

**Driving Online and Offline Sales:  
The Cross-Channel Effects of Digital versus Traditional Advertising**

November 6, 2011

Isaac M. Dinner  
Assistant Professor of Marketing  
IE Business School,  
Calle de Serrano 105, Madrid, Spain, [isaac.dinner@ie.edu](mailto:isaac.dinner@ie.edu)

Harald J. van Heerde  
Professor of Marketing  
University of Waikato, Department of Marketing,  
Private Bag 3105, Hamilton 3240, New Zealand, [heerde@waikato.ac.nz](mailto:heerde@waikato.ac.nz)

Scott A. Neslin  
Albert Wesley Frey Professor of Marketing  
Dartmouth College, Tuck School of Business,  
Dartmouth College, Hanover, NH, USA, [scott.neslin@dartmouth.edu](mailto:scott.neslin@dartmouth.edu)

*Acknowledgments.* We are grateful for the constructive feedback received from participants at the presentation of this paper at the 2011 Marketing Science Conference (Rice University) and at marketing seminars at the University of North Carolina at Chapel Hill, the University of Pittsburgh, and the University of New South Wales. We also thank Don Lehmann and Carl Mela for their excellent suggestions on an earlier draft of this paper. We would also like to thank an anonymous retailer for sharing their marketing and transaction data with us, and to Sha Yang for her insights and for sharing unpublished data regarding the effectiveness of search advertising.

**Driving Online and Offline Sales:  
The Cross-Channel Effects of Digital versus Traditional Advertising**

*ABSTRACT*

Today's marketing environment is characterized by a surge in multichannel shopping and ever more choice in advertising channels. This requires firms to understand how both digital and traditional advertising drive sales within the same channel (e.g., digital advertising affecting online sales) and across channels (e.g., digital advertising affecting offline sales). We develop a Dynamic Linear Model (DLM) to measure these effects. The model addresses: (1) the endogeneity of advertising, (2) dynamic advertising effects, (3) a multivariate dependent variable (4) heterogeneity across markets, and (5) competitive advertising effects. It decomposes advertising's impact into customer counts and spend. We calibrate the model using data from a large, upscale retailer. We estimate elasticities for traditional (offline), online display (banner) and online search advertising on online and offline sales. Further, we develop and test hypotheses on how advertising impacts own- and cross-channel sales. We find that cross-channel effects exist and are important. For example, much of the impact of digital advertising can be attributed to its effect on the offline channel, primarily because of the impact on customer count.

**KEYWORDS:** advertising elasticity, search advertising, banner advertising, multi-channel, cross effects, bayesian methods, dynamic linear model, return on investment

## *INTRODUCTION*

Many brick and mortar retailers have opened online stores and found these stores dramatically increase sales. For example, while Macy's June 2011 year-to-year revenue growth in physical stores was reported to be 6.7%, online sales increased by 45% (Macy's Press Release 2011). This growth in the online sales channel has been rivaled by the growth in online advertising spending. That is, while the global financial crisis forced most companies to cut their marketing budgets (McKinsey 2009; The Financial Times 2008), internet advertising grew 15% from 2009 to 2010, to the point where online advertising spend was over \$26 Billion in 2010 (IAB 2010).

The proliferation of the online purchase channel, combined with dual advertising outlets across online and offline channels, creates opportunities but also an increasingly complex problem for retailers. Firms like Macy's now have to balance advertising across traditional media (television, print, etc.) with more recent online techniques such as search and display advertising (banner ads) to propel purchases across both physical and internet stores.

While many marketing campaigns are put in place to drive sales to their native channel (e.g. email campaigns aim to drive online sales (Ansari et al. 2008)), it is natural to consider the impact that advertising will have across channels. Specifically, online advertising may have an impact on offline sales (a "cross effect") as well as an impact on online sales (an "own effect"). This raises questions as to what percentage of the sales impact will be native and what will be cross-channel. Is the cross-effect of online advertising on offline sales smaller or larger than the impact of traditional advertising on online sales?

From a managerial perspective, knowing the full return from all advertising expenditures is necessary to developing an efficient marketing plan. So, estimating the cross-channel effects

of advertising is crucial as 1) solely measuring own channel effects may undervalue advertising effects and 2) understanding the size of cross-effects may have strong implications for setting strategy across channels. For example, if a retailer wants to build sales at a physical store, it may in fact be better to utilize online advertising.

While the cross-channel effects of advertising are important there is also a need to understanding *how* advertising impacts cross-channel sales. Will advertising increase the number of customers drawn to a channel or simply generate more revenue per customer? Answers to this question will help firms understand which tools they can use to increase channel “traffic” and which they can use to make the average purchase larger and more profitable.

The purpose of this paper is to develop a model to estimate sales in both online and offline channels as a function of online and traditional advertising. The model allows us to estimate the total sales elasticity of traditional, online search and online display advertising, as well as decompose this elasticity into advertising’s impact on customer counts and average customer spend. By measuring the relative effectiveness of digital advertising media compared with traditional media, our model addresses one of the key research priorities identified by the Marketing Science Institute (2010). Our task is particularly challenging for a number of reasons, including (i) endogeneity of advertising, (ii) dynamic advertising effects, (iii) a multivariate dependent variable, (iv) heterogeneity across markets and (v) competitive advertising effects. To tackle these challenges we create a Dynamic Linear Model (DLM) (e.g., West and Harrison 1999, p. 284; Ataman et al. 2008; Ataman et al. 2010; Van Heerde et al. 2004) and estimate it using Bayesian techniques.

There are five central contributions to this paper. First, we develop a unique model to confront the issue of estimating the influence of multiple advertising media across channels.

Second, we decompose the total sales impact into customer counts and average customer spend. Third, we apply theory on decision making in a multichannel environment to develop hypotheses about how advertising will impact cross-channel sales. Fourth, we contribute to the knowledge base by estimating the model and testing our hypotheses on a large, upscale retailer. Finally, we expound on the managerial implications of our findings on the cross-channel impact of digital and traditional advertising media.

### *LITERATURE REVIEW*

We distinguish between “own-channel” effects, where offline advertising affects offline sales or online advertising affects online sales, and “cross-channel effects,” where offline advertising affects online sales or online advertising affects offline sales. Research in this area can be categorized into these two groups, which we display in Figure 1. While there is an abundance of research in the former group (particularly in the offline channel), there is a small, but growing area of research investigating cross-channel advertising effects to which this paper contributes. See Table 1 for a selected list of research.

**<Insert Figure 1 and Table 1 About Here>**

#### *Own-Channel Advertising Effects*

The abundance of research investigating the impact of offline advertising on offline sales has generated two major meta-analyses. The first, Assmus, Farley and Lehmann (1984), came well before the common use of the web channel, and analyzes the results of 128 studies. They find an average short-term elasticity of 0.22 and an average carryover effect of 0.46. Recently, Sethuraman, Tellis, and Briesch (2011) cover the 25 years following Assmus et al. (1984). Using a larger sample of 751 estimates they focus on advertising from print, TV and “aggregate” media to find an average short-term elasticity of 0.12. They find that elasticities are lower in mature

markets and conclude that the decrease in the effectiveness of advertising over the past 25 years is due to “increased competition, ad clutter, the advent of the Internet as an alternate information source.” (p. 460).

The relatively recent advent of the online channel means that research about the effect of online advertising on online sales is sparser. Manchanda et al. (2006) determine the likelihood of consumer repurchase due to online display exposure and find an elasticity of 0.02. Trusov et al. (2009) measure the impact of event marketing on signups on a major social networking site. They estimate a short-run elasticity of only 0.002 for event marketing on signups; online word-of-mouth had an effect of 0.14. While in the context of a social network a new user may be considered a “sale,” it is not exactly the same as revenue which comes from selling a particular product. Other research in the web space has focused on dependent variables such as consideration sets (e.g., Naik and Peters 2009) or used independent variables other than advertising, such as online communications (Sonnier et al. 2011).

#### *Cross-Channel Advertising Effects*

Some researchers have begun to measure the cross-channel effects of online activities on offline sales (Pauwels et al. 2011; Van Nierop et al. 2011) as well as the impact of offline activities on online sales and consideration (Trusov et al. 2009; Naik and Peters 2009). In addition, some research measures both online and offline cross-channel effects. For example, Ansari et al. (2008) measure the impact of catalog and email contact on purchase incidence, choice and quantity. However, this model uses customer contact methods that are different from advertising.

A report by Abraham (2008) utilizes 18 comScore studies to group customers into four categories based on seeing online display ads, search ads, both or none. Regardless of dependent

variable (web visits, online and offline sales revenue), he finds that groups that saw no type of online ad provided the least revenue and visits, followed by customers who only saw display ads, a large jump to customers who only saw search ads, and the best performance from customers that saw both search and display ads. Interestingly, the biggest impact was on offline buyers, implying that the cross-channel effects from advertising may be more important than the own internet channel effects. However, these results do not provide a method on how to analyze the current resource allocation or decompose impact into customer count versus spend.

Finally, Wiesel et al. (2011) use vector autoregression to estimate how customers move through the “purchase funnel”. In this project they work with a family-run business-to-business company in the Netherlands to help the company optimize the allocation of fax, flyer and search advertising expenditures. As part of the study they find a very high sales elasticity for Adwords (4.35), and much lower sales elasticity of 0.05 and 0.04 for Flyers and Faxes, respectively. They find cross-effects such as 73% of the profit impact of Adwords is due to offline sales and 20% of the profit impact from direct mail flyers is due to online sales. Their work is an important step in measuring cross-effects of online and offline media, particularly relevant to the B2B arena. We build on this work by decomposing the sales impact into customer count and customer spend effects, by including competitive effects, by considering heterogeneity across markets, and by focusing on mass advertising communications in a B2C context.

In summary, while own-channel advertising effects have been studied extensively, researchers are beginning to make headway in considering cross-channel advertising effects. However, no paper has considered the sales decomposition we propose or addressed endogeneity, competitive advertising, heterogeneous markets, and dynamics that are endemic to empirically studying the impact of advertising. In addition, no study has examined, and

compared, three major types of advertising – traditional (offline), online display, and online search. Thus, we are able to provide deeper insights into how offline and online advertising generates sales both within and across channels. Further, our Dynamic Linear Model is uniquely suited to address the modeling challenges faced when studying advertising effectiveness.

### *CONCEPTUAL FRAMEWORK*

Research has conventionally focused on the impact of traditional advertising media within the offline channel. Today, the establishment of the internet channel allows for the analysis of marketing effectiveness from the online to the offline world and vice versa, creating a situation where advertising in one channel may have cross-channel impacts. As discussed above, research in this area is rather sparse, and we hope to contribute by developing a model to assess advertising effectiveness across channels while also decomposing the revenue to understand better *how* advertising impacts the number of customers, as well as customer spend. Building on the literature on revenue decomposition (Lam et al. 2001; Pauwels et al. 2002; Van Heerde and Bijmolt 2005) we separate total revenue into 4 components: 1) offline customer count, 2) offline customer spend, 3) online customer count and 4) online customer spend (see Figure 2). Estimating the impact of different media on this decomposition tells us which advertising media generate channel traffic and which generate spend per customer, allowing managers to develop advertising strategies that tie more closely to objectives and coordinate more effectively. E.g., Medium X may be needed to generate traffic on Channel A while Medium Y may be needed to generate spend on Channel B.

**<Insert Figure 2 About here>**

Figure 2 implies the following equation for total sales revenues:

$$(1) \quad \textit{Revenues} = \textit{OfflineRevenue} + \textit{OnlineRevenue}$$

$$= \text{OfflineCustomerCount} \times \text{OfflineCustomerSpend} \\ + \text{OnlineCustomerCount} \times \text{OnlineCustomerSpend}$$

The elasticity of revenues with respect to a particular advertising vehicle can be obtained by computing  $\left(\frac{\partial \text{Revenues}}{\partial \text{Advertising}}\right) \times \left(\frac{\text{Advertising}}{\text{Revenues}}\right)$ , which can be shown to equal (see Appendix 1):

$$(2) \quad \eta_{\text{Advertising} \rightarrow \text{Revenues}} = \overline{\% \text{Revenues}}_{\text{online}} (\eta_{\text{Advertising} \rightarrow \text{online, count}} + \eta_{\text{Advertising} \rightarrow \text{online, spend}}) + \\ \overline{\% \text{Revenues}}_{\text{offline}} (\eta_{\text{Advertising} \rightarrow \text{offline, count}} + \eta_{\text{Advertising} \rightarrow \text{offline, spend}})$$

where  $\overline{\% \text{Revenues}}_{\text{online}}$  is the mean percentage of the firm's revenues that are from online, and  $\overline{\% \text{Revenues}}_{\text{offline}}$  is the mean percentage of the firm's revenues that are from offline. Equation (2) shows that a particular advertising medium will have a strong impact on overall revenues to the extent that it increases customer counts and customer spend in channels that are a significant component of overall revenues.

Figure 3 combines Figures 1 and 2 to show the conceptual framework that drives our research. We examine the impact of three types of advertising (offline, online search, and online display) on the components of equation (1).

**<Insert Figure 3 About Here>**

### *HYPOTHESES*

We draw on three theories from the literature on multichannel customer management to develop hypotheses:

*Customer decision-making in a multichannel environment:* Based on information-processing theories of consumer decision-making, Neslin et al. (2006) developed a framework for how customers make purchase decisions in a multichannel environment. We adopt a simplified version of their framework, shown in Figure 4.

**<Insert Figure 4 about here>-**

Customers proceed from need recognition to search and finally purchase, a standard progression in information-processing theories (e.g., see Engel et al. 1995, pp. 146-152). However, in a multichannel environment, the search and purchase phases can take place in a variety of channels. For example, the customer may develop a need for a new television, search on the Internet to gather information about the various features and brands, and then make the purchase in a consumer electronics store. Advertising can influence all three phases of the process. It can stimulate need recognition, determine in which channel the customer searches, and influence in which channel the customer chooses to make the purchase.

*Research shopping:* Consumer research shopping is defined as searching for information in Channel A but purchasing in Channel B (Blattberg, Kim, and Neslin 2008, p. 652; Verhoef, Neslin, and Vroomen 2007, p. 129). Initial research by DoubleClick (2004) found that internet search followed by retail store purchase was the most common form of research shopping. Verhoef et al. (2007) replicated these findings and identified three factors that influence research shopping: (1) Attribute advantage: Customers have more favorable attribute perceptions for Channel A as a search channel than they do for Channel B; however, Channel B excels on purchase channel attributes. (2) Lack of channel lock-in: Channel A may not “lock in” the customer from search to purchase, i.e., the customer searches on the channel but easily moves to Channel B to make the purchase. Internet cart abandonment is a common example. (3) Channel synergy: The customer who searches on Channel A may have a more favorable purchase experience on Channel B. E.g., the customer learns about the features of a TV on the Internet so that he or she is able to ask more intelligent questions of sales personnel in the store.

*Channel/marketing congruency*: Blattberg et al. (2008) conjecture there is a “channel/marketing congruency” effect, “whereby marketing and channels link well if they use similar technologies.” For example, Ansari et al. (2008, p. 647) observed that emails are more likely to induce the customer to choose the internet rather than the telephone for purchase. As email and web stores use consistent technology (i.e., the internet), emails are more likely to route customers to an online store. Blattberg et al. (2008) did not develop this conjecture further, but their idea is consistent with a shopping cost argument. The consumer incurs lower shopping costs by staying in the same channel – i.e., responding to a display ad by searching and then buying on the Internet is a lower cost shopping strategy than research shopping. The congruency conjecture does not rule out research shopping, but it says that the similarity in technology between advertising and channels is an important factor. This means that web-based advertising will encourage customers to embark and stay on a web-based shopping experience.

In formulating our hypotheses, we also distinguish between “cross-effects” and “own-effects.” A cross-effect is when offline advertising affects online sales or online advertising affects offline sales. An own-effect is when offline advertising affects offline sales or online advertising affects online sales. In addition, our work studies two aspects of consumer behavior – customer counts, and customer spend. The above theories apply more directly to customer counts than to customer spend. That is, advertising may stimulate need recognition (“Maybe I need to buy a new winter coat?”), but the actual dollars spent on the coat is determined by the consumer’s budget, taste, and possibly promotions that offer economic incentives to spend more. This is consistent with research that finds customer count is more malleable with respect to advertising than customer spend. For example, Ansari et al. (2008) found that emails and

catalogs influenced purchase incidence and channel choice, but had no significant impact on order size. We therefore state:

H1: For display, search, and traditional advertising, cross and own-effects will be larger for customer count than for customer spend.

Considering that effects will be larger for customer count than for spend, our remaining hypotheses focus on customer count.

The most basic implication of the above theories is they provide a reason why cross-channel advertising effects exist. From the multichannel decision process perspective, advertising can explicitly encourage the customer to shop at a particular channel, for example a print ad may include an internet address and extol the virtues of internet shopping. From a research shopping perspective, advertising can initiate customer search on Channel A, but then the customer may find it convenient to search on Channel B. This would be the online→offline research shopping example described earlier. Both of these examples are positive cross-effects – more customers will purchase online in the first case; offline the second case. Note, however, that channel congruency can create negative cross-effects. For example, the customer might normally purchase offline, but an online ad encourages the customer to complete the entire process online. The customer has been diverted from offline to online purchasing, creating a negative cross-effect on offline purchasing. In any case, we have our fundamental hypothesis:

H2: Cross-channel advertising effects exist, wherein online advertising influences offline purchases, and offline advertising influences online purchases.

Research shopping encourages cross-effects, whereby channel congruency encourages own-effects. The question is which effect is stronger. Wiesel et al. (2011) show a strong online→offline effect, in contrast to Ansari et al.'s (2008) congruency results. While we acknowledge that this argument is debatable, for our empirical test we follow Ansari et al. (2008) and conjecture that the shopping costs that drive congruency will outweigh research shopping, and that overall, own-effects should be larger than cross-effects. This hypothesis is also suggested by the dominance of own-effects over cross-effects in pricing research. In that case, a decrease in price for the focal brand allows the focal brand loyalists to buy that brand. Likewise we hypothesize that congruency lowers switching costs, allowing focal channel loyalists to purchase on that channel. Hence:

H3: For all media, own-effects will be larger than cross-effects.

The above theories also have implications on the relative magnitudes of cross-effects. First we consider the relationship of traditional vs. display advertising. The key finding from the research shopping literature is that online→offline is more prevalent than offline→online. This is because online is a superior search channel while offline is a superior purchase channel; websites are less able to lock in the customer than a retail store and online search makes offline purchases more efficient. This yields:

H4: The cross-effect from display advertising to offline purchasing is larger than the cross-effect from traditional advertising to online purchasing.

Search advertising is different in kind than offline and display advertising because it targets consumers in the search stage of the decision process. A customer who sees a search ad already has need recognition and is likely engaged in “goal-directed” as compared to “experiential” shopping (Novak and Hoffman 2003). With goal-directed shopping, the customer has a particular objective in mind, whereas experiential shopping is more exploratory. There is evidence that the internet is especially amenable to goal-directed shopping (Wolfenbarger and Gilly 2001; To et al. 2007; Bridges and Florsheim 2008). The goal-directed shopper is much closer to executing a purchase, and search advertising helps by routing her or him to a website where the purchase can be made. Of course it is still possible that the customer clicks through, gathers information, and then makes the purchase offline. However, the above arguments suggest that the own-effect of search advertising will be much stronger than the cross-effect, compared to less-directive display advertising. We therefore have:

H5: The differential effect on online versus offline purchasing will be larger for search advertising than for display advertising.

#### *MODEL DEVELOPMENT*

We start with the decomposition of revenues into customer count and spend. Repeating equation (1) with subscripts to differentiate between markets and time periods, we have:

$$(3) \quad REVENUES_{mt} = OfflineCustomerCount_{mt} \times OfflineCustomerSpend_{mt} + \\ OnlineCustomerCount_{mt} \times OnlineCustomerSpend_{mt}$$

where  $REVENUES_{mt}$  is revenues in market  $m$  ( $m=1, \dots, M$ ) in week  $t$  ( $t=1, \dots, T$ ),

$OfflineCustomerCount_{mt}$  is the number of offline store customers,  $OfflineCustomerSpend_{mt}$  is the

average expenditures by offline store customers,  $OnlineCustomerCount_{mt}$  is the number of online store customers, and  $OnlineCustomerSpend_{mt}$  is the average expenditures by online store customers. To obtain the elasticities required in equation (2), we need to model the effects of offline (traditional) and online (display & search) advertising on the four terms on the right-hand side of equation (3). These four terms are the dependent variables in our model.

*Challenges and Solutions.* In developing this model, we face a number of challenges. Table 2 lists these challenges and the solutions we employ.

**<Insert Table 2 about here>**

A first challenge is that advertising is likely to be endogenous. That is, a manager may plan weekly advertising expenditures based on demand shocks known to her or him but unknown to the researcher. This creates a correlation between the regressor (advertising) and the error term, which leads to a biased advertising coefficient if not addressed (e.g., Sethuraman et al. 2011). One way we address this issue is to capitalize on exogenous variation in advertising in the data. In our data set, display advertising has been experimentally varied by the focal company in a number of test markets (more details in the next section). However, this accounts for only a small portion of display advertising, and traditional and search advertising have not been varied experimentally at all. Therefore, we use a Bayesian Instrumental Variable approach (Rossi et al., 2005, Section 7.1) to address the endogeneity of advertising.

A second challenge is that advertising effects are likely to carry over to the next period(s) (e.g. Leone 1995). The carry-over rate may be different for different media (traditional, display, search), because they serve different purposes in the purchase process. On top of that, carry-over may differ even for a given advertising medium across the four dependent variables. To cope with this challenge, we use Adstock, or goodwill, for the effect of each advertising medium on each dependent variable (Broadbent 1984). Adstock is the cumulative value of a brand's

advertising at a given point in time. Over time, advertising builds a stock of consumer goodwill which subsequently decays with the time since the previous exposure (e.g., Broadbent 1984; Danaher et al. 2008). We cast the model as a Dynamic Linear Model (West and Harrison 1999), because this allows us to readily model Adstock variables (separately for each medium) and estimate their carry-over effects on the four dependent variables.

A next challenge is that there may be heterogeneity in response parameters across markets. Depending on (unobserved) consumer and market characteristics, some markets may show a stronger response to advertising and other marketing activities than other markets. We cope with this challenge by adopting a Hierarchical Bayesian specification (e.g., Chib and Greenberg 1995; Montgomery 1997).

Next, we need to accommodate for error correlations between the four dependent variables. For instance, a higher number of store customers may mean a more crowded store environment. This may lead to less personal service per customer, which may reduce the amount spent per customer (Van Heerde and Bijmolt 2005). Also, adverse weather may reduce the number customers to physical stores yet increase the number of online customers. For these reasons, we allow for a full error covariance matrix for the number of online (offline) customers and their average expenditures, both within markets and across markets.

A final challenge is that the model needs to accommodate competitive effects. Advertising by a competitor may adversely affect the focal company if the ads entice potential customers to switch. On the other hand, if competitor advertising serves as a need recognition reminder, it could have a positive effect on focal company's performance (Schultz and Wittink 1976). Our model controls for competitor advertising efforts, both for traditional and online media. We also allow for the possible endogeneity of the competitor advertising variables.

*Model Specification.* We specify a system of equations for the four dependent variables: the log of the number of store and online customers and their respective log average expenditures. We use a log-log model because it allows us to interpret the coefficients (for log independent variables) as elasticities. The model equations are:

$$(4) \quad \ln \text{OfflineCustomerCount}_{mt} = \text{TraditionalAdStock}_{1,mt} + \text{OnlineDisplayAdStock}_{1,mt} + \text{OnlineSearchAdStock}_{1,mt} + X'_{mt}\beta_{1,m} + v_{1,mt}$$

$$(5) \quad \ln \text{OfflineCustomerSpend}_{mt} = \text{TraditionalAdStock}_{2,mt} + \text{OnlineDisplayAdStock}_{2,mt} + \text{OnlineSearchAdStock}_{2,mt} + X'_{mt}\beta_{2,m} + v_{2,mt}$$

$$(6) \quad \ln \text{OnlineCustomerCount}_{mt} = \text{TraditionalAdStock}_{3,mt} + \text{OnlineDisplayAdStock}_{3,mt} + \text{OnlineSearchAdStock}_{3,mt} + X'_{mt}\beta_{3,m} + v_{3,mt}$$

$$(7) \quad \ln \text{OnlineCustomerSpend}_{mt} = \text{TraditionalAdStock}_{4,mt} + \text{OnlineDisplayAdStock}_{4,mt} + \text{OnlineSearchAdStock}_{4,mt} + X'_{mt}\beta_{4,m} + v_{4,mt}$$

where  $\text{TraditionalAdStock}_{i,mt}$ ,  $\text{OnlineDisplayAdStock}_{i,mt}$ , and  $\text{OnlineSearchAdStock}_{i,mt}$  are the effects of the, respectively, traditional, online display, and online search AdStock variables in market  $m$  in week  $t$  on dependent variable  $i$  ( $i=1$  for the log of the number of store customers,  $i=2$  for the log average expenditures for store customers,  $i=3$  for the log of the number of online customers, and  $i=4$  for the log average expenditures for online customers).  $X'_{mt}$  (defined below) is a 1 by  $K$  row vector of control variables including an intercept and cross-advertising expenditures, and the  $v_{i,mt}$  are error terms.  $\beta_{i,m}$  are the effects of control variables.

*AdStock.* Adstock is a latent construct defined as follows:

$$(8) \quad \text{TraditionalAdStock}_{i,mt} = \lambda_{i,1m} \text{TraditionalAdStock}_{i,mt-1} + \eta_{i,1m} \ln(\text{TraditionalAdvertising}_t + 1) + \omega_{i,1mt}$$

$$(9) \quad \text{OnlineDisplayAdStock}_{i,mt} = \lambda_{i,2m} \text{OnlineDisplayAdStock}_{i,mt-1} + \eta_{i,2m} \ln(\text{OnlineDisplayAdvertising}_t + 1) + \omega_{i,2mt}$$

$$(10) \quad \text{OnlineSearchAdStock}_{i,mt} = \lambda_{i,3m} \text{OnlineSearchAdStock}_{i,mt-1} + \eta_{i,3m} \ln(\text{OnlineSearchAdvertising}_t + 1) + \omega_{i,3mt}$$

where  $TraditionalAdvertising_t$  is the dollar expenditures on traditional advertising in week  $t$ ;  $OnlineDisplayAdvertising_t$  is the display advertising spend, and  $OnlineSearchAdvertising_t$  is the search advertising spend. We take the logs of 1 plus these amounts to avoid taking the log of zero in case of zero expenditures. The carry-over coefficient is  $\lambda_{i,jm}$  for medium  $j$ , market  $m$ , dependent variable  $i$ .  $\omega_{i,jmt}$  is the i.i.d. error term, which allows for stochastic behavior of AdStock:  $\omega_{i,jmt} \sim N(0, W_{i,jm})$ .  $\eta_{i,jm}$  is the short-run advertising elasticity for medium  $j$ , market  $m$ , dependent variable  $i$  (this can be seen upon substitution of (8) - (10) into (4) - (7)), and  $\eta_{i,jm}/(1 - \lambda_{i,jm})$  is the long-run advertising elasticity if  $0 \leq \lambda_{i,jm} < 1$ .

*Instrumental Variables.* To control for endogeneity, we use a Bayesian Instrumental Variable approach (Rossi et al. 2005, Section 7.1). We regress the own and competitor advertising variables in the model on instrumental variables (operationalized in the data section):

$$(11) \ln(TraditionalAdvertising_t + 1) = \gamma_{0,1t} + InstrumentalVars'_t \gamma_1 + v_{5,t}$$

$$(12) \ln(OnlineDisplayAdvertising_t + 1) = \gamma_{0,2t} + InstrumentalVars'_t \gamma_2 + v_{6,t}$$

$$(13) \ln(OnlineSearchAdvertising_t + 1) = \gamma_{0,3t} + InstrumentalVars'_t \gamma_3 + v_{7,t}$$

$$(14) \ln(CompetitorTraditionalAdvertising_t + 1) = \gamma_{0,4t} + InstrumentalVars'_t \gamma_4 + v_{8,t}$$

$$(15) \ln(CompetitorDisplayAdvertising_t + 1) = \gamma_{0,5t} + InstrumentalVars'_t \gamma_5 + v_{9,t}$$

We correlate the errors of the endogenous regressors in equations (11)-(15) with the errors of the main model (4) - (7). In particular, we specify:

$$(16) \quad v_t = (v_{1,1t}, v_{2,1t}, \dots, v_{4,Mt}, v_{5,t}, v_{6,t}, v_{7,t}, v_{8,t}, v_{9,t})' \sim N(0, V),$$

where  $V$  is a full covariance matrix. Finally, we specify time-varying intercepts  $\gamma_{0,kt}$  ( $k=1, \dots, 5$ ) to allow for flexible yet parsimonious advertising patterns not captured by the rest of the covariates:

$$\gamma_{0,kt} = \gamma_{0,kt-1} + \omega_{0k,t}, \text{ where the i.i.d. errors } \omega_{0k,t} \sim N(0, W_{0k}).$$

*Dynamic Linear Model.* To estimate the model, we rewrite equations (4)-(16) as a Dynamic Linear Model (DLM) (e.g., West and Harrison 1999, p. 284; Ataman et al. 2010; Van Heerde et al. 2004). We need to organize the variables and parameters such that we obtain the observation equation in this format:

$$(17) \quad y_t = F_t' \theta_t + X_t' \beta + v_t$$

and the state equation in this format:

$$(18) \quad \theta_t = G \theta_{t-1} + Z_t' \eta + \omega_t.$$

The observation equation is organized by stacking equations (4) - (16) as follows:

$$\underbrace{\begin{pmatrix} \ln \text{OfflineCustomerCount}_{1t} \\ \ln \text{OfflineCustomerSpend}_{1t} \\ \ln \text{OnlineCustomerCount}_{1t} \\ \ln \text{OnlineCustomerSpend}_{1t} \\ \vdots \\ \ln \text{OfflineCustomerCount}_{Mt} \\ \ln \text{OfflineCustomerSpend}_{Mt} \\ \ln \text{OnlineCustomerCount}_{Mt} \\ \ln \text{OnlineCustomerSpend}_{Mt} \\ \ln(\text{TraditionalAdvertising}_t + 1) \\ \ln(\text{OnlineDisplayAdvertising}_t + 1) \\ \ln(\text{OnlineSearchAdvertising}_t + 1) \\ \ln(\text{CompetitorTraditionalAdvertising}_t + 1) \\ \ln(\text{CompetitorDisplayAdvertising}_t + 1) \end{pmatrix}}_{y_t} = \underbrace{\begin{pmatrix} I_{4M} \otimes (1 \ 1 \ 1) & 0 \\ 0 & I_5 \end{pmatrix}}_{F_t'} \underbrace{\begin{pmatrix} \text{TraditionalAdStock}_{1,1t} \\ \text{OnlineDisplayAdStock}_{1,1t} \\ \text{OnlineSearchAdStock}_{1,1t} \\ \vdots \\ \vdots \\ \text{TraditionalAdStock}_{4,Mt} \\ \text{OnlineDisplayAdStock}_{4,Mt} \\ \text{OnlineSearchAdStock}_{4,Mt} \\ \gamma_{0,1t} \\ \gamma_{0,2t} \\ \gamma_{0,3t} \\ \gamma_{0,4t} \\ \gamma_{0,5t} \end{pmatrix}}_{\theta_t} + \underbrace{\begin{pmatrix} I_4 \otimes X'_{1t} & \cdots & I_4 \otimes X'_{Mt} & 0 \\ 0 & & I_5 \otimes (\text{InstrumentalVars}'_t) & \begin{pmatrix} \beta_{1,1} \\ \beta_{2,1} \\ \vdots \\ \beta_{4,M} \\ \gamma_1 \\ \vdots \\ \gamma_5 \end{pmatrix} \end{pmatrix}}_{X_t'} \underbrace{\begin{pmatrix} \beta_{1,1} \\ \beta_{2,1} \\ \vdots \\ \beta_{4,M} \\ \gamma_1 \\ \vdots \\ \gamma_5 \end{pmatrix}}_{\beta} + v_t,$$

where  $I_L$  is an identity matrix of size  $L$  and  $\otimes$  is the Kronecker product.

The state equation is operationalized as:

$$\begin{aligned}
& \underbrace{\begin{pmatrix} \text{TraditionalAdStock}_{1,1t} \\ \text{OnlineDisplayAdStock}_{1,1t} \\ \text{OnlineSearchAdStock}_{1,1t} \\ \vdots \\ \vdots \\ \text{TraditionalAdStock}_{4,Mt} \\ \text{OnlineDisplayAdStock}_{4,Mt} \\ \text{OnlineSearchAdStock}_{4,Mt} \\ \gamma_{0,1t} \\ \gamma_{0,2t} \\ \gamma_{0,3t} \\ \gamma_{0,4t} \\ \gamma_{0,5t} \end{pmatrix}}_{\theta_t} = \underbrace{\begin{pmatrix} \lambda_{1,11} & & & & \\ & \ddots & & & \\ & & \lambda_{4,3M} & & \\ & & & \ddots & \\ & & & & I_5 \end{pmatrix}}_G \underbrace{\begin{pmatrix} \text{TraditionalAdStock}_{1,1t-1} \\ \text{OnlineDisplayAdStock}_{1,1t-1} \\ \text{OnlineSearchAdStock}_{1,1t-1} \\ \vdots \\ \vdots \\ \text{TraditionalAdStock}_{4,Mt-1} \\ \text{OnlineDisplayAdStock}_{4,Mt-1} \\ \text{OnlineSearchAdStock}_{4,Mt-1} \\ \gamma_{0,1t-1} \\ \gamma_{0,2t-1} \\ \gamma_{0,3t-1} \\ \gamma_{0,4t-1} \\ \gamma_{0,5t-1} \end{pmatrix}}_{\theta_{t-1}} + \\
& + \underbrace{\left( \begin{bmatrix} I_{4M} \otimes \begin{pmatrix} \ln(\text{TraditionalAdvertising}_t + 1) \\ \ln(\text{OnlineDisplayAdvertising}_t + 1) \\ 0 \\ \ln(\text{OnlineSearchAdvertising}_t + 1) \end{pmatrix} \end{bmatrix} \right)}_{Z_t'} \underbrace{\begin{pmatrix} \eta_{1,11} \\ \vdots \\ \eta_{4,31} \\ \vdots \\ \eta_{1,1M} \\ \vdots \\ \eta_{4,3M} \end{pmatrix}}_{\eta} + \\
& + \underbrace{\begin{pmatrix} \omega_{1,11t} \\ \omega_{2,11t} \\ \vdots \\ \omega_{0,5t} \end{pmatrix}}_{\omega_t}, \text{ where } \omega_t \sim N(0, W), \text{ with } W \text{ a diagonal covariance matrix.}
\end{aligned}$$

*Parameter shrinkage.* The market-specific parameters are assumed to be draws from national hypermeans:  $\beta_m = (\beta_{1,m}, \beta_{2,m}, \beta_{3,m}, \beta_{4,m})' \sim N(\bar{\beta}, V_\beta)$ ,  $\lambda_m = (\lambda_{1,1m}, \dots, \lambda_{4,3m}) \sim N(\bar{\lambda}, V_\lambda)$ ,  $\eta_m = (\eta_{1,1m}, \dots, \eta_{4,3m})' \sim N(\bar{\eta}, V_\eta)$ , where  $V_\beta$  ( $4K$  by  $4K$ ),  $V_\lambda$  ( $12$  by  $12$ ), and  $V_\eta$  ( $12$  by  $12$ ) are full covariance matrices (Chib and Greenberg 1995).

*Estimation.* The DLM is estimated using forward-filtering, backward sampling (Carter & Kohn 1994, Frühwirth-Schnatter 1994). Time-invariant regression parameters and the variance matrices are estimated with Hierarchical Bayes / Gibbs sampling with uninformative priors. We conducted simulation tests that confirm we can retrieve the model parameters (see Appendix 2).

## DATA DESCRIPTION

We use data from a major American clothing retailer that wishes to remain anonymous. The store offers a large assortment of high-end clothing and accessories. They participate in 25 US markets. We have two years of weekly market-level sales data covering September 2008-August 2010. This provides 2575 aggregate in-store sales observations and 2575 aggregate online sales observations (25 markets  $\times$  103 weeks; we lose the first week to operationalize lagged effects). For each market and week our data include the number of unique customers and the average spend per customer. Operationalization and source of all variables is in Table 3.

**<Insert Tables 3 and 4 about here>**

### *Advertising Spend*

The retailer undertakes both online and offline advertising. Their traditional (offline) advertising is a mix of radio, print media, television and billboard campaigns that averages 33% of total advertising expenditures. These campaigns are relatively short and unique, making generalizable analysis by any one type difficult. We therefore aggregate these expenditures across the various traditional media. Full descriptive statistics are in Table 4.

The retailer spends 38% of its advertising budget on display advertising (banners) and 29% on search advertising. Most of the display advertising is general in nature, and either advertises the store name or a specific event. In addition, during a twelve-week period, five of the markets received targeted display advertisements specifically mentioning the regional store. As spending for these advertisements is included with our national number we add a dummy variable (*DummyTestBannerAdvertising<sub>mt</sub>*), signifying when the particular region received these targeted ads.

### *Control Variables*

The retailer implements a large number of promotions. These range from in-store discount sales to promotional events to special discounts for targeted customers. We operationalize  $Promotion_{mt}$  as the sum of all promotion-days in the given week. For example, if party dresses are discounted 7 days, men's shoes are discounted 4 days and there is a one-day promotional event for a specific designer, then this week would have 12 promotion-days. The average week had 9.6 promotion days, and there was variation in promotions across markets.

Other variables include economic environment, seasonality and trend. Specifically, we use unemployment at the market level as a proxy for the macro-economic conditions ( $UnemploymentRate_{mt}$ ). We also include a Christmas season dummy variable ( $ChristmasDummy_t$ ) to account for any natural increases in purchase behavior at this time. To account for any systematic changes in customer count or spend that occur within our time period we also use a trend variable ( $Trend_t$ ). We also account for the advertising expenditures from the firm's primary competitor. This information, from TNS, includes the amount spent on traditional advertising ( $CompetitorTraditionalAdvertising_t$ ) and online display advertising ( $CompetitorOnlineDisplayAdvertising_t$ ).

### *Instrumental Variables*

We instrument for the five endogenous own and competitor advertising variables in our model (equations 11-15). The vector of instrumental variables,  $InstrumentalVars'_t$ , includes traditional and online display advertising expenditures of five retailers that are non-direct competitors of the focal firm (the focal firm is in a much higher price bracket). We picked these ten (five non-direct competitors times two types of advertising – traditional and online) instrumental variables because they are unlikely to influence sales of our focal retailer, hence

will not be correlated with the error term of the main model (4) - (5). However, to the extent that clothing retailers advertise at similar times within the year as well as over time, they should be correlated with our endogenous variables. These ten instrumental variables are not used elsewhere in the system, which is sufficient for five endogenous regressors. The instruments appear to be sufficiently strong (the incremental R-squares range from 0.25 to 0.41).

### *MODEL RESULTS*

The estimation approach was outlined earlier and is also described in Appendix 2. We run the Gibbs sampler for 20,000 draws and retain each 10th draw of the last 10,000 draws. Visual inspection confirms convergence of the Gibbs chain. We next used Raftery and Lewis's (1996) test for MCMC convergence (implemented in the "coda" package in R). This test confirms that the burn-in and inference samples are sufficiently large. The procedure resulted in 1,000 draws used for inference of the posterior distribution.<sup>1</sup>

#### *Control Variable Impact*

We first discuss the impact of the control variables to assess the validity of the model (Table 5). As expected, an increase in unemployment decreases the number of customers purchasing both in-store (-0.228) and online (-0.287), but does not impact the amount of spend per customer. Similarly, the Christmas season increases the number of customers (offline = 0.257, online = 0.322), but does not affect order size. Promotions have a positive effect on the number of store customers (0.013) and in-store customer spend (0.009). The effect of promotions is weaker for the number of online customers (0.010), which is plausible as the promotions are more comprehensive and salient in-store. Finally, *traditional* advertising by the competitor

---

<sup>1</sup> We also estimated a non-dynamic version of the model with two-stage least squares. The effects are qualitatively similar to those from the Dynamic Linear Model, and the results are available on request.

significantly decreases both the number of customers (offline = -0.194, online = -0.216) and spend across channels (offline = -0.147, online = -0.230). However, *online* advertising by the competitor will in fact increase both online customer spend and the online customer count. This may reflect online advertising's ability to stimulate need recognition and initiate information search (see also Wiesel et al. 2011).

**<Insert Table 5 about here>**

#### *Elasticities and Carry-Over Coefficients*

Table 6 shows the short-term advertising elasticities ( $\eta$ ). Traditional advertising has positive significant elasticities for the number of store customer count (0.009) as well for online customer spend (0.03), but a negative cross-elasticity for online customer count (-0.021). Online display advertising has a positive impact on online customer count (0.045), on online customer spend (0.014), and a positive cross-effect on offline customer count (0.052). Online search advertising has a significant impact on all dependent variables: own-effects on online customer count (0.153) and online customer spend (-0.044), as well as cross-effects on store customer count (0.107) and in-store customer spend (0.077).

**<Insert Tables 6-8 about here>**

The carryover coefficients ( $\lambda$ ) for advertising are in Table 7. Carryover effects are quite consistent across advertising media and dependent variables (0.57-0.74). The highest carryover is for the online customer count (0.65-74), followed by store customer count (0.61-65). Carryover is comparable for online and offline customer spend (0.59-0.63). Also, in all cases online display advertising has the lowest carryover. An alternative way to interpret the carry-over effect is to consider how long it will take for 90% of the effect to dissipate (also listed in Table 7). Using the duration interval we see that 90% of the advertising effects dissipate in 4.4-7.8 weeks.

Combining the carryover coefficients and the short term coefficients we can then calculate the long-term advertising elasticities (Table 8). These long-term elasticities are around 2.5 - 3 times larger than the short-term elasticities, and their qualitative insights are similar.

**<Insert Table 8 about here>**

To estimate the elasticity of each advertising medium on individual channel sales, we sum the elasticities for customer count and customer spend in that channel (Equation (2)). This is shown in Table 9 (short-term) and Table 10 (long-term). For example, for traditional advertising, the total short-term store sales elasticity is  $0.009 + 0.002 = 0.011$ . Likewise its total short-term online sales elasticity is  $-0.021 + 0.030 = 0.009$ . When we weight these two impacts by the mean fraction of sales through each channel (Equation (2)), the total impact of traditional advertising is 0.011. Online display advertising has a short-term store elasticity of 0.053 and online elasticity of 0.058, for a total short-term impact of 0.054. Search advertising has a short-term store elasticity of 0.184 and online of 0.108, for a total short-term impact of 0.172. Note that the total elasticities are close in magnitude to the store elasticities. This is because store sales are roughly 85% of total sales, so the elasticity of a medium through the store channel largely determines its total elasticity. Overall, search has a higher total elasticity than display, which in turn is more effective than traditional advertising. The strong results for search advertising are consistent with Wiesel et al. (2011). Table 10 shows that traditional advertising has a total long-term elasticity of 0.025, display advertising's elasticity is 0.141 and search's elasticity is 0.491.

**<Insert Tables 9 and 10 about here>**

#### *Heterogeneity in Parameters*

A benefit of this model is that we are able to account for heterogeneity of individual elasticities across markets. We create histograms of the elasticity of each advertising medium's

effect to demonstrate the differences in the median elasticities. In Figure 5 we demonstrate these histograms for the count dependent variables, the primary item of interest. We see that inter-market differences are not drastically different, but have a nontrivial range. For example, the range of display's elasticity on online customer count more than doubles from 0.03 to 0.07.

**<Insert Figure 5 about here>**

### *Testing Hypotheses*

To test our hypotheses we compare the estimated elasticities based on their posterior draws from the Bayesian estimation routine. For each hypothesis we calculate the pertinent difference per draw, and check whether the 95% posterior density for the difference includes or excludes zero. Our hypothesis tests are shown in Table 11.

**<Insert Table 11 about here>**

For H1 we compare the elasticity effects for count and order size for each of the six combinations of advertising medium with channel purchase. All four online display and search advertising elasticities demonstrate higher effects for customer counts than for customer spend. However, both traditional advertising elasticities (offline and online) are not significantly different for count versus spend. This provides partial support for our hypothesis that effects will be larger for customer counts than for customer spend.

For H2, we test if the cross effects on customer count (traditional→online, display→offline and search→offline) are non-zero. We find that all three of these values are non-zero, confirming H2 that advertising in one channel has impacts on the other. Interestingly, while both online→offline elasticities are positive, the impact of traditional advertising on online customer count is negative. This may be a case where traditional advertising is driving transactions that might have occurred online to the stores, cannibalizing online customer counts.

H3 tests if own-effects will be larger than the cross effects. We find that this is supported for search and traditional advertising, but not for display advertising. We confirm H4, that the cross-effect from display advertising to offline purchasing is larger in magnitude than the cross effect from offline advertising to online purchasing. Finally, we also confirm H5, that the differential effect on online versus offline purchasing is larger for search advertising than display advertising.

#### *Return on Investment*

ROI is the return from investing a dollar of advertising on the profit of each of the channels, calculated as:  $ROI = \eta_{\text{Advertising}} \times (\text{Average Profit in Channel} / \text{Average Advertising Spend}) - 1$ . We use a 35% profit margin because the company's annual report states that the cost of goods sold are around 65% of revenues. Table 12 shows that while traditional advertising has a negative short-term ROI for offline sales (-\$0.32), it does return a profit in the long term (\$0.92). This is consistent with traditional methods being only moderately effective for enhancing sales (e.g., Ataman et al. 2010). Second, both display and search ads are very effective offline, particularly in the long term (\$6.38 for display, \$35.72 for search), yet less so on online spending (\$0.78 for display, \$4.45 for search). This difference across sales channel is mainly due to the fact that baseline offline sales are much higher than baseline online sales (Table 4). Thus, a given percentage increase in offline sales contributes much more in magnitude.

The ROIs for display and especially search advertising are quite large, especially in the long term. However, Wiesel et al. (2011, Table 4) provide a most direct comparison and obtain similar results. They find an ROI of 56.72 Euros for the long-term, multichannel return from search advertising. An industry study reported search ROIs of £17 - £27 for a British travel agency (Murphy 2008). We also found industry data supporting an average search campaign ROI

of 6.03 Yuan for a Chinese cosmetics e-tailer (Yang 2011). A final point is that many companies evaluate online search ROI by assessing click-throughs and conversion. This is a short-term calculation that may be appropriate under certain conditions, but does not include cross-channel effects nor the long-term effect, factors we find to be quite important in our application.

### *Benchmark Model Comparison*

We compare our focal model to a benchmark model with only own-channel effects. This benchmark only includes the effects of offline advertising on offline customer counts and spends and of online (search and display) advertising on online customer counts and spends. Other than restricting the cross-channel advertising effects to zero, the benchmark is identical to the focal model. For both models we evaluate holdout sample performance by calculating the log likelihood of the one-step-ahead forecasts (West and Harrison 1999, p. 393-394). The difference in the mean likelihood across draws, which is the log Bayes factor, is 50.2 in favor of the focal model. Since this number is larger than 2, we conclude there is strong evidence (West and Harrison 1999, p. 393-394) for the existence of cross-channel effects of advertising.

## DISCUSSION

Our results demonstrate that the impact of advertising a) is not just within a single purchase channel, but across multiple channels, b) has differential effects on customer counts and spend across channels and c) for online search and online display have larger returns compared to traditional advertising.

Measuring the impact of advertising on only a single channel does not fully account for its total impact, underestimating the overall effect. For example, many firms that use Google AdWords multiply the probability that a customer who clicks their ad will purchase or “convert”

(e.g. 2%) times the amount they earn from that purchase (e.g. \$50) to determine a maximum amount to pay for that click (e.g.  $2\% \times \$50 = \$1.00$ ). However, by only assuming instantaneous effects within the internet channel, firms are omitting two factors – a) the long-term effect of the search ad impressions on online sales and b) the effect of the search ad impressions on offline sales. Our results demonstrate that these cross-effects exist, which supports H2. For example, the immediate online ROI from search advertising for our firm is only \$0.62. However, the long-term online impact from this advertising is \$4.45. This long-term effect may be due to click-throughs that translate into sales later than in the current period, or to increases in awareness or attitudes due to impressions (Rutz et al. 2011). Even more impressive is the ROI on offline sales of \$35.72. This is profit that cannot be accounted for with standard own-channel metrics. Similar cases can be made for online display advertising, which has an immediately *unprofitable* online ROI of -\$0.35, but a profitable long term offline ROI of \$6.38. Interestingly, for our firm, traditional advertising partially cannibalizes online sales, yielding a negative long-term ROI of -\$1.07. However, this is partially offset by positive long-term returns for offline sales (\$0.92).

In addition, these cross-effects differ across channels. For example, results from testing H3 find that traditional and online search advertising have larger own-effects than cross effects, while there is no difference between own and cross-effects for online display advertising. Further, as supported by H4, the cross-effect from online display advertising to offline purchasing is larger in magnitude than the cross-effect from offline advertising to online purchasing. This asymmetry suggests, for our retailer, that research shopping is relatively more powerful than channel congruency.

Prior research has shown that while advertising and other promotions can have an effect on customer counts, there is little to no effect on the level of customer spend (Ansari et al. 2008),

which we consider in H1. The rationale for this theory is that firms have an easier time collecting customers from other retailers than convincing a consumer to reallocate part of a fixed budget. However, our decomposition shows that in some cases advertising does have an impact on customer spend. For instance, while online display is successful at increasing customer counts in both channels, it also increases the spend of online customers. Interestingly, the transactions that search ads bring to offline stores are significantly larger in quantity (0.077), while transactions in the online channel are significantly lower (-0.044). That is, online search enhances order size in the offline channel but decreases order size in the online channel. By only considering the customer counts, we might conclude that online search has more of an effect on the online channel, but including both customer count and spend allows us to better understand the mechanisms and see that search advertising has a greater effectiveness on the offline channel.

We observe that for our firm, online search advertising has the greatest effectiveness, followed by online display and then traditional advertising. While the cause for these differences is not definite, we can make some conjectures. Despite online display and traditional advertising being of a similar nature, i.e. focused on impression-based awareness, it is possible that our retailer's display advertising is better targeted than the traditional advertising. In fact, further discussions with our firm's managers revealed that 40-50% of its display ads were "re-targeted," i.e., served to customers who had recently visited the firm's website. Also, the online environment enables customization of advertisements based on web content (Hauser et al. 2009) while billboards and magazines cannot be customized. In contrast to both traditional and display advertising, search advertising targets customers who are further along in the purchase decision process (Figure 4). That is, online search advertising can be used to target customers undertaking active product searches as opposed to merely stimulating need recognition. This is further

supported from H5, where we see that search advertising has a relatively larger impact on online purchasing where the consumer is close to making a decision.

## CONCLUSION

In this paper we develop a Dynamic Linear Model to estimate how traditional and digital advertising impact sales, customer counts and spend across online and offline media while accounting for (i) endogeneity of advertising, (ii) dynamic advertising effects, (iii) a multivariate dependent variable, (iv) heterogeneity across markets and (v) competitive advertising effects. Our model addresses one of the key research priorities identified by the Marketing Science Institute (2010) by assessing the relative effectiveness of digital advertising media compared with traditional media. For the specific retailer we studied, we find that advertising cross-effects are large, particularly from online advertising to offline sales, and that advertising has differential effects on customer counts and spend across channels. This sales decomposition allows the firm to align its advertising spend with its advertising objectives. For example, display advertising is an effective means to increase offline customer count, whereas traditional advertising improves online customer spend (Table 8).

Undoubtedly, many of our specific empirical findings are due to the particular industry we examined: clothing/accessories retailing. However, the model can easily be used to examine other industries and to measure the impact of other types of advertising or consumer contact types. This should encourage further research to replicate our results in other settings as well as measure the impact of new media such as mobile devices.

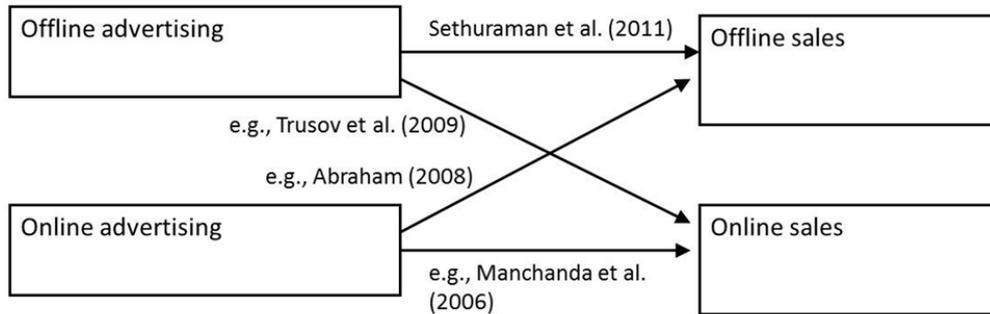
## REFERENCES

- Abraham, Magid (2008), "The Off-Line Impact of Online Ads," *Harvard Business Review*, 86 (April), 28.
- Ansari, Asim, Carl Mela, and Scott A. Neslin (2008), "Customer Channel Migration," *Journal of Marketing Research*, 45 (1), 60-76
- Assmus, Gert, John U. Farley and Donald R. Lehmann (1984), "How Advertising Affects Sales: Meta-Analysis of Econometric Results," *Journal of Marketing Research*, 21 (1), 65074.
- Ataman, Berk, Harald J. van Heerde, and Carl. F. Mela (2010), "The Long-term Effect of Marketing Strategy on Brand Sales," *Journal of Marketing Research*, 47 (October), 866-882.
- Ataman, Berk, Carl. F. Mela, and Harald J. van Heerde (2008), "Building Brands," *Marketing Science*, 27 (November–December), 1036–1054.
- Blattberg, Robert C., Byung-Do Kim, and Scott A. Neslin (2008), *Database Marketing: Analyzing and Managing Customers*, New York: Springer Science+Business Media, LLC.
- Bridges, Eileen, and Renée Florsheim (2008), "Hedonic and Utilitarian Shopping Goals: The Online Experience," *Journal of Business Research*, 61 (4), 309-314.
- Broadbent, Simon S. (1984), "Modeling with Adstock," *Journal of the Market Research Society*, 26 (4), 295-312.
- Carter, C. and R. Kohn (1994), "On Gibbs Sampling for State Space Models," *Biometrika*, 81 (3), 541–53.
- Chib, Siddhartha and Edward Greenberg (1995), "Hierarchical Analysis of SUR Models with Extensions to Correlated Serial Errors and Time-Varying Parameter Models," *Journal of Econometrics*, 68 (2), 339-60.
- Danaher, Peter J., André Bonfrer, and Sanjay Dhar (2008), "The Effect of Competitive Advertising Interference on Sales for Packaged Goods," *Journal of Marketing Research*, 45 (April), 211-25.
- DoubleClick (2004), "Multi-channel Shopping Study – Holiday 2003", New York: DoubleClick.
- Engel, James F., Roger D. Blackwell, and Paul W. Miniard (1995), *Consumer Behavior*, 8<sup>th</sup> Edition, Fort Worth, TX: The Dryden Press.
- Frühwirth-Schnatter, Sylvia (1994), "Data Augmentation and Dynamic Linear Models," *Journal of Time Series Analysis*, 15 (2), 183–202.
- Goldfarb, Avi, and Catherine Tucker (2011), "Online Display Advertising Targeting and Obtrusiveness," *Marketing Science*, 30 (3), 389-404.
- Hauser, John R., Glen L. Urban, Guilherme Liberali and Michael Braun (2009) "Website Morphing," *Marketing Science*, 28 (2), 202-223.
- IAB (2010) "IAB Internet Advertising Revenue Report," May 23, 2010, [http://www.iab.net/resources/adrevenue/pdf/IAB\\_PwC\\_2010\\_full\\_year.pdf](http://www.iab.net/resources/adrevenue/pdf/IAB_PwC_2010_full_year.pdf)
- Lam, Shun Y., Mark Vandenbosch, John Hulland, and Michael Pearce (2001), "Evaluating Promotions in Shopping Environments: Decomposing Sales Response into Attraction, Conversion, and Spending Effects," *Marketing Science*, 20 (2), 194-215.
- Leone, Robert P. (1995), "Generalizing What is Known about Temporal Aggregation and Advertising Carryover," *Marketing Science*, 14 (3, Part 2 of 2), G141-G150

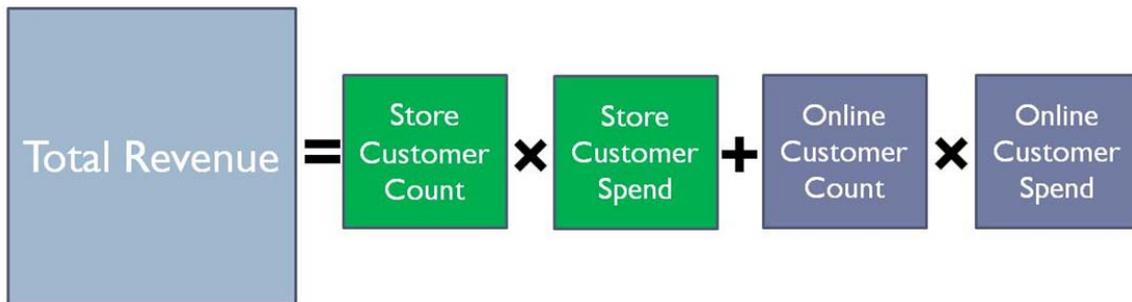
- Lodish, Leonard M., Magid Abraham, Stuart Kalmenson, Jeanne Livelsberger, Beth Lubetkin, Bruce Richardson and Mary Ellen Stevens (1995), "How T.V. Advertising Works: A Meta-Analysis of 389 Real World Split Cable T.V. Advertising Experiments," *Journal of Marketing Research*, 32 (May), 125-139.
- McKinsey (2009), "The Crisis: A New Era in Management," *McKinsey Quarterly*, 2009 (1).
- Macy's (2011) "Macy's, Inc. Same-Store Sales Up 6.7% in June," July 7, 2011, <http://phx.corporate-ir.net/phoenix.zhtml?c=84477&p=irol-newsArticle&ID=1582922>
- Manchanda, Puneet, Jean-Pierre Dubé, K.Y. Goh, and Pradeep Chintagunta (2006), "The Effect of Banner Advertising on Internet Purchasing," *Journal of Marketing Research*, 43 (1), 98-108.
- Marketing Science Institute (2010), "Research Priorities", <http://www.msi.org/research/index.cfm?id=43>.
- Montgomery, Alan L. (1997), "Creating Micro-Marketing Pricing Strategies Using Supermarket Scanner Data," *Marketing Science*, 16 (4), 315-337.
- Murphy, David (2008), "A Measured Approach," *Marketing*, March 26, 47-49.
- Naik, Prasad A., and Kay Peters (2009), "A Hierarchical Marketing Communications Model of Online and Offline Media Synergies," *Journal of Interactive Marketing*, 23 (4), 288-299.
- Neslin, Scott A., D. Grewal, R. Leghorn, V. Shankar, M. L. Teerling, J. S. Thomas, P. C. Verhoef (2006), "Challenges and Opportunities in Multichannel Management." *Journal of Service Research*, 9 (2) 95-113.
- Novak, Thomas P., and Donna L. Hoffman (2003), "The Influence of Goal-Directed and Experiential Activities on Online Flow Experiences," *Journal of Consumer Psychology*, 13 (1&2), 3-16.
- Pauwels, Koen, Dominique M. Hanssens, and S. Siddarth (2002), "The Long-Term Effects of Price Promotions on Category Incidence, Brand Choice, and Purchase Quantity," *Journal of Marketing Research*, 39 (November), 421-39.
- Pauwels, Koen, Peter S.H. Leeflang, Marije L. Teerling, and K.R. Eelko Huizingh (2011), "Does Online Information Drive Offline Revenues? Only for Specific Products and Consumer Segments!" *Journal of Retailing*, 87 (1), 1-17.
- Raftery, Adrian E., and Steven M. Lewis (1996), "Implementing MCMC," in *Markov Chain Monte Carlo in Practice*, Walter R. Gilks, Sylvia Richardson, and David J. Spiegelhalter, eds. London: Chapman and Hall.
- Rossi, Peter. E., Greg M. Allenby, Robert McCulloch (2005), *Bayesian Statistics and Marketing*, John Wiley & Sons, New York.
- Rutz, Oliver J., Michael Trusov, and Randolph E. Bucklin (2011), "Modeling Indirect Effects of Paid Search Advertising: Which Keywords Lead to More Future Visits?," *Marketing Science*, 30 (4), 646-665.
- Schultz, Randall L. and Dick R. Wittink (1976), "The measurement of industry advertising effects," *Journal of Marketing Research*, 13(1), 71-75.
- Sethuraman, Raj, Gerard J. Tellis, Richard A. Briesch (2011), "How Well Does Advertising Work? Generalizations from Meta-analysis of Brand Advertising Elasticities," *Journal of Marketing Research*, 48 (3), 457-471.
- Sonnier, Garrett P., Leigh McAlister, and Oliver J. Rutz (2011), "A Dynamic Model of the Effect of Online Communications on Firm Sales," *Marketing Science*, 30 (4), 702-716.

- The Financial Times (2008), "Marketing Budget Cuts Predict to Deepen," July 13.  
<http://www.marketingcharts.com/interactive/us-online-advertising-market-to-reach-50b-in-2011-3128/>
- Trusov, Michael, Randolph E. Bucklin, and Koen Pauwels (2009), "Effects of Word-of-Mouth Versus Traditional Marketing: Findings from an Internet Social Networking Site," *Journal of Marketing*, 73 (September), 90-102.
- To, Pui-Lai, Chechen Liao, and Tzu-Hua Lin (2007), "Shopping Motivations on Internet: A Study Based on Utilitarian and Hedonic Value," *Technovation*, 27 (12), 774-787.
- Van Heerde, Harald J. and Tammo H.A. Bijmolt (2005), "Decomposing the Promotional Revenue Bump for Loyalty Program Members versus Non-Members," *Journal of Marketing Research*, 42 (4), 443-457.
- Van Heerde, Harald J., Carl F. Mela, and Puneet Manchanda (2004), "The Dynamic Effect of Innovation on Market Structure," *Journal of Marketing Research*, 41 (May), 166-83.
- Van Nierop, J.E.M., P.S.H. Leeflang, M.L. Teerling and K.R.E. Huizingh (2011), "The impact of the Introduction and Use of an Informational Website on Offline Customer Buying Behavior," *International Journal of Research in Marketing*, 28 (2), 155-165.
- Verhoef, Peter C., Scott A. Neslin, and Björn Vroomen (2007), "Multichannel Customer Management: Understanding the Research-Shopper Phenomenon," *International Journal of Research in Marketing*, 24 (2), 129-148.
- West, Mike and Jeff Harrison (1999), *Bayesian Forecasting and Dynamic Models*, 2nd ed. New York: Springer-Verlag.
- Wiesel, Thorsten, Koen Pauwels, and Joep Arts (2011), "Marketing's Profit Impact: Quantifying Online and Offline Funnel Progression," *Marketing Science*, 30 (4), 604-611.
- Wolfenbarger, Mary, and Mary C. Gilly (2001), "Shopping Online for Freedom, Control, and Fun," *California Management Review*, 43 (2), 34-55.
- Yang, Sha (2011), Personal correspondence on search advertising ROI, Marshall School of Business, University of Southern California, October 24, 2011.

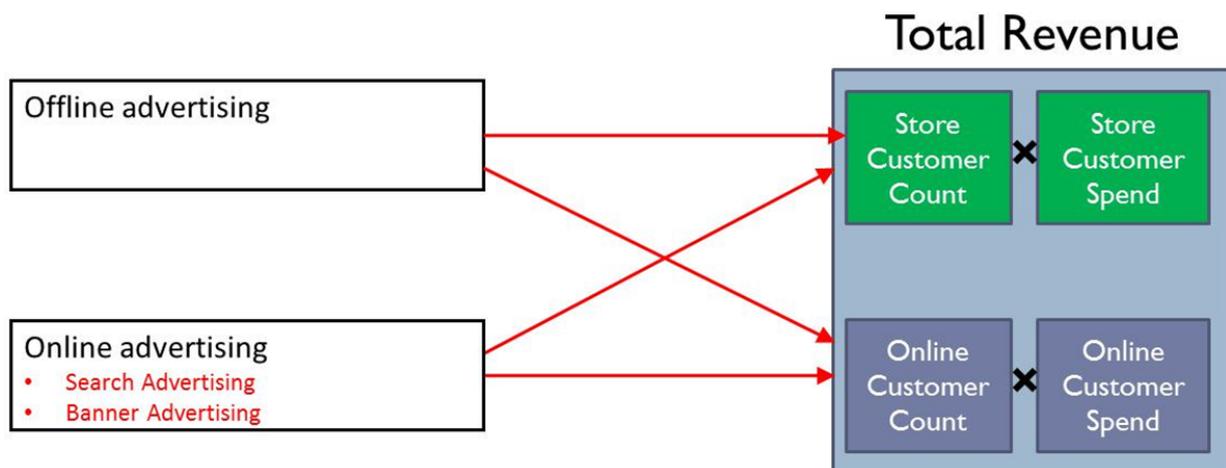
**Figure 1:**  
**Literature (Selected Studies) about the Effects of**  
**On- and Offline Advertising on On- and Offline Sales**



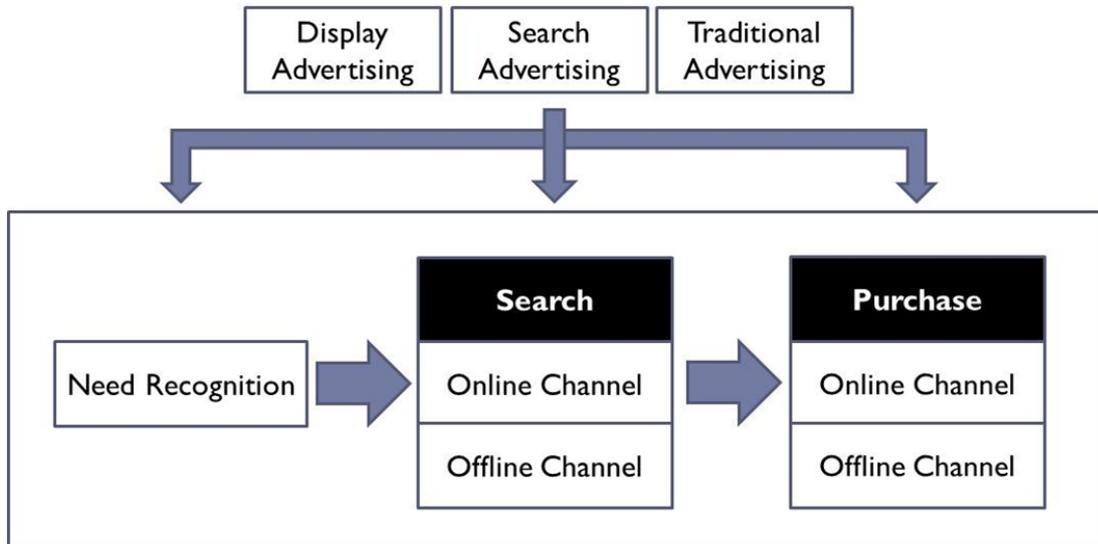
**Figure 2:**  
**Total Revenues and its Constituent Components**



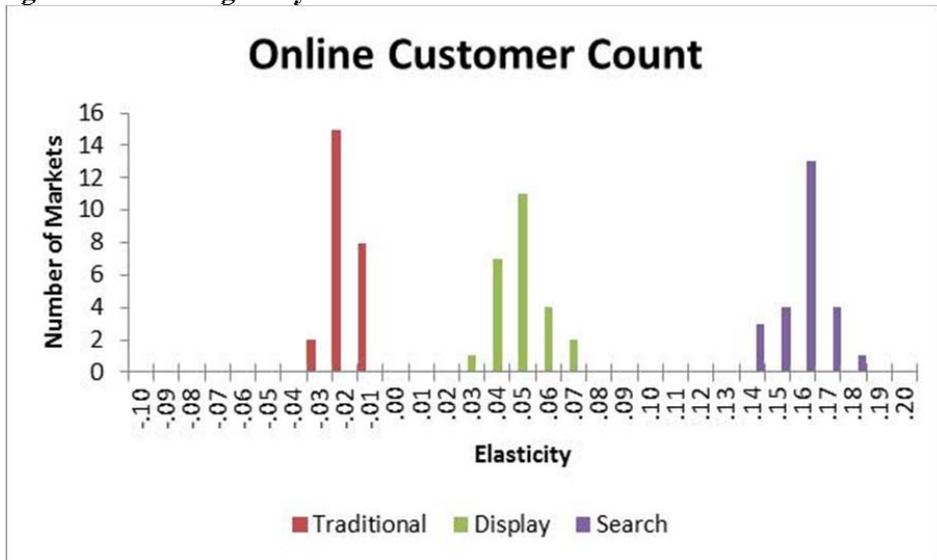
**Figure 3:**  
**Conceptual Framework for the Effects of**  
**On- and Offline Advertising on On- and Offline Sales**



**Figure 4:**  
**Framework Describing how Customers Make Purchase Decisions in a Multichannel Environment**



**Figure 5:**  
**Histograms of Heterogeneity in the Elasticities of Store and Online Customer Counts**



**Table 1:  
Research that Investigates Offline and Online Media Effectiveness**

Paper	Dependent Variable	Sales Decomposition	Own-Channel Effects		Cross-Channel Effects	
			Offline→ Offline	Online→ Online	Offline→ Online	Online→ Offline
Assmus et al. 1984	Sales		Yes			
Sethuraman et al. 2011	Sales		Yes			
Lodish et al. 1995	Sales		Yes			
Naik and Peters 2009	Consideration			Yes	Yes	
Sonnier et al. 2011	Sales			Yes		
Pauwels et al. 2011	Sales					Yes
Manchanda et al. 2006	Sales			Yes		
Rutz et al. 2011	Website visitors			Yes		
Van Nierop et al. 2011	Sales					Yes
Trusov et al. 2009	Sales			Yes	Yes	
Abraham 2008	Sales		Yes	Yes	Yes	Yes
Wiesel et al. 2011	Sales Leads, Quotes, Profits		Yes	Yes	Yes	Yes
Ansari et al. 2008	Incidence, Choice, Quantity		Yes	Yes	Yes	Yes
This Paper	Sales Customer Count, Customer Spend	Yes	Yes	Yes	Yes	Yes

**Table 2:  
Model Challenges and Solutions**

<b>Challenge</b>	<b>Solution</b>
Endogeneity in Advertising	Bayesian Instrumental Variable estimation
Dynamic Advertising Effects	Dynamic Linear Model for AdStock for traditional advertising, online display advertising, and online search advertising
Heterogeneity Across markets	Hierarchical Bayes
Multivariate dependent variable across revenue components and markets	Fully correlated errors
Competitive effects	Competitor advertising effects

**Table 3:  
Variable Operationalizations**

<b>Variable</b>	<b>Operationalization</b>	<b>Source</b>
OfflineSales	Aggregate dollar value of store sales in a particular market.	Company
OfflineCustomerCount	Total number of unique in-store transactions	Company
OfflineCustomerSpend	Average dollar value per in-store transaction	Company
OnlineSales	Aggregate dollar value of online sales in a particular market	Company
OnlineCustomerCount	Total number of unique online transactions	Company
OnlineCustomerSpend	Average dollar value per online transaction	Company
TraditionalAdvertising	Total national dollar spend on Newspaper, Magazine, Radio, Television and Billboard advertising	TNS
OnlineDisplayAdvertising	Total national dollar spend on internet display advertising	TNS
OnlineSearchAdvertising	Total national dollar spend on search advertising	Company
DummyTestBannerAdvertising	A binary variable that is 1 during weeks of additional geo-targeted banner advertising	Company
ChristmasDummy	A binary variable that is 1 for weeks including December (i.e. the last 5 weeks of the year)	Company
Promotion	The number of individual promotion days in a week. Multi-day promotions are counted as many days as the promotion runs	Company
UnemploymentRate	Unemployment rate in the associated market	Bureau of Labor Statistics
CompetitorTraditionalAdvertising	Total national dollar spend on internet traditional advertising by the closest competitor	TNS
CompetitorOnlineDisplayAdvertising	Total national dollar spend on display advertising by the closest competitor	TNS
Market	Geographic region representing a service area as defined by the company	Company

**Table 4:**  
**Descriptive Statistics (per period)\***

<b>Variable</b>	<b>Mean</b>	<b>StdDev</b>
Offline Sales (Per Region)	324,685.3	304,340.3
OfflineCustomerCount (Per Region)	1,129.9	950.4
OfflineCustomerSpend (Average, Per Region)	287.4	77.7
Online Sales (Per Region)	71,425.7	74,192.6
OnlineCustomerCount (Per Region)	421.0	437.8
OnlineCustomerSpend (Average, Per Region)	169.6	24.5
TraditionalAdvertising (National)	47,861.8	52,384.8
OnlineDisplayAdvertising (National)	55,235.8	36,265.1
OnlineSearchAdvertising (National)	41,358.4	15,721.2
DummyTestBannerAdvertising	0.023	0.151
ChristmasDummy	0.097	0.296
Promotion	9.6	5.7
UnemploymentRate (%)	8.8	2.4
CompetitorTraditionalAdvertising	76,455.3	52,720.7
CompetitorOnlineDisplayAdvertising	351.6	709.5

\*: Customer Count, Customer Spend, Advertising and Competitor Advertising are all comparatively scaled to prevent the identity of the retailer.

N = 2575

**Table 5:**  
**The Effects of Control Variables ( $\beta$ ) on the Four Dependent Variables**

<b>Independent variable</b>	<b>Store Customer Count</b>	<b>Store Customer Spend</b>	<b>Online Customer Count</b>	<b>Online Customer Spend</b>
Intercept	4.635**	5.413**	3.267**	8.698**
DummyTestBannerAdvertising	0.039*	-0.013	0.035	-0.002
ChristmasDummy	0.257**	-0.039	0.322**	0.051
Promotion	0.013**	0.009*	0.010*	0.004
Trend	0.001	-0.002	-0.004	0.006
UnemploymentRate	-0.228**	0.078	-0.287**	0.059
CompetitorTraditionalAdvertising	-0.194**	-0.147**	-0.216**	-0.230**
CompetitorOnlineDisplayAdvertising	0.010	0.020**	0.025**	-0.005

\*\* : 95% highest posterior density excludes zero; \* 90% highest posterior density excludes zero.

**Table 6:**  
**Short-Term Advertising Elasticities ( $\eta$ ) for the Four Dependent Variables**

<b>Independent variable</b>	<b>Store Customer Count</b>	<b>Store Customer Spend</b>	<b>Online Customer Count</b>	<b>Online Customer Spend</b>
Traditional Advertising	0.009**	0.002	-0.021 **	0.030**
Online Display Advertising	0.052**	0.001	0.045 **	0.014**
Online Search Advertising	0.107**	0.077**	0.153 **	-0.044**

\*\* : 95% highest posterior density excludes zero; \* 90% highest posterior density excludes zero.

**Table 7:**  
**Advertising Carry-Over Coefficients ( $\lambda$ ) and Duration Intervals for the Four Dependent Variables**

<b>Independent variable</b>	<b>Store Customer Count</b>		<b>Store Customer Spend</b>		<b>Online Customer Count</b>		<b>Online Customer Spend</b>	
	90% duration interval		90% duration interval		90% duration interval		90% duration interval	
	$\lambda$	(weeks)	$\lambda$	(weeks)	$\lambda$	(weeks)	$\lambda$	(weeks)
Traditional Advertising	0.63	4.9	0.60	4.5	0.74	7.8	0.60	4.5
Online Display Advertising	0.61	4.7	0.59	4.4	0.65	5.3	0.57	4.1
Online Search Advertising	0.65	5.4	0.63	5.1	0.68	6.0	0.61	4.7

**Table 8:**  
**Long-Term Advertising Elasticities [ $\eta/(1-\lambda)$ ] for the Four Dependent Variables**

<b>Independent variable</b>	<b>Store Customer Count</b>	<b>Store Customer Spend</b>	<b>Online Customer Count</b>	<b>Online Customer Spend</b>
Traditional Advertising	0.025**	0.006	-0.082 **	0.077**
Online Display Advertising	0.135**	0.002	0.126 **	0.032**
Online Search Advertising	0.304**	0.210**	0.477 **	-0.115**

\*\* : 95% highest posterior density excludes zero; \* 90% highest posterior density excludes zero.

The long-term elasticity ( $\eta/(1-\lambda)$ ) is computed for each draw and the table shows its posterior median.

**Table 9:  
Total Short-Term Advertising Elasticities and Their Decomposition**

Independent variable	Store Customer Count	Store Customer Spend	Total Store Sales Elasticity	Online Customer Count	Online Customer Spend	Total Online Sales Elasticity	Total Sales Elasticity <sup>1</sup>
	(a)	(b)	(a+b)	(c)	(d)	(c+d)	$\frac{\%Revenues_{offline} * (a+b) + \%Revenues_{online} * (c+d)}{\%Revenues_{online} * (c+d)}$
Traditional Advertising	0.009	0.002	0.011	-0.021	0.030	0.009	0.011
Online Display Advertising	0.052	0.001	0.053	0.045	0.014	0.058	0.054
Online Search Advertising	0.107	0.077	0.184	0.153	-0.044	0.108	0.172

<sup>1</sup> In the calculation we use  $\%Revenues_{offline} = 0.85$  ( $\%Revenues_{online} = 0.15$ ), because in the application 85% of sales goes through the offline channel.

**Table 10:  
Total Long-Term Advertising Elasticities and Their Decomposition**

Independent variable	Store Customer Count	Store Customer Spend	Total Store Sales Elasticity	Online Customer Count	Online Customer Spend	Total Online Sales Elasticity	Total Sales Elasticity <sup>1</sup>
	(a)	(b)	(a+b)	(c)	(d)	(c+d)	$\frac{\%Revenues_{offline} * (a+b) + \%Revenues_{online} * (c+d)}{\%Revenues_{online} * (c+d)}$
Traditional Advertising	0.025	0.006	0.031	-0.082	0.077	-0.005	0.025
Online Display Advertising	0.135	0.002	0.138	0.126	0.032	0.158	0.141
Online Search Advertising	0.304	0.210	0.514	0.477	-0.115	0.362	0.491

<sup>1</sup> In the calculation we use  $\%Revenues_{offline} = 0.85$  ( $\%Revenues_{online} = 0.15$ ), because in the application 85% of sales goes through the offline channel.

**Table 11:  
Hypothesis Testing**

H1: For display, search, and traditional advertising, cross- and own-effects will be larger for customer counts than for customer spend.			
Conditions	$\eta_{\text{display} \rightarrow \text{offline}, \text{count}} > \eta_{\text{display} \rightarrow \text{offline}, \text{ordersize}}$	0.052 > 0.001	Yes*
	$\eta_{\text{search} \rightarrow \text{offline}, \text{count}} > \eta_{\text{search} \rightarrow \text{offline}, \text{ordersize}}$	0.107 > 0.077	Yes
	$\eta_{\text{traditional} \rightarrow \text{offline}, \text{count}} > \eta_{\text{traditional} \rightarrow \text{offline}, \text{ordersize}}$	0.009 > 0.002	No
	$\eta_{\text{display} \rightarrow \text{online}, \text{count}} > \eta_{\text{display} \rightarrow \text{online}, \text{ordersize}}$	0.045 > 0.014	Yes
	$\eta_{\text{search} \rightarrow \text{online}, \text{count}} > \eta_{\text{search} \rightarrow \text{online}, \text{ordersize}}$	0.153 > -0.044	Yes
	$\eta_{\text{traditional} \rightarrow \text{online}, \text{count}} > \eta_{\text{traditional} \rightarrow \text{online}, \text{ordersize}}$	-0.021 > 0.030	No
H2: There will be cross-effects, whereby online advertising influences offline purchases, and offline advertising influences online purchases.			
Conditions	$\eta_{\text{display} \rightarrow \text{offline}, \text{count}} \neq 0$	0.052 $\neq$ 0	Yes
	$\eta_{\text{search} \rightarrow \text{offline}, \text{count}} \neq 0$	0.107 $\neq$ 0	Yes
	$\eta_{\text{traditional} \rightarrow \text{online}, \text{count}} \neq 0$	-0.021 $\neq$ 0	Yes
H3: For all media, own-effects will be larger than cross-effects.			
Conditions	$\eta_{\text{display} \rightarrow \text{online}, \text{count}} > \eta_{\text{display} \rightarrow \text{offline}, \text{count}}$	0.045 > 0.052	No
	$\eta_{\text{search} \rightarrow \text{online}, \text{count}} > \eta_{\text{search} \rightarrow \text{offline}, \text{count}}$	0.153 > 0.107	Yes
	$\eta_{\text{traditional} \rightarrow \text{offline}, \text{count}} > \eta_{\text{traditional} \rightarrow \text{online}, \text{count}}$	0.009 > -0.021	Yes
H4: The cross-effect from display advertising to offline purchasing is larger than the cross-effect from traditional advertising to online purchasing.			
Condition:	$\eta_{\text{display} \rightarrow \text{offline}, \text{count}} > \eta_{\text{traditional} \rightarrow \text{online}, \text{count}}$	0.052 > -0.021	Yes
H5: The differential effect on online versus offline purchasing will be larger for search advertising than for display advertising.			
Condition:	$\eta_{\text{search} \rightarrow \text{online}, \text{count}} - \eta_{\text{search} \rightarrow \text{offline}, \text{count}} > \eta_{\text{display} \rightarrow \text{online}, \text{count}} - \eta_{\text{display} \rightarrow \text{offline}, \text{count}}$	0.153 - 0.107 > 0.045 - 0.052	Yes

\* A "Yes" means that the difference in the predicted direction and that the 95% posterior density of the relevant difference excludes zero. A "No" means that the 95% posterior density includes zero or that the difference is opposite to the predicted direction.

**Table 12:  
Profit Return on Investment**

<b>Short Term ROI</b>	<b>Traditional</b>	<b>Online Display</b>	<b>Online Search</b>
Offline Sales	-\$0.32	\$1.84	\$12.15
Online Sales	-\$0.88	-\$0.35	\$0.62
<b>Long Term ROI</b>	<b>Traditional</b>	<b>Online Display</b>	<b>Online Search</b>
Offline Sales	\$0.92	\$6.38	\$35.72
Online Sales	-\$1.07	\$0.78	\$4.45

## APPENDIX 1: ELASTICITY DECOMPOSITION

The elasticity of revenues with respect to a particular advertising vehicle can be obtained by

computing  $\left(\frac{\partial \text{Revenues}}{\partial \text{Advertising}}\right) \times \left(\frac{\text{Advertising}}{\text{Revenues}}\right)$ , which is equal to:

$$\begin{aligned}
 &= \left( \frac{\partial [\text{OnlineCustomerCount} \times \text{OnlineCustomerSpend} + \text{OfflineCustomerCount} \times \text{OfflineCustomerSpend}]}{\partial \text{Advertising}} \right) \\
 &\times \left( \frac{\text{Advertising}}{\text{Revenues}} \right) \\
 &= \left( \left( \frac{\partial \text{OnlineCustomerCount}}{\partial \text{Advertising}} \right) \times \text{OnlineCustomerSpend} + \left( \frac{\partial \text{OnlineCustomerSpend}}{\partial \text{Advertising}} \right) \times \text{OnlineCustomerCount} + \right. \\
 &\left. \left( \frac{\partial \text{OfflineCustomerCount}}{\partial \text{Advertising}} \right) \times \text{OfflineCustomerSpend} + \left( \frac{\partial \text{OfflineCustomerSpend}}{\partial \text{Advertising}} \right) \times \text{OfflineCustomerCount} \right) \times \\
 &\quad \left( \frac{\text{Advertising}}{\text{Revenues}} \right) \\
 &= \left( \left( \frac{\partial \text{OnlineCustomerCount}}{\partial \text{Advertising}} \frac{\text{Advertising}}{\text{OnlineCustomerCount}} \right) \times \text{OnlineCustomerCount} \times \text{OnlineCustomerSpend} + \right. \\
 &\quad \left( \frac{\partial \text{OnlineCustomerSpend}}{\partial \text{Advertising}} \frac{\text{Advertising}}{\text{OnlineCustomerSpend}} \right) \times \text{OnlineCustomerCount} \times \text{OnlineCustomerSpend} + \\
 &\quad \left( \frac{\partial \text{OfflineCustomerCount}}{\partial \text{Advertising}} \frac{\text{Advertising}}{\text{OfflineCustomerCount}} \right) \times \text{OfflineCustomerCount} \times \text{OfflineCustomerSpend} + \\
 &\quad \left. \left( \frac{\partial \text{OfflineCustomerSpend}}{\partial \text{Advertising}} \frac{\text{Advertising}}{\text{OfflineCustomerSpend}} \right) \times \text{OfflineCustomerCount} \times \text{OfflineCustomerSpend} \right) \times \\
 &\quad \left( \frac{1}{\text{Revenues}} \right)
 \end{aligned}$$

If we calculate the elasticity at the mean (online and offline) revenues, we obtain:

$$\begin{aligned}
 &= \overline{\% \text{Revenues}}_{\text{online}} (\eta_{\text{Advertising} \rightarrow \text{online, count}} + \eta_{\text{Advertising} \rightarrow \text{online, spend}}) + \\
 &\quad \overline{\% \text{Revenues}}_{\text{offline}} (\eta_{\text{Advertising} \rightarrow \text{offline, count}} + \eta_{\text{Advertising} \rightarrow \text{offline, spend}})
 \end{aligned}$$

## APPENDIX 2: MODEL ESTIMATION

For completeness sake, we repeat the model below. The observation equation of our Dynamic Linear Model is:

$$(A1) \quad y_t = F_t' \theta_t + X_t' \beta + v_t; \quad v_t \sim N(0, V),$$

and the state equation is:

$$(A2) \quad \theta_t = G \theta_{t-1} + Z_t' \eta; \quad \omega_t \sim N(0, W).$$

For parameter vector  $\beta = (\beta_1, \dots, \beta_m, \dots, \beta_M, \gamma_1, \dots, \gamma_5)$ , we assume that the market-specific parameters are draws from national hypermeans:

$$\beta_m = (\beta_{1,m}, \beta_{2,m}, \beta_{3,m}, \beta_{4,m})' \sim N(\bar{\beta}, V_\beta). \text{ Parameters } \lambda_m \text{ are on the diagonal of } G, \text{ and } \lambda_m = (\lambda_{1,1m}, \dots, \lambda_{4,3m})' \sim N(\bar{\lambda}, V_\lambda). \text{ Parameter vector } \eta = (\eta_1, \dots, \eta_m, \dots, \eta_M)', \text{ and } \eta_m = (\eta_{1,1m}, \dots, \eta_{4,3m})' \sim N(\bar{\eta}, V_\eta).$$

We estimate model (A1)-(A2) by a combination of forward filtering, backward sampling algorithm (Carter and Kohn 1994, Frühwirth-Schnatter 1994) and Gibbs sampling (Chib and Greenberg 1995). The forward filtering equations assume that  $\beta$ ,  $V$ ,  $\lambda_m$ ,  $\eta$ , and  $W$  are known, an assumption we shall relax shortly (West and Harrison 1999, p. 103-104).

### Step 1) Forward filtering, backward sampling.

*Forward filtering:*

Posterior at  $t-1$ :

$$(A3) \quad \theta_{t-1} | D_{t-1} \sim N(m_{t-1}, C_{t-1}) \text{ where } D_{t-1} \text{ is the information available at } t-1.$$

Prior at  $t$ :

$$(A4) \quad \theta_t | D_{t-1} \sim N(a_t, R_t) \text{ where } a_t = Z_t' \eta + G m_{t-1} \text{ and } R_t = G C_{t-1} G' + W.$$

One-step forecast at  $t$ :

$$(A5) \quad \tilde{y}_t | D_{t-1} \sim N(f_t, Q_t) \text{ where } \tilde{y}_t = y_t - X_t' \beta, f_t = F_t' a_t \text{ and } Q_t = F_t' R_t F_t + V.$$

Posterior at  $t$ :

$$(A6) \quad \theta_t | D_t \sim N(m_t, C_t), \text{ where } m_t = a_t + A_t(\tilde{y}_t - f_t), C_t = R_t - A_t Q_t A_t', \text{ and } A_t = R_t F_t Q_t^{-1}.$$

*Backward sampling*

The backward sampling part samples the parameters  $\theta_t$  as described by West and Harrison (1999, p. 570). We simulate the individual state vectors  $\theta_T, \theta_{T-1}, \dots, \theta_0$  as follows:

$$(A7) \quad \text{Sample } \theta_T \text{ from } (\theta_T | D_T) \sim N(m_T, C_T), \text{ then}$$

(A8) For each  $t = T-1, T-2, \dots, 1, 0$ , sample  $\theta_t$  from  $p(\theta_t | \theta_{t+1}, D_t)$ , where the conditioning value of  $\theta_{t+1}$  is the value just sampled. The required conditional distributions are:

$$(\theta_t | \theta_{t+1}, D_t) \sim N(g_t, M_t), \text{ where } g_t = m_t + B_t(\theta_{t+1} - a_{t+1}) \text{ and } M_t = C_t - B_t R_{t+1} B_t', \text{ with } B_t = C_t G R_{t+1}^{-1}. \text{ We use diffuse starting values: } m_0 = 0 \text{ and } C_0 = 10 * I.$$

As noted above, the process assumes that  $\beta, V, \lambda_m, \eta$ , and  $W$  are known. By using Gibbs sampling techniques, we can sample from each of these distributions described below.

**Step 2)  $V$ .** We allow for a full covariance matrix between the errors of the observation equations. We specify the prior on the covariance matrix  $V$  as Inverse Wishart  $(v_V, S_V)$ . The full conditional distribution for  $V$  is Inverse Wishart  $\sim$

$$(v_V + T, S_V + \sum_{t=1}^T (y_t - F_t' \theta_t - X_t' \beta)(y_t - F_t' \theta_t - X_t' \beta)').$$

We use a diffuse prior for  $V$ , with  $v_V = \dim(V) + 2$  and  $S_V = I_{\dim(V)}$ .

**Step 3)  $W$ .** We can write the element of (A2) referring to market  $m$ , element  $k$  as:

$\theta_{kmt} = \theta_{kmt-1} * \lambda_{km} + Z'_{kmt} \eta_{km} + \omega_{kmt}; \omega_{kmt} \sim N(0, W_{km})$ . The prior for  $W_{km}$  is Inverse Gamma  $(v_W/2, S_W/2)$ . The conditional distribution for  $W_{km}$  is Inverse Gamma  $\sim ((v_W + T)/2, S_W/2 + \sum_{t=1}^T (\theta_{kmt} - \theta_{kmt-1} * \lambda_{km} - Z'_{kmt} \eta_{mk})^2 / 2)$ . We choose a diffuse prior:  $v_W = 1$  and  $S_W = 10^{-4}$ .

**Step 4)  $\beta$ .** We assume a normal prior for  $\beta = (\beta_1, \dots, \beta_m, \dots, \beta_M, \gamma_1, \dots, \gamma_5)'$ :

$\beta \sim N(\mu_{0\beta}, \Sigma_{0\beta})$ . Rearranging equation (A1) yields  $\dot{y}_t \equiv y_t - F_t' \theta_t = X_t' \beta + \nu_t$ , where  $\nu_t \sim N(0, V)$ . As this is a standard SUR regression equation with normal errors, the full conditional posterior distribution is given by  $\beta \sim N(\mu_\beta, \Sigma_\beta)$ , where

$\Sigma_\beta = (\Sigma_{0\beta}^{-1} + \sum_{t=1}^T X_t V^{-1} X_t')^{-1}$  and  $\mu_\beta = \Sigma_\beta (\Sigma_{0\beta}^{-1} \mu_{0\beta} + \sum_{t=1}^T X_t V^{-1} \dot{y}_t)$ . We use a hierarchical prior for  $\beta_m$  and a diffuse prior for  $\gamma_1, \dots, \gamma_5$ :  $\mu_{0\beta} = (\bar{\beta}, \dots, \bar{\beta}, \dots, \bar{\beta}, 0, \dots, 0)$ , and  $\Sigma_{0\beta} =$

$$\begin{bmatrix} I_M \otimes V_\beta & 0 \\ 0 & 10^3 I \end{bmatrix}.$$

**Step 5)  $\bar{\beta}$ .** We assume a normal prior for  $\bar{\beta} \sim N(\mu_{0\bar{\beta}}, \Sigma_{0\bar{\beta}})$ . The posterior is given by

$\bar{\beta} \sim N(\mu_{\bar{\beta}}, \Sigma_{\bar{\beta}})$ , where  $\Sigma_{\bar{\beta}} = (\Sigma_{0\bar{\beta}}^{-1} + D_\beta' (I_{\dim(\bar{\beta})} \otimes V_\beta^{-1}) D_\beta)^{-1}$  and  $\mu_{\bar{\beta}} = \Sigma_{\bar{\beta}} (\Sigma_{0\bar{\beta}}^{-1} \mu_{0\bar{\beta}} + D_\beta' (I_{\dim(\beta)} \otimes V_\beta^{-1}) \beta)$ , where  $D_\beta = \iota_M \otimes I_{\dim(\bar{\beta})}$ . We use a diffuse prior for  $\bar{\beta}$ :  $\mu_{0\bar{\beta}} = 0$  and  $\Sigma_{0\bar{\beta}} = 10^3 I$ .

**Step 6)  $V_\beta$ .** We specify the prior on the covariance matrix  $V_\beta$  as Inverse Wishart  $(\nu_{V_\beta}, S_{V_\beta})$ . The full conditional distribution for  $V_\beta$  is Inverse Wishart  $\sim$

$$(\nu_{V_\beta} + M, S_{V_\beta} + \sum_{m=1}^M (\beta_m - \bar{\beta})(\beta_m - \bar{\beta})').$$

We use a diffuse prior for  $V_\beta$ , with  $\nu_{V_\beta} = \dim(V_\beta) + 2$  and  $S_{V_\beta} = 10^{-2} I_{\dim(V_\beta)}$ .

**Step 7)  $\lambda_m$ .** We assume a normal hierarchical prior for element  $k$ :  $\lambda_{km} \sim N(\bar{\lambda}_k, V_{\lambda_k})$ . We can write the element of (A2) referring to market  $m$ , element  $k$  as:  $\theta_{kmt} = \theta_{kmt-1} * \lambda_{km} + Z'_{kmt} \eta_{km} + \omega_{kmt}$ ;  $\omega_{kmt} \sim N(0, W_{km})$ . We can re-arrange this equation to obtain:  $\dot{\theta}_{kmt} \equiv \theta_{kmt} - Z'_{kmt} \eta_{km} = \theta_{kmt-1} * \lambda_{km} + \omega_{kmt}$ , which is a regression model with normal errors. Given that the prior and likelihood are normal, the full conditional posterior distribution is given by

$\lambda_{km} \sim N(\mu_{\lambda_{km}}, \Sigma_{\lambda_{km}})$ , where  $\Sigma_{\lambda_{km}} = (V_{\lambda_k}^{-1} + W_{km}^{-1})^{-1}$  and

$$\mu_{\lambda_{km}} = \Sigma_{\lambda_{km}} (V_{\lambda_k}^{-1} \bar{\lambda}_k + W_{km}^{-1} \sum_{t=1}^T \theta_{mt-1} \dot{\theta}_{mt}).$$

**Step 8)  $\eta_m$ .** We assume a normal hierarchical prior for element  $k$ :  $\eta_{km} \sim N(\bar{\eta}_k, V_{\eta_k})$ . We can write the element of (A2) referring to market  $m$ , element  $k$  as:  $\theta_{kmt} = \theta_{kmt-1} * \lambda_{km} + Z'_{kmt} \eta_{km} + \omega_{kmt}$ ;  $\omega_{kmt} \sim N(0, W_{km})$ . We can re-arrange this equation to obtain:  $\ddot{\theta}_{mt} \equiv \theta_{mt} - \theta_{kmt-1} * \lambda_{km} = Z'_{kmt} \eta_{km} + \omega_{kmt}$ , which is a regression model with normal errors. Given that the prior and likelihood are normal, the full conditional posterior distribution is given by

$\eta_{km} \sim N(\mu_{\eta_{km}}, \Sigma_{\eta_{km}})$ , where  $\Sigma_{\eta_{km}} = (V_{\eta_k}^{-1} + W_{km}^{-1})^{-1}$  and

$$\mu_{\eta_{km}} = \Sigma_{\eta_{km}} (V_{\eta_k}^{-1} \bar{\eta}_k + W_{km}^{-1} \sum_{t=1}^T Z'_{kmt} \ddot{\theta}_{mt}).$$

**Step 9)  $\bar{\lambda}$ .** We assume a normal prior for  $\bar{\lambda} \sim N(\mu_{0\bar{\lambda}}, \Sigma_{0\bar{\lambda}})$ . The posterior is given by

$\bar{\lambda} \sim N(\mu_{\bar{\lambda}}, \Sigma_{\bar{\lambda}})$ , where  $\Sigma_{\bar{\lambda}} = (\Sigma_{0\bar{\lambda}}^{-1} + D'_\lambda (I_{\dim(\bar{\lambda})} \otimes V_\lambda^{-1}) D_\lambda)^{-1}$  and  $\mu_{\bar{\lambda}} = \Sigma_{\bar{\lambda}} (\Sigma_{0\bar{\lambda}}^{-1} \mu_{0\bar{\lambda}} + D'_\lambda (I_{\dim(\lambda)} \otimes V_\lambda^{-1}) \lambda)$ , where  $D_\lambda = \iota_M \otimes I_{\dim(\bar{\lambda})}$ . We use a diffuse prior for  $\bar{\lambda}$ :  $\mu_{0\bar{\lambda}} = 0$  and  $\Sigma_{0\bar{\lambda}} = 10^3 I$ .

**Step 10)  $\bar{\eta}$ .** We assume a normal prior for  $\bar{\eta} \sim N(\mu_{0\bar{\eta}}, \Sigma_{0\bar{\eta}})$ . The posterior is given by

$\bar{\eta} \sim N(\mu_{\bar{\eta}}, \Sigma_{\bar{\eta}})$ , where  $\Sigma_{\bar{\eta}} = (\Sigma_{0\bar{\eta}}^{-1} + D'_\eta (I_{\dim(\bar{\eta})} \otimes V_\eta^{-1}) D_\eta)^{-1}$  and  $\mu_{\bar{\eta}} = \Sigma_{\bar{\eta}} (\Sigma_{0\bar{\eta}}^{-1} \mu_{0\bar{\eta}} + D'_\eta (I_{\dim(\eta)} \otimes V_\eta^{-1}) \eta)$ , where  $D_\eta = \iota_M \otimes I_{\dim(\bar{\eta})}$ . We use a diffuse prior for  $\bar{\eta}$ :  $\mu_{0\bar{\eta}} = 0$  and  $\Sigma_{0\bar{\eta}} = 10^3 I$ .

**Step 11)  $V_\lambda$ .** We specify the prior on the covariance matrix  $V_\lambda$  as Inverse Wishart  $(v_{V_\lambda}, S_{V_\lambda})$ . The full conditional distribution for  $V_\lambda$  is Inverse Wishart  $\sim$

$$(v_{V_\lambda} + M, S_{V_\lambda} + \sum_{m=1}^M (\lambda_m - \bar{\lambda})(\lambda_m - \bar{\lambda})').$$

We use a diffuse prior for  $V_\lambda$ , with  $v_{V_\lambda} = \dim(V_\lambda) + 2$  and  $S_{V_\lambda} = 10^{-2} I_{\dim(V_\lambda)}$ .

**Step 12)  $V_\eta$ .** We specify the prior on the covariance matrix  $V_\eta$  as Inverse Wishart  $(v_{V_\eta}, S_{V_\eta})$ . The full conditional distribution for  $V_\eta$  is Inverse Wishart  $\sim$

$$\left( v_{V_\eta} + M, S_{V_\eta} + \sum_{m=1}^M (\eta_m - \bar{\eta})(\eta_m - \bar{\eta})' \right).$$

We use a diffuse prior for  $V_\eta$ , with  $v_{V_\eta} = \dim(V_\eta) + 2$  and  $S_{V_\eta} = 10^{-2} I_{\dim(V_\eta)}$ .

We programmed steps 1-12 in Gauss 11. Extensive tests with simulated data show that the code retrieves the model parameters from data sets generated according to these parameters. The code is available on request. For the empirical model estimation, we use OLS-based starting values for the parameters.

## References

- Carter, C. and R. Kohn (1994), "On Gibbs Sampling for State Space Models," *Biometrika*, 81 (3), 541–553.
- Chib, Siddhartha and Edward Greenberg (1995), "Hierarchical Analysis of SUR Models with Extensions to Correlated Serial Errors and Time-Varying Parameter Models," *Journal of Econometrics*, 68 (2), 339-60.
- Frühwirth-Schnatter, Sylvia (1994), "Data Augmentation and Dynamic Linear Models," *Journal of Time Series Analysis*, 15 (2), 183–202.
- West, Mike and Jeff Harrison (1999), *Bayesian Forecasting and Dynamic Models*, 2nd ed. New York: Springer-Verlag.