these effects are moderated by car visibility (i.e., how likely people are to see what others are driving in a given geographic area) and whether they vary by receiving groups, sending groups, vehicle types, and price tier. By examining this moderation and variation, we provide further evidence that these effects are driven by visual influence.

We find that people are more likely to purchase a new car if people around them have done so recently. These effects vary according to gender, particularly that of the receiving group. Sales to men have a larger impact on both men and women but more so for male-oriented vehicle types such as pickup trucks; however, for female-oriented CUVs, men are more greatly influenced by women than men. In addition, these effects depend on visibility. Sales of automobiles that consumers do (not) see do (not) influence their purchase behavior. Moreover, variables pertaining to car visibility have statistically significant correlations with our estimates of visual influence, and they vary in the hypothesized direction.

If these effects are due to influence, one might wonder whether they are driven by people observing others' behavior (i.e., visual influence) or by a more network-based process such as word of mouth (Godes and Mayzlin 2004, 2009). Consumers might buy a new car because they see others driving them, for example, or because one of their social ties buys a new one and tells them about it. Although it is difficult to completely tease apart these possibilities in the current investigation, several points suggest that a visual observation story is more likely. First, our results show that the effects are moderated by factors that should affect the visibility of others' behavior (i.e., they are stronger in areas where observing others car purchases should be easier and more likely). While it is possible that these factors also affect aspects of social networks, and thus the likelihood that information spread through word of mouth (Berger and Schwartz 2011), this seems less likely. Second, although a person's social ties do buy new cars every so often, people see the new cars of others they do *not* know (and thus do not speak to) much more frequently. Third, even among those consumers who are influenced to purchase a new car through word of mouth, it is unlikely that most do so the very next month; therefore, because our model captures month-to-month effects, the majority of such consumers are not affecting our results. In summary, while word of mouth may have a stronger effect on behavior, most cars people see on a given day are owned by others they do not know or speak to, and car purchases are generally made with a considerable lag. Consequently, the frequency of visual influence should likely be higher, at least in the domain of cars.

A major contribution to our study is the use of careful statistical and empirical controls. By saturating the model (in a good way) with intercepts, seasonality, time trends, automobile covariates, near and far zip code effects, and heterogeneity, we mitigated the likelihood of finding a spurious effect that was not due to visual influence.

More than providing a test, we have demonstrated effects in automobiles—a product category in which this may be difficult to find beyond niche models (Narayanan and Nair 2011). Functional considerations play a large role in automobile purchases, as do a variety of constraints (e.g., budget, prior leases). Furthermore, new car purchases are high-involvement decisions to which consumers devote a large amount of thought and expend substantial amounts of time on attribute and price research. Thus, while visible and conspicuous, cars are in many ways functional goods, and the very fact that we found an effect for automobiles serves as a further robustness check. It also suggests that visual effects may be larger, and easier to find, in product categories that are not so functional but are equally visible and salient. Further research to identify how visual influence varies across product categories (from, e.g., toilet paper at the low end to, e.g., women's purses at the high end) and individuals (are some people more prone to influence than others and does this correlate with known individual-level variables?) is of definite future interest.

Another subject of further research, which requires more granular data than those used here, would be to examine people within a small geographic area (e.g., a county, a city). Using pinpointed addresses, researchers could track how the effects of a purchase of a particular make and model propagate throughout a neighborhood and determine how this varies across particular makes and models. This would serve to supplement, in a local fashion, the more global and fundamentally different results discussed here. Such work is the focus of much recent research on word-of-mouth effects. Such data might also allow for the consideration of the when and/or what to buy questions (Bucklin, Siddarth, and Silva-Risso 2008; Gupta 1988) and could augment the how much to buy question discussed here.

Understanding and acknowledging the effects of visual influence has the potential to help managers in the design and evaluation of advertising and product launch campaigns, inventory management, brand management, and other crucial areas. In particular, whether these visual influence and social identity effects exist and how strong they are has important implications for consumer decision making and thus marketing efforts. For example, we show that the ROI of an advertising campaign might be substantially underestimated if some sales are due directly to it but others are due indirectly to it through visual influence (Figure 7). Such effects are also relevant for manufacturers and retailers in managing sales, inventory, and distribution, in which the ability to forecast sales volume drives profitability.

In conclusion, our research suggests that new car purchases are shaped by visual influence. The models presented here help shed light on exactly how these effects work and how they vary by car visibility, social groups, and vehicle types. Further process-oriented experiments to uncover the psychological mechanism would be a worthwhile supplement to the effects we observe here.

APPENDIX A: PRIORS AND IMPLEMENTATION

Because social groups within various zip codes and counties are likely to have heterogeneous preferences for new car purchase volume, we allow the parameters of our model to vary at the county-group level (or below). We embed these assumptions into a Bayesian hierarchical model by allowing each county-group's parameters to come from a common distribution, allowing our model to "borrow strength" across all county-group combinations. The likelihood is specified in the main text (see the "Model Specification" and "Visual Effect Specification" subsections and, in particular, Equation 1). In this Appendix, we specify the priors and the sampling procedure. In general, we use standard noninformation priors. Although we assume a priori independence of various parameters, as is common, our Bayesian model allows for a posteriori dependence between these components as induced by the data.

We begin with the zip code group intercept terms, which we give a normal prior:

$$\alpha_{z,g} \sim \text{Normal}(0, 100^2). \tag{A1}$$

We note that a standard deviation of 100 on the log scale is extremely uninformative. We use a similar formulation for the monthly seasonal effects:

$$\alpha_{c(z),m(t)} \sim \text{Normal}(\overline{\alpha}_{m(t)}, \sigma_m^2) \tag{A2}$$

$$\overline{\alpha}_{m(t)} \sim \text{Normal}(0, 100^2), \text{ and} \tag{A3}$$

$$\sigma_m \sim \text{Uniform}(0, 100). \tag{A4}$$

Each county has its own set of monthly seasonal intercepts. However, a given month, say September, is shrunk to a common mean shared by all counties. Furthermore, for identifiability, we set $\overline{\alpha}_{\text{January}} = 0$ and $\alpha_{c(z),\text{January}} = 0$ for all counties.

For the $\vec{\beta}_{c(z),g} = (\beta_{c(z),g,1}, \beta_{c(z),g,2}, \beta_{c(z),g,3})$ parameters governing the cubic time trend, we also use normal priors with shrinkage to the group-level means as given by

$$\beta_{c(z),g,i} \sim \text{Normal}(\overline{\beta}_i, \sigma_{\beta,i}^2), \tag{A5}$$

$$\overline{\beta}_i \sim \text{Normal}(\vec{0}, 100^2), \text{ and} \tag{A6}$$

$$\sigma_{\beta,i} \sim \text{Uniform}(0, 100), \tag{A7}$$

where $i \in \{1, 2, 3\}$. The formulation for the elements of $\vec{\gamma}_{c(z),g}$ vector is identical *mutatis mutandis*.

Recall that we parameterize the $\delta^{\text{focal}}_{c(z),g,g'}$ as $\delta^{\text{focal}}_{c(z),g,g'} = \delta^{\text{focal}}_{0,c(z),g,g'} + \vec{b} \times \vec{x}_c$. The prior for the elements of $\vec{b}$ is given by

$$b_i \sim \text{Normal}(0, 100^2), \tag{A8}$$

whereas the $\delta^{\text{focal}}_{0,c(z),g,g'}$ have the prior given by

$$\delta^{\text{focal}}_{0,c(z),g,g'} \sim \text{Normal}(\overline{\delta}^{\text{focal}}_{g,g'}, \sigma^2_{g,g'}), \tag{A9}$$

$$\overline{\delta}^{\text{focal}}_{g,g'} \sim \text{Normal}(\vec{0}, 100^2), \text{ and} \tag{A10}$$

$$\sigma_{g,g'} \sim \text{Uniform}(0, 100). \tag{A11}$$

The prior for the $\delta^{\text{partner}}_{c(z),g,g'}$ is identical *mutatis mutandis* to the prior for the $\delta^{\text{focal}}_{0,c(z),g,g'}$ given previously. When implementing the third level of the design presented in the "Visual Effect Specification" subsection, we set (1) $\delta^{\text{focal}}_{0,c(z),g,g'} = \delta^{\text{focal}}_{0,c(z),g}$, $\overline{\delta}^{\text{focal}}_{g,g'} = \overline{\delta}^{\text{focal}}_{g}$, and $\sigma_{g,g'} = \sigma_g$ for all g' in the case that visual effects only vary by receiving group; (2) $\delta^{\text{focal}}_{0,c(z),g,g'} = \delta^{\text{focal}}_{0,c(z),g'}$, $\overline{\delta}^{\text{focal}}_{g,g'} = \overline{\delta}^{\text{focal}}_{g'}$, and $\sigma_{g,g'} = \sigma_{g'}$ for all g in the case that visual effects only vary by sending group; and (3) $\delta^{\text{focal}}_{0,c(z),g,g'} = \delta^{\text{focal}}_{0,c(z)}$, $\overline{\delta}^{\text{focal}}_{g,g'} = \overline{\delta}^{\text{focal}}$, and $\sigma_{g,g'} = \sigma$ for all g, g' in the case that visual effects do not vary by either receiving or sending group.

Finally, we have $\epsilon_{z,g,t} \sim \text{Normal}(\vec{0}, \sigma^2_{\epsilon,})$, as discussed in the main text. This requires a prior for σ^2_ϵ, and we use $\sigma_\epsilon \sim$ Uniform(0, 100).

We sample from the full posterior distribution using Markov chain Monte Carlo (Chib and Greenberg 1995; Gelfand 1996; Gelman et al. 2003). We implement the Markov chain Monte Carlo algorithm in WinBUGS (Spiegelhalter, Thomas, and Best 1999) running four independent chains each for 7500 iterations, discarding the first 5000 as burn-in, and thinning every 10 iterations.

APPENDIX B: REDUCED FORM MODELS

In this Appendix, we describe the robustness test we performed to address the possibility that some of our automobile covariates $\vec{u}_{z,g,t}$ (e.g., price, APR) are endogenous. These tests, which require estimation of reduced form models, serve to demonstrate that equilibrium quantities increase with visible prior sales and that equilibrium prices do not decrease as a result of increased visible prior sales. Consequently, we can conclude that the demand curve must have shifted out and thus that the observed increase in sales associated with increased visible prior sales is driven at least in part by demand-side phenomena.

A completely generic reduced form estimates the following quantity and price equations:

$$\text{(B1)} \qquad Q = a_0 + \vec{a}_1 X^D + \vec{a}_2 X^S + \nu$$
$$P = b_0 + \vec{b}_1 X^D + \vec{b}_2 X^S + \eta$$

where P is price, Q is quantity, X^D and X^S are demand and supply covariates respectively, and ν and η are error terms. The estimates $\hat{a}$ and $\hat{b}$ do not serve to estimate the effects of the covariates on the underlying supply and demand curves; rather, they estimate the effect of each covariate on the equilibrium quantity and equilibrium price, respectively.

As we discussed in the "Model" section, our α and β parameters capture a variety of demand covariates. In contrast, the supply curve should respond to changes in costs of productions (e.g., raw materials, labor, energy). We believe there is little variation beyond that captured by zip-group intercepts, county-group seasonality, and time trends. Indeed, existing research finds that "executives responsible for short- to medium-run manufacturing and pricing decisions for automobiles indicate that, in practice, these decisions are not made on the basis of small changes to manufacturing costs" (Busse, Knittel, and Zettelmeyer 2012).

In our case, quantity Q is the variable $y_{z,g,t}$ analyzed throughout. Thus, the quantity equation is a special case of our main quantity model (i.e., Equation 1 with the G^2 logistic, county-relative effects selected by DIC). In particular, we have a reduced form Q equation when we reestimate this model with $\vec{\gamma}_{c(z),g}$ (which parameterizes the $\vec{u}_{z,g,t}$) fixed at $\vec{0}$. We note that, technically, we need only to fix at zero the components of $\vec{\gamma}_{c(z),g}$ that parameterize endogenous covariates; in our analysis, we performed the strictest test and assumed that all elements of $\vec{u}_{z,g,t}$ were potentially endogenous and thus fixed the whole $\vec{\gamma}_{c(z),g}$ vector to $\vec{0}$.

In typical settings, there is a single price P and thus a single P equation. However, our unique automobile setting features many price-related variables P, including the stated price itself, the manufacturer's rebate, and the APR. Indeed, as mentioned previously, we assume that potentially all of the nine $\vec{u}_{z,g,t}$ variables (i.e., the variables listed in Table 1) could be endogenous. Consequently, for each $u_{z,g,t,i}$, where i ranges over the nine variables, we use a right-hand side reduced form specification that is analogous to that used for Q. Namely, we estimate

$$\text{(B2)} \qquad u_{z,g,t,i} = \alpha_{z,g} + \alpha_{c(z),m(t)} + \beta_{c(z),g,1}t + \beta_{c(z),g,2}t^2 + \beta_{c(z),g,3}t^3 + v_{focal} + v_{partner} + \epsilon_{z,g,t},$$

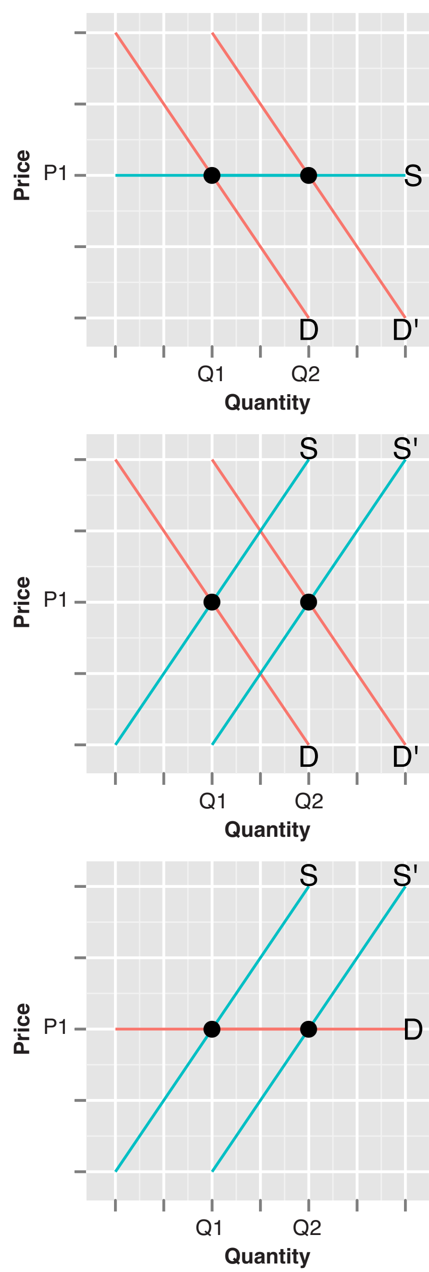
with $\epsilon_{z,g,t} \sim N(0, \sigma^2)$.

When we estimate these models, we find the following: First, the estimated $\bar{\delta}^{focal}_{g,g'}$ and $\bar{\delta}^{partner}_{g,g'}$ for the quantity model are nearly identical to those estimated without forcing $\vec{\gamma}_{c(z),g} = 0$ (i.e., Figure 2 looks the same regardless of whether $\vec{\gamma}_{c(z),g}$ is forced to $\vec{0}$.). Second, the estimated $\bar{\delta}^{focal}_{g,g'}$ and $\bar{\delta}^{partner}_{g,g'}$ for each $u_{z,g,t,i}$ are statistically no different than zero (i.e., a plot such as Figure 2 for these coefficients would show all coefficients overlapping zero). Together, these results imply that (1) equilibrium quantity increased and (2) the various variables pertaining to equilibrium price remained approximately the same.

There are three supply and demand curve configurations that are consistent with this result, and they are depicted in Figure B1. Under the first, the shift in quantity is purely

Appendix B1

THE THREE SUPPLY AND DEMAND CURVE CONFIGURATIONS CONSISTENT WITH REDUCED FORM ESTIMATES



demand driven. Under the second, there are both demand and supply effects. Under the third, there are only supply effects. Because we can rule out the third scenario (i.e., a horizontal demand curve is implausible in the market for all cars), we can confidently conclude that our observed effects are driven, at least in part, by demand.[12]

REFERENCES

Abrams, Dominic, Margaret Wetherell, Sandra Cochrane, Michael A. Hogg, and John C. Turner (1990), "Knowing What to Think by Knowing Who You Are: Self-Categorization and the Nature of Norm Formation, Conformity and Group Polarization," *British Journal of Social Psychology*, 29 (2), 97–119.

Asch, Solomon Eliot (1956), "Studies of Independence and Conformity: A Minority of One Against a Numerous Majority," *Psychological Monographs*, 70 (416), 1–70.

Banerjee, Abhijit V. (1992), "A Simple Model of Herd Behavior," *Quarterly Journal of Economics*, 107 (3), 797–817.

Banerjee, Sudipto, Bradley P. Carlin, and Alan E. Gelfand (2004), *Hierarchical Modeling and Analysis for Spatial Data*. Boca Raton, FL: Chapman and Hall/CRC.

Bell, David R. and Sangyoung Song (2007), "Neighborhood Effects and Trial on the Internet: Evidence from Online Grocery Retailing," *Quantitative Marketing and Economics*, 5 (4), 361–400.

Bem, Sandra L. (1981), "Gender Schema Theory: A Cognitive Account of Sex Typing," *Psychological Review*, 88 (4), 354–64.

Berger, Jonah and Chip Heath (2007), "Where Consumers Diverge from Others: Identity-Signaling and Product Domains," *Journal of Consumer Research*, 34 (August), 121–34.

——— and ——— (2008), "Who Drives Divergence? Identity Signaling, Outgroup Dissimilarity, and the Abandonment of Cultural Tastes," *Journal of Personality and Social Psychology*, 95 (3), 593–607.

——— and Lindsay Rand (2008), "Shifting Signals to Help Health: Using Identity Signaling to Reduce Risky Heath Behaviors," *Journal of Consumer Research*, 32 (5), 509–518.

——— and Eric Schwartz (2011), "What Drives Immediate and Ongoing Word of Mouth?" *Journal of Marketing Research*, 48 (October), 869–80.

Bikhchandani, Sushil, David Hirshleifer, and Ivo Welch (1992), "A Theory of Fads, Fashion, Custom and Cultural Change as Informational Cascades," *Journal of Political Economy*, 100 (5), 992–1926.

Brown, Larry, Tony Cai, and Anirban DasGupta (2006), "On Selecting an Optimal Transformation," technical report, University of Pennsylvania.

Bucklin, Randolph E., Sivaramakrishnan Siddarth, and Jorge M. Silva-Risso (2008), "Distribution Intensity and New Car Choice," *Journal of Marketing Research*, 45 (August), 473–86.

Burger, Jerry M., Shelley Soroka, Katrina Gonago, Emily Murphy, and Emily Somervell (2001), "The Effect of Fleeting Attraction on Compliance to Requests," *Personality Social Psychology Bulletin*, 27 (12), 1578–86.

Burnkrant, Robert E. and Alain Cousineau (1975), "Informational and Normative Social Influence in Buyer Behavior," *Journal of Consumer Research*, 2 (3), 206–215.

Busse, Meghan, Christopher R. Knittel, and Florian Zettelmeyer (2012), "Pain at the Pump: The Effect of Gasoline Prices on New and Used Automobile Markets," *American Economic Review*, forthcoming.

———, Duncan Simester, and Florian Zettelmeyer (2010), "The Best Price You'll Ever Get: The 2005 Employee Discount Pricing Promotions in the U.S. Automobile Industry," *Marketing Science*, 29 (2), 268–90.

Cai, Hongbin, Yuyu Chen, and Hanming Fang (2009), "Observational Learning: Evidence from a Randomized Natural Field Experiment," *American Economic Review*, 99 (3), 864–82.

Chib, Siddhartha and Edward Greenberg (1994), "Understanding the Metropolis Algorithm," *The American Statistician*, 49 (4), 327–35.

Cialdini, Robert B. (2008), *Influence: Science and Practice*. Boston: Pearson.

DasGupta, Anirban (2008), *Asymptotic Theory of Statistics and Probability*. New York: Springer.

Dasgupta, Srabana, Sivaramakrishnan Siddarth, and Jorge M. Silva-Risso (2007), "Lease or Buy: A Structural Model of the Vehicle Acquisition Decision," *Journal of Marketing Research*, 44 (August), 490–502.

Deaux, Kay, Anne Reid, Kim Mizrahi, and Kathleen A. Ethier (1995), "Parameters of Social Identity," *Journal of Personality and Social Psychology*, 68 (2), 280–91.

Deutsch, Morton and Harold B. Gerard (1955), "A Study of Normative and Informational Social Influences upon Individual Judgment," *Journal of Abnormal Psychology*, 51 (3), 629–36.

Dubé, Jean-Pierre, Günter J. Hitsch, and Puneet Manchanda (2005), "An Empirical Model of Advertising Dynamics," *Quantitative Marketing and Economics*, 3 (2), 107–144.

Gelfand, Alan E. (1996), "Model Determination Using Sampling Based Methods," in *Markov Chain Monte Carlo in Practice*. Boca Raton, FL: Chapman and Hall (CRC), 142–62.

Gelman, Andrew, John B. Carlin, Hal S. Stern, and Donald B. Rubin (2003), *Bayesian Data Analysis*, 2d ed. Boca Raton, FL: Chapman and Hall/CRC.

Godes, David and Dina Mayzlin (2004), "Using Online Conversations to Study Word-of-Mouth Communication," *Marketing Science*, 23 (4), 545–60.

——— and ——— (2009), "Firm-Created Word-of-Mouth Communication: Evidence from a Field Test," *Marketing Science*, 28 (4), 721–39.

Grinblatt, Mark, Matti Keloharrju, and Seppo Ikaheimo (2008), "Social Influence and Consumption: Evidence from the Automobile Purchases of Neighbors," *The Review of Economics and Statistics*, 90 (4), 735–53.

Gupta, Sunil (1988), "Impact of Sales Promotions on When, What, and How Much to Buy," *Journal of Marketing Research*, 25 (November), 342–55.

Heffetz, Ori (2011), "A Test of Conspicuous Consumption: Visibility and Income Elasticities," *Review of Economics and Statistics*, 93 (4), 1101–1117.

Johansson, Johny K. (1973), "A Generalized Logistic Function with an Application to the Effect of Advertising," *Journal of the American Statistical Association*, 68 (344), 824–27.

Joshi, Yogesh V., David J. Reibstein, and Z. John Zhang (2009), "Optimal Entry Timing in Markets with Social Influence," *Management Science*, 55 (6), 926–39.

Laurent, Gilles, Jean-Noel Kapferer, and Françoise Roussel (1995), "The Underlying Structure of Brand Awareness Scores," *Marketing Science*, 14 (3), G170–G179.

Lenk, Peter J., Wayne S. DeSarbo, Paul E. Green, and Martin R. Young (1996), "Hierarchical Bayes Conjoint Analysis: Recovery of Partworth Heterogeneity from Reduced Experimental Designs," *Marketing Science*, 15 (2), 173–91.

Manski, Charles F. (1993), "Identification of Endogenous Social Effects: The Reflection Problem," *Review of Economic Studies*, 60 (3), 531–42.

Narayanan, Sridhar and Harikesh S. Nair (2011), "Estimating Causal Installed-Base Effects: A Bias-Correction Approach," technical report, Stanford University Graduate School of Business.

Rossi, Peter E., Greg M. Allenby, and Rob McCulloch (2003), "Bayesian Statistics and Marketing," *Marketing Science*, 22 (3), 304–328.

[12]We caution the reader not to interpret the visibility finding as necessarily arising *solely* from a shift in the demand curve. While this is one possibility, our results are also consistent with shifts in both the demand and supply curves.

Scott-Morton, Fiona, Florian Zettelmeyer, and Jorge Silva-Risso (2001), "Internet Car Retailing," *Journal of Industrial Economics*, 49 (4), 501–519.

———, ———, and ——— (2003), "Consumer Information and Price Discrimination: Does the Internet Affect the Pricing of New Cars to Women and Minorities?" *Quantitative Marketing and Economics*, 1 (1), 65–92.

Shang, Jen, Americus Reed II, and Rachael Croson (2008), "Identity Congruency Effects on Donations," *Journal of Marketing Research*, 45 (June), 351–61.

Sherif, Muzafer and Carl I. Hovland (1961), *Social Judgment: Assimilation and Contrast Effects in Communication and Attitude Change*. New Haven, CT: Yale University Press.

Shriver, Scott K. (2010), "Network Effects in Alternative Fuel Adoption: Empirical Analysis of the Market for Ethanol," technical report, Stanford University Graduate School of Business.

Silva-Risso, Jorge and Irina Ionova (2008), "A Nested Logit Model of Product and Transaction-Type Choice for Planning Automakers' Pricing and Promotions," *Marketing Science*, 27 (4), 545–66.

Simmel, Georg (1908), "Fashion," *International Quarterly*, 10 (1),130–35.

Spiegelhalter, David J., Nicola G. Best, Bradley P. Carlin, and Angelika van der Linde (2002), "Bayesian Measures of Model Complexity and Fit," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 64 (4), 583–639.

———, Andrew Thomas, and Nicky G. Best (1999), *WinBUGS Version 1.2 User Manual*. Cambridge, UK: MRC Biostatistics Unit.

Srinivasan, Shuba, Koen Pauwels, Jorge M. Silva-Risso, and Dominique M. Hanssens (2004), "Product Innovations, Promotions, and Long-Term Firm Value," *Journal of Marketing*, 68 (October),142–56.

Van den Bulte, Christophe and Yogesh V. Joshi (2007), "New Product Diffusion with Influentials and Imitators," *Marketing Science*, 26 (3), 400–421.

——— and Gary Lilien (2001), "Medical Innovation Revisited: Social Contagion Versus Marketing Effort," *American Journal of Sociology*, 106 (5), 1409–1435.

White, Katherine and Darren Dahl (2006), "To Be or Not Be? The Influence of Dissociative Reference Groups on Consumer Preferences," *Journal of Consumer Psychology*, 16 (4), 404–414.

——— and ——— (2007), "Are All Outgroups Created Equal? Consumer Identity and Dissociative Influence," *Journal of Consumer Research*, 34 (4), 525–36.