Submitted to *SSRN*
manuscript draft

The Influence of Context on Online Shopping Behavior: The Case of Concurrent TV Consumption

Oliver Hinz
Goethe University Frankfurt, Faculty of Economics and Business Administration, Theodor-W.-Adorno-Platz 4, 60323 Frankfurt am Main, Germany, ohinz@wiwi.uni-frankfurt.de

Shawndra Hill
Microsoft Research NYC, 641 Avenue of the Americas, New York, NY 10023 U.S.A. shawndra@microsoft.com

Amit Sharma
Microsoft Research India, 9 Lavelle Road, Bengaluru, Karnataka 560001 India, amshar@microsoft.com

It is widely agreed that people use digital devices while watching TV, a phenomenon called "second screening". However, there is little work that explains the impact this shared attention across screens has on business. In this paper, we study the impact of concurrent TV viewership on online shopping behavior. Our first study analyzes the impact of TV consumption on online shopping behavior at the aggregate level using a panel of 100,000 US customers over a period of 24 months. We address potential endogeneity problems with an instrumental variable approach. Our second study proposes a novel approach to realizing individual-level analysis over thousands of consumers and products. We identify a sample of consumers who watched a certain TV show and analyze their Amazon shopping carts in the context of TV consumption. We then compare these results to the shopping behavior of the same sample of consumers one week before they watched the TV show. Both studies reveal that while TV consumption is correlated with a higher likelihood of online shopping in general, the causal effect of TV watching differs with respect to the complexity of the products purchased. If a TV program appeals to a large number of viewers, this results in fewer sales of high-complexity products and more sales of low-complexity products. New insights on this effect and information about its magnitude could help marketers better allocate their TV ad spending and make better online shopping recommendations for consumers who concurrently consume TV.

Key words: Online Sales, TV, Contextual Information, Personalization, Attention Economy, Instrumental Variable Approach, Big Data Analysis

---

1. Introduction

While both anecdotal evidence and data from research has established that TV watchers frequently engage in second screening, which is the concurrent use of a mobile device, laptop, or other computer while watching TV, it remains unclear whether this activity is beneficial or detrimental for online retailers. On the one hand, people are at home when they are second screening, so they are more likely to shop online as this is generally a context-specific activity (i.e., people are more likely to shop online if they are at home). Logic dictates that as the number of people shopping online increases, sales of products will also increase, indicating that second screening may be favorable for online retailers. On the other hand, according to the concept of the attention economy, people have a finite amount of attention and if they are watching TV, they may not have sufficient stores of attention to effectively engage in online shopping. This could perhaps lead to fewer sales, indicating that second screening may be unfavorable for online retailers.

The attention economy has been addressed extensively, both in the popular press and by research studies, and is an interesting concept to consider in the context of second screening behavior. The basis of the attention economy concept is that the amount of attention an individual has is fixed and thus the information presented to people needs to be managed (Davenport and Beck 2001, Goldhaber 1997, Simon 1996, Shapiro and Varian 1998). For example, if consumers receive numerous advertising messages, much of the information will fall on deaf ears, leading to inefficient marketing. During second screening, TV competes with social media like Facebook and Twitter, which in turn competes with online shopping. This competition may mean that a TV watcher's attention is spread thin and that one of these activities, in some cases online shopping, receives less attention. While the attention economy has been widely studied in general, only limited research has examined how second screening in particular affects online shopping. Past work has shown that people are less likely to participate in online auctions when they are watching TV, suggesting a negative relationship between TV viewership and this type of online sales. While this finding is important, it represents a gap in that it addresses only auctions and not other types of online shopping. Further work is required to assess how and to what degree second screening affects online shopping. To address this gap and build on existing related research, we seek to address the following research topics in this paper.

1. Does an increase in TV viewing affect conventional forms of online shopping?
2. If this is the case, does its influence vary across product types?

We hypothesize that people have an attention budget, thus making it difficult to focus on shopping for complex products when watching TV. This would mean that TV viewership has a stronger negative effect on sales of complex products compared to sales of non-complex products. Looking at this variance and relating it to the attention economy is a novel contribution of our research.

Previous research has shown that contextual factors can substantially influence consumer behavior. Context can influence preferences for brands or products because of perceived complementarities between context and choice Belk 1975. Further, individuals may behave differently and prefer different products or brands under different contexts (Lussier and Olshavsky 1979, Klein and Yadav 1989). In this case, we are interested in how context influences whether people buy products online, namely the effect of concurrent TV consumption on online shopping behavior and whether this context has an equal impact on all types of products.

It is clear from previous work that context matters, but new technologies now allow us to infer an individual's context from new data sources. With the availability of new types of data such as clickstream data, we can construct novel data sets to test the connections between TV viewership and online sales.

While our data sources are currently rather unique, further technology improvements will allow tracking individuals across devices and enable capturing this potentially important contextual information on a broader basis in the near future. This development is currently a hot industry topic in discussions such as those surrounding Alphonso, a start-up that collects TV-viewing data for advertisers through various apps that integrate the company's software. Using smartphone microphones, Alphonso's software can detail what people watch by identifying audio signals in TV ads and shows (Maheshwari 2017). While this development certainly warrants a vivid discussion on privacy, this is beyond the scope of the current article, which seeks to better understand the ways in which this contextual information is so valuable to industry.

For this purpose, we use two unique data sources that allow us to identify the effect of second screening on online shopping behavior, and we find variance in response based on the complexity of products. Concretely, our results indicate that recommendations or advertisements on websites should differ, dependent on whether an individual is watching TV and engaging in second screening behavior. Our findings on the influence of concurrent TV consumption and internet usage on shopping behavior could thus be used to optimize personalization.

The remainder of this paper is organized as follows: First, we review the existing literature related to the influence and importance of second screening contexts on sales. The subsequent section discusses our first empirical study, which analyzes the impact of TV consumption on online shopping behavior at an aggregate level. We examine a panel of 100,000 individuals over 24 months and address potential endogeneity problems with an instrumental variable approach. The next section presents our second, complementary approach, which examines individual-level behavior. We conclude with a discussion and summary of the theoretical contributions and managerial implications of our findings and provide suggestions for further research.

2. Previous Research

In the following, we discuss prior research related to the content of this article. We survey papers that examine the interplay between regular TV programs and online sales, highlighting the research gap this paper aims to close in order to improve our understanding of the impact of second screening contexts on online shopping behavior.

2.1. TV and Sales

There is a long history of research on the influence of TV advertising on sales (e.g., Lodish et al. 1995). We deliberately do not integrate our paper in this stream as we are not interested in the effectiveness or efficiency of advertising. In contrast, we believe that the attractiveness of TV consumption itself may influence online shopping behavior. This has not been examined to the best of our knowledge.

Only Hinz et al.(2015), who examined the impact of TV on online auctions in Germany over a period of about five years, have addressed a closely related research question. The authors found a negative causal effect of TV viewership on online auction sales, even after controlling for covariates. We extend this work, examining the impact of TV viewership not on auctions, which typically require a high level of attention, but instead on general online retail shopping, by examining a dataset provided by comScore containing information on the behavior of over 100,000 consumers in the United States.

Moreover, we address a potential problem in Hinz et al.'s (2015) work caused by uncontrolled heterogeneity due to location, specifically that TV viewership patterns from a set of users in one part of the country may be correlated with the online auction behavior of users from another part of the country. For example, residents in rural areas may stay at home and watch TV after work since there are few entertainment options available, while urban dwellers may spend their after-work hours outside the house, at restaurants for example, which limits their online auction participation during this time. This relationship could, in principle, lead to a spurious correlation problem because TV viewership of the rural population could produce a false negative correlation with the online activity of city dwellers, theoretically leading to a reported negative relationship. We address this problem, as the dataset of our first study consists of sales and TV viewership data at the ZIP-code level. To avoid the spurious correlation problem caused by differences in TV watching behavior between rural and urban dwellers, we restrict our analysis to consumers in the New York Designated Market Areas (DMA) defined by Nielsen.

In addition, with our second unique dataset that tracks individuals on a large internet browser, we move from solely aggregate-level analysis to individual-level analysis as well. This multi-method approach, which combines aggregate and individual-level data analyses, should help build confidence in the robustness of our findings.

Furthermore, we scrutinize the problem in more detail in both empirical studies and examine the effect of the TV consumption on different product types with respect to their documented characteristics. For this purpose, we rely on the classification results presented by Lovett et al. (2014), who offered 136 measures of brand characteristics for nearly 700 of the top U.S. brands.

These measures cover a broad range of characteristics, including brand personality, consumer satisfaction, age, and complexity. Especially for our purpose, the complexity of a product sold by a brand may be a moderator for online sales in the context of concurrent TV consumption. If a TV show attracts the attention of a consumer, it is possible that sales of high-complexity products may suffer from this loss of attention. On the other hand, sales of low-complexity products may not be affected, or consumers may make use of their time watching TV to order low-complexity products online. Accordingly, we expect subtle differences with respect to the influence of TV viewership on sales of high- and low-complexity products.

2.2. Attention Economy

The reason for this hypothesis is founded in the concept of the attention economy (Simon 1996). Simon argued that a world rich in information leads to a scarcity of "whatever that information consumes," in this case the attention of humans. Therefore, the idea behind the attention economy is that attention and hence the information that demands our attention need to be managed efficiently to avoid information overload (Davenport and Beck 2001, Goldhaber 1997, Simon 1996, Shapiro and Varian 1998). So far, one stream of research and practice has dealt with the problem of how to allocate information more efficiently, developing or examining applications to better control or customize information (Huberman and Wu 2007, Shapiro and Varian 1998. Falkinger (2007) developed a theoretical model describing the structure of competition for attention. Assuming an information-rich world with limited attention available, he found that international integration and progress in information technologies tends to decrease global diversity and the attention level of individuals.

From a marketing perspective, research on the attention economy is relevant, as marketers struggle with the problem of information overload. Consumers cannot process all the incoming information aimed at them, including advertising. Already in the 1990s, Kroeber-Riel (1987) found that only 5% of advertising reached its intended recipients. As a recently developed communication channel, the internet breaks the advertising mold. Consumers have access to a plethora of easily retrievable information such as news and ads. Consequently, media channels have to compete for the attention of consumers. Accordingly, attention to the internet appears to affect other communication channels such as newspapers and TV. Dimmick et al. (2004) showed that the internet has displaced traditional media in the daily news domain, with the largest displacement in newspapers and TV. Thus, attention to newspapers and TV providing daily news has shifted to the internet, leading to lower sales of printed media.

This paper focuses on the attention effects of traditional media, such as the effect of TV on the internet. Insights may be of high value as TV viewers are increasingly simultaneously using the internet, or second screening. According to a study by Accenture (2015), 87% of consumers use a second screen device while watching TV. However, academic studies are rather scarce in this area.

2.3. Product Attributes and Purchases

Previous studies have examined the interplay between sales and various product characteristics, but none have looked specifically at how second screening and TV watching context affect sales of products with different product characteristics. For example, Kushwaha and Shankar (2013) examined how the value of customer purchases varies by product category, namely hedonistic and utilitarian, and low- and high-risk products. While the authors do not explicitly evaluate high- and low-complexity products, their work demonstrates that online sales differ by product category, indicating that product attributes do affect sales. The division of goods into hedonic and utilitarian is the most common categorization in the literature examining consumer purchases. Differences between hedonic and utilitarian products and how these affect purchases of such goods have been commonly investigated in the consumer marketing and decision-making literature (e.g., Dhar and Wertenbroch 2000, Babin et al. 1994). However, the literature is lacking regarding how differences between low- and high-complexity goods affects retail purchases, whether online or in brick-and-mortar stores. While this topic has received little attention, it is particularly interesting, considering that people often search online while they are watching TV in response to ads. This paper is unique because it integrates consideration of product attributes with the effect of the attention economy on online sales.

Another important aspect of this paper is how people consider and shop for highly complex goods compared to less complex goods. Previous work has proposed a two-stage process for consideration of products, in which consumers evaluate different product attributes depending on the stage (Moe 2006). In the first stage, the goal is to reduce a large set of products to a smaller set; in the second stage, a much smaller number of alternatives is evaluated. Accordingly, simpler decision rules can be used earlier in the process, and more complete rules when making the final purchase decision, indicating that the final decision requires a significant amount of attention. Because there are more factors to consider when shopping for complex goods, making a purchase decision about such products may require more attention than for simple, less complex goods. It could be the case that people are unable to spread their attention across TV viewing and purchasing; on the other hand, people do engage in second screening so they may use this time to make purchases online even if they are splitting their attention.

3. Empirical Study 1: Aggregate-level Data Analysis

We start with an analysis at the aggregate level, which involves following a panel of users for two years and relating their online shopping behavior to TV consumption in the area of observation. Beside revealing the plain association between online shopping behavior and TV consumption, we attempt to identify the causal influence of TV consumption on online shopping through the use of instrumental variables.

3.1. Data and Descriptives

We compile data from three sources covering a two-year time span (2012-2013). First, we collect online transactions using comScore data from Wharton Research Data Services. These data are generated from a panel of about 100,000 users who opted in to the use of tracking software on their computers. Second, we acquired TV viewership data from Rentrak (now ComScore) that reports hourly viewership data in the United States by DMA. We use the data for the New York DMA. These Rentrak data are collected through cable TV partners that provide viewership data on many millions of TV viewers in the United States. Finally, we collect weather data for the New York area.

We aggregate the data using a two-hour base because of an observed a problem in the comScore data, namely a lag in the logging of data immediately before and after midnight. The number of transactions drops a few minutes before midnight and then peaks a few minutes after midnight. It appears that the transactions are queued because other processes that begin on the database server delay the immediate logging of data. The queue is then alleviated a few minutes later. To solve this problem, we aggregate on a two-hour base with a single period between 11 pm and 1 am. Thereby, we address the problem of delayed logging around midnight since we assign all delayed and lagged transactions to the same period.

We arrive at a dataset with 8,772 observation periods (12 observations per day for two years). We use a rich dataset from Rentrak on TV behavior in different regions of the United States. We add weather information for the New York area to the dataset. We use precipitation and temperature, as these can easily influence online sales, as well as temperature as a squared term, because online sales could increase if it is too hot or too cold to go outside. We further integrate dummies for public holidays in New York, extraordinary TV events like the Super Bowl, and uncommon winter weather like the extreme blizzard in February 2013 in the northeast United States, as these events could have a direct impact on online sales and we need to control for their potential effect.

As pointed out earlier, we distinguish products with respect to their complexity because we believe that high-complexity products may especially suffer from a limited attention budget on the consumer side. We classify products as high- (low-) complexity if the complexity score in Lovett et al. (2014) is higher (lower) than the average of all brands plus (minus) one standard deviation.

We then calculate the number of high/low-complexity products per time period and arrive at the distribution over time depicted in Figure 1. Total sales and sales of low-complexity products by time of day (left) peak between 7 and 11pm, which is an expected result.

TV viewership is also a function of time, which is represented in Figure 2. As seen in the figure, TV viewership peaks between 7 pm and 11 pm and is higher on Saturdays and Sundays than on weekdays, which is similar to results from online sales in Figure 1.

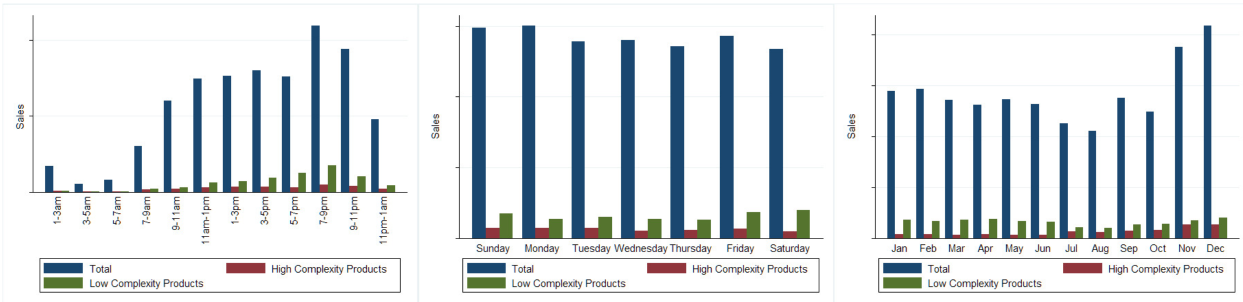

Figure 1 Total Sales and Sales of High and Low Complexity Products by Time of Day (left), Day of Week (middle), and Month (right)

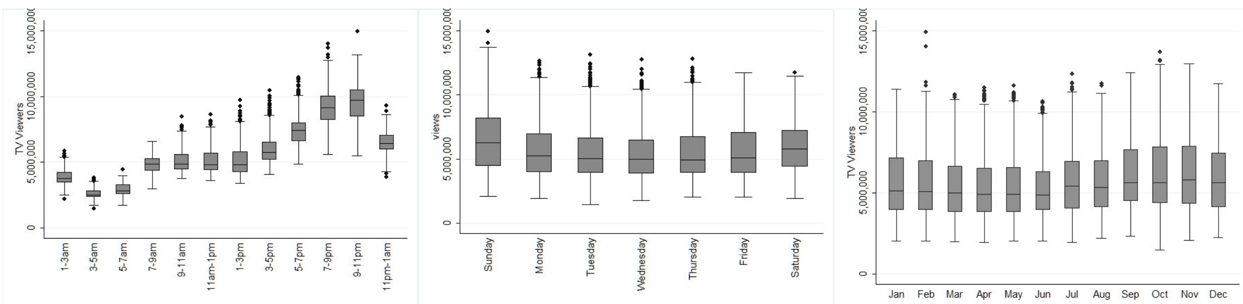

Figure 2 Number of TV Viewers by Time of Day (left), Day of Week (middle), and Month (right)

3.2. Model Specification

The dependent variable is the number of sales of either high- or low-complexity products in a given time period. To examine the effect of TV on online sales, we use the number of TV viewers, *TVViewer*, in a given period as the main explanatory variable of interest. To control for weather effects, we use *Precipitation* in mm and *Temperature* and *Temperature²* in degrees Celsius. We further include time variables to control for Weekday, Month, Time of Day, and Public Holiday effects. Equation (1) summarizes our basic model:

$$\begin{aligned} Sales_{lowC/highC,t} = {} & \alpha_0 + \alpha_1 \cdot TVViewer_t + \alpha_2 \cdot Precipitation + \alpha_3 \cdot Temperature_t + \\ & \alpha_4 \cdot Temperature_t^2 + \alpha_5 \cdot Superbowl_t + \alpha_6 \cdot Blizzard + \alpha_7 \cdot + PublicHoliday + \\ & \sum_{i=1}^{7} \alpha_{i+7} \cdot Weekday_{i,t} + \sum_{i=1}^{12} \alpha_{i+14} \cdot Month_{i,t} + \sum_{i=1}^{12} \alpha_{i+26} \cdot TimeofDay_{i,t} + \varepsilon_t \end{aligned} \tag{1}$$

3.3. Identification and Endogeneity

Spurious correlation is a common problem with time-series data. For technologically intensive goods, for example, the cost to produce and thus price generally decrease over time due to technological advances, while quantity increases over time. These correlations make it difficult to determine the extent to which increasing quantities are due to a growing user base or simply due to lower prices (Gowrisankaran and Stavins 2004).

In the context of our model, we emphasize that the problem is not econometric identification, which can always be achieved by choosing appropriately parsimonious functional forms, but the identification of causal effects on sales. In particular, the number of TV viewers may be endogenous. Therefore, we first need to consider potentially omitted variables and second, determine if there may be some effect of the dependent variable (sales) on the independent variables, causing so-called reverse causality bias.

Omitted variable bias results from correlations between the omitted cause, X_t, and included variables (Liu 2006). In our model, for example, the unobserved time spent at home, X_t, is likely to bias the results in a simple OLS regression. The condition of being at home could lead to a higher likelihood of online shopping and, at the same time, is likely to be correlated with the number of TV viewers. This may bias our inference for the effect of a specific TV program on sales. Additionally, there may be some effect of sales on the number of TV viewers, for example, due to people who turn off the TV when they are shopping online, which may additionally bias the results.

For causal identification, we employ a combination of strategies. First, we include variables that capture seasonality as well as daily, weekly, and monthly patterns. However, there is also the risk that the time dummies will overcontrol for system-specific factors that are legitimately part of the complementarity system we are examining. Thus, the coefficient estimates from such models may underestimate the true effects of the complements if we do not introduce orthogonal variance.

Second, we use an instrumental variable (IV) that is orthogonal to the terms of our system to identify variation in the number of TV viewers. It is always difficult to find valid instruments. Our explanatory variable of interest that is potentially endogenous is the number of TV viewers. It is easy to imagine that there are factors that influence the likelihood of simultaneously watching TV and shopping online. If we cannot control for these factors (e.g., the likelihood of being at home), the estimates are likely to be biased.

We thus need to find an instrumental variable that fulfills two requirements: Firstly, the instrument needs to be correlated with the endogenous explanatory variables conditional on the other covariates (i.e., something that is correlated with the number of TV viewers). This requirement can be easily tested by examining the significance of the IV in the first stage of a two-stage model.

Secondly, the IV should not correlate with the error term in the explanatory equation, which means that the instrument will not suffer from the same problem as the original predicting variable. However, it cannot be statistically tested whether this is true because the condition involves the unobservable residual. This condition has to be taken on faith, which is why theory or facts are very important for the analysis to be convincing.

A good IV should not be predictable, so that no actor (online consumer) driving the equation could anticipate and proactively change his behavior. Furthermore, the IV should not be influenced by the dependent variable (sales) or the endogenous explanatory factor (number of TV viewers), so it will thus be truly exogenous. For our research question, a disaster can constitute a valid IV. The occurrence of a disaster cannot be predicted and is usually covered in TV programs with special broadcasts. If these special broadcasts attract enough TV viewers, this could serve as a shock to the system. We can then use this shock to partial out the causal effect of TV viewership on online sales, if these sales are not directly influenced above and beyond the effect through the attention attracted by the TV program. We discuss potential confounding effects later.

During our observation period, the Boston Marathon bombing, a terrorist attack that occurred during the Boston Marathon on April 15, 2013, with three people killed and 264 injured, serves as such a disaster as it received heavy media coverage. We use this event as the first IV and a dummy variable for the eight hours of TV coverage after the explosion. As a second IV, we use eight hours on April 19 during which TV media reported on the unprecedented manhunt for the bombing suspect, with thousands of law enforcement officers searching a 20-block area of Watertown, Massachusetts.

Using two IVs has advantages, namely that it allows overidentification of the model and permits conducting the Sargan-Hansen test of overidentifying restrictions. If we observe a strong rejection of the null hypothesis with the Sargan-Hansen test, this would raise doubts about the validity of the estimates. These considerations lead to the two-stage model summarized by Equations (2) and (3):

$$\begin{aligned} TVViewerEst_t = {} & \beta_0 + \beta_1 \cdot IV1_t + \beta_2 \cdot IV2_t + \beta_3 \cdot Precipitation + \beta_4 \cdot Temperature_t + \\ & \beta_5 \cdot Temperature_t^2 + \beta_6 \cdot SuperBowl_t + \beta_7 \cdot Blizzard + \beta_8 \cdot PublicHoliday + \\ & \sum_{i=1}^{7} \alpha_{i+8} \cdot Weekday_{i,t} + \sum_{i=1}^{12} \alpha_{i+15} \cdot Month_{i,t} + \sum_{i=1}^{12} \alpha_{i+27} \cdot TimeOfDay_{i,t} + \varepsilon_t \end{aligned} \tag{2}$$

$$\begin{aligned} Sales_{lowC/highC,t} = {} & \gamma_0 + \gamma_1 \cdot TVViewerEst_t + \gamma_2 \cdot Precipitation + \gamma_3 \cdot Temperature_t + \\ & \gamma_4 \cdot Temperature_t^2 + \gamma_5 \cdot Superbowl_t + \gamma_6 \cdot Blizzard + \gamma_7 \cdot + PublicHoliday + \\ & \sum_{i=1}^{7} \gamma_{i+7} \cdot Weekday_{i,t} + \sum_{i=1}^{12} \gamma_{i+14} \cdot Month_{i,t} + \sum_{i=1}^{12} \gamma_{i+26} \cdot TimeofDay_{i,t} + \varepsilon_t \end{aligned} \tag{3}$$

3.4. Estimation Results

3.4.1. Ordinary Least Squares Estimates We first report the OLS estimates for descriptive purposes and then the estimates generated by the instrumental variable regression. For both low- and high-complexity products, we observe a positive effect of TV viewing on sales (p<.01, see Model (1) in Table 1 and Table 2). This means that online shopping sales increase with the number of TV viewers. This effect is most likely driven by the fact that people are more likely to watch

TV and shop online when they are at home. However, this does not indicate a causal effect of the attractiveness of a particular TV program on sales. We address this endogeneity issue with more sophisticated analysis methods in the following sections.

However, the OLS estimates show that online sales of low-complexity products are the highest between 5 pm and 9 pm and peak on Saturdays. In contrast, high-complexity products are sold more often during the week and experience a drop on Saturdays. We also see that sales in general are lower in the summer months compared December and that high-complexity products are less likely to be sold when the temperature is very high ($p<.1$). Interestingly, the Super Bowl increases the number of low-complexity products sold. This can be explained by looking at the classification by Lovett et al. (2014). According to this, many food and beverage brands are classified as low-complexity goods, and we typically expect a high number of orders for food and beverages during the Super Bowl.

3.4.2. Instrumental Variable Regression We estimate the model outlined in Equations (2) and (3) using extended instrumental variable regressions (see Baum and M. E. 2010) with heteroskedastic and autocorrelation consistent (HAC) standard errors and covariance estimation in STATA.

In the first stage, we determine the effect of exogenous events on the number of TV viewers (i.e., we estimate equation (2)). If these events are unrelated to online shopping above and beyond the effect caused by the attention attracted by the TV show, this approach helps us identify the causal effect of concurrent TV consumption on online shopping sales.

The estimates for the first stage (see Table 7 in Appendix) clearly show that the two events (i.e., the Boston Marathon bombing and the subsequent reporting on the hunt for the suspect) significantly increase the number of TV viewers ($p<.01$), which is the necessary condition for a valid IV approach. The first-stage estimates also show TV watching behavior over time. The number of TV viewers peaks between 7 pm and 9 pm and is usually higher on Saturdays and Sundays ($p<.01$). We observe the highest number of TV viewers in the fall months of September, October, and November. The number of viewers is also higher on public holidays ($p<<.01$), when on average 180,000 more people watch TV than on regular days in the New York area. Overall, the results provide face validity.

To test the suitability of our instruments, we further run an under-identification test, which is an LM test of whether the equation is identified (i.e., that the excluded instruments are "relevant"), meaning it is correlated with the endogenous regressors. Because we drop the i.i.d. assumption and use HAC statistics, we apply the Kleibergen and Paap (2006) rk LM statistic. We can reject the null hypothesis, which indicates that the matrix of regressors and instruments is of full column rank (i.e., the model is not underidentified (p<.05)). However, a rejection of the null hypothesis of this under-identification test does not rule out the possibility of weak instruments Hall et al. (1996). This problem arises when the excluded instruments are correlated with the endogenous regressors but only weakly (see e.g., Stock and Yogo (2005) for further discussion). We thus apply a weak instruments test based on the Kleibergen-Paap Wald rk F statistic and compare the values with the corresponding critical values compiled by Stock and Yogo (2005). The Kleibergen-Paap rk Wald F statistic is 49.541, which clearly indicates that our instruments are not within the set of weak instruments as defined by Stock and Yogo, in terms of both relative bias to OLS and bias in second-stage significance.

Since we have two exogenous instruments for endogenous TV viewership, we can also check Hansen's J statistic. The null hypothesis of the Hansen test is that the instruments are valid (i.e., uncorrelated with the error term) and that the excluded instruments are correctly excluded from the estimated equation. The value of Hansen's J statistic is 0.032 with $p \geq .85$, which is far from the value that would allow rejection of the null hypothesis, indicating that our instrument set is appropriate. Overall, all the statistical tests conducted allow us to conclude that the chosen IVs are appropriate.

We thus use these first-stage estimates in the second stage (see equation (3)) to identify the causal effect of TV viewership on online shopping sales. The F-values for all models (2), (3), and (4) in Table 1 (sales of low-complexity products) and Table 2 (sales of high-complexity products) allow us to reject the null hypothesis that the sets of coefficients are jointly zero (p<.01). In the following, we discuss the results of the full model (4) in Table 1 and Table 2 that incorporates all control variables, beginning with a discussion of low-complexity products.

For low-complexity products such as food and beverages, we observe a positive, albeit weakly significant effect of TV attractiveness on sales (p<.1). This means that the context of TV consumption can cause higher sales of this particular product type because it can be seen as a complement to TV. Obviously, this is in line with the popular assumption that higher online activity through second screening comes along with TV consumption. However, this is true only for a special subset of products, where the decision does not require significant cognitive effort by the consumer. Such purchase decisions can be made easily when watching TV. If we trust our IV approach, an additional one million TV viewers in the NY region would lead to a 0.883 increase of online sales of low-complexity products in a two-hour period (p<.1) in our sample. While this may seem economically insignificant, it should be considered in relative terms. If the number of TV viewers increases by 1%, the number of low-complexity products sold increases on average by 5%.

We still observe that sales are lower in the late summer and fall months compared to December but also observe that the effect of the Super Bowl becomes insignificant. This means that this effect is fully absorbed by the number of TV viewers, and that there is no effect above and beyond this attractiveness effect that leads to a higher number of TV viewers and ultimately to higher sales of low-complexity products.

However, this observed relationship does not hold for all product types. This finding is evident when looking at Table 2, where the dependent variable captures the number of high-complexity products sold. For high-complexity products, the relationship between TV viewers and sales is negative and highly significant ($p<.01$). If the TV program attracts many consumers, the likelihood for online sales of high-complexity products decreases significantly. Obviously, consumers are not able to spread their attention between TV and a complex purchase decision. A 1% increase of TV viewers thus leads to a decrease of sales by 0.7% for high-complexity products.

Among the control variables, we observe that there is an optimum temperature for online sales of high-complexity products. This relationship follows an inverted U-shape and by setting the first derivative to 0, we determine that the maximum is about 30°C, or 86°F. Interestingly, this finding is in line with research from medicine, where Grether (1973) suggested 80°F as the ideal environmental temperature for optimum performance of vigilance tasks and 85°F as the ideal temperature for optimum performance of other tasks Hancock and Vasmatzidis (2003).

Furthermore, the results show that the Super Bowl has a negative effect on sales of high-complexity products through the very high number of TV viewers but also a positive direct effect ($p<.05$) that counteracts the attention effect. This observation may be driven by a subsample of the population that is not interested in football and who uses the time to make complex purchase decisions. We revisit this point in our individual-level analysis in Section 4.

	(1) OLS Descriptive Estimates		(2) IV results, Only IV		(3) IV results, (2) & Time Controls		(4) IV results, (3) & Weather Controls	
Viewers in Mil	0.169***	(0.013)	0.687***	(0.277)	0.900*	(0.465)	0.883*	(0.466)
01:00-02:59 (0/1)	-0.541***	(0.058)			1.923	(1.5669)	1.882	(1.583)
03:00-04:59 (0/1)	-0.354***	(0.063)			3.131	(2.219)	3.052	(2.224)
05:00-06:59 (0/1)	-0.428***	(0.059)			2.814	(2.065)	2.741	(2.070)
07:00-08:59 (0/1)	-0.533***	(0.064)			1.319	(1.182)	1.281	(1.185)
09:00-10:59 (0/1)	0.473***	(0.067)			1.203	(1.071)	1.169	(1.074)
11:00-12:59 (0/1)	-0.201***	(0.068)			1.468	(1.066)	1.434	(1.070)
13:00-14:59 (0/1)	-0.105	(0.076)			1.511	(1.034)	1.478	(1.037)
15:00-16:59 (0/1)	-0.001	(0.078)			1.000	(0.645)	0.982	(0.647)
17:00-18:59 (0/1)	0.058	(0.090)			0.006	(0.114)	0.013	(0.112)
21:00-22:59 (0/1)	-0.539***	(0.106)			-2.197**	(1.065)	-2.153**	(1.063)
23:00-00:59 (0/1)	-0.591***	(0.068)			0.042	(0.414)	0.033	(0.416)
Sunday (0/1)	-0.222***	(0.063)			-0.662**	(0.300)	-0.668**	(0.300)
Monday (0/1)	-0.247***	(0.060)			-0.127	(0.100)	-0.152*	(0.092)
Tuesday (0/1)	-0.150**	(0.063)			0.075	(0.157)	0.046	(0.146)
Wednesday (0/1)	-0.197***	(0.059)			0.094	(0.193)	0.066	(0.184)
Thursday (0/1)	-0.222***	(0.060)			0.035	(0.174)	0.010	(0.166)
Friday (0/1)	-0.012	(0.063)			0.208	(0.160)	0.194	(0.155)
January (0/1)	-0.047	(0.077)			0.145	(0.146)	0.189	(0.177)
February (0/1)	-0.065	(0.079)			0.177	(0.173)	0.260	(0.229)
March (0/1)	0.003	(0.078)			0.337	(0.234)	0.272	(0.195)
April (0/1)	0.055	(0.084)			0.370	(0.237)	0.229	(0.152)
May (0/1)	-0.005	(0.092)			0.298	(0.231)	0.101	(0.124)
June (0/1)	-0.004	(0.106)			0.414	(0.305)	0.192	(0.172)
July (0/1)	-0.286**	(0.112)			-0.169	(0.132)	-0.388***	(0.148)
August (0/1)	-0.319***	(0.105)			-0.244**	(0.109)	-0.466***	(0.155)
September (0/1)	-0.290***	(0.100)			-0.555***	(0.168)	-0.755**	(0.326)
October (0/1)	-0.280***	(0.082)			-0.683***	(0.253)	-0.841**	(0.380)
November (0/1)	-0.146*	(0.082)			-0.410**	(0.186)	-0.457**	(0.223)
Public Holiday (0/1)	0.067	(0.099)			-0.077	(0.146)	-0.057	(0.139)
Super Bowl (0/1)	2.309***	(0.630)					-0.104	(1.892)
Blizzard (0/1)	-0.596**	(0.238)					-1.986**	(0.955)
Temperature deg. C	-0.004	(0.006)					0.026	(0.021)
Temperature deg. C^2	0.000	(0.000)					-0.001	(0.000)
Constant	0.260**	(0.115)	-3.265**	(1.572)	-5.329	(3.545)	-5.281	(3.620)
F	45.799		6.134		32.391		29.266	
R^2	0.124		0.002		0.138		0.140	
RMSE	1.376		1.827		1.630		1.615	

Table 1 Estimation Results for Low-Complexity Goods (Full Model - Results with and without Instrumental Variable)

	(1) OLS Descriptive Estimates		(2) IV results, Only IV		(3) IV results, (2) & Time Controls		(4) IV results, (3) & Weather Controls	
Viewers in Mil	0.075***	(0.007)	-0.072***	(0.014)	-0.133***	(0.028)	-0.124***	(0.026)
01:00-02:59 (0/1)	0.010	(0.031)			-0.688***	(0.105)	-0.666***	(0.100)
03:00-04:59 (0/1)	0.057*	(0.030)			-0.934***	(0.141)	-0.892***	(0.133)
05:00-06:59 (0/1)	0.031	(0.028)			-0.891***	(0.132)	-0.852***	(0.124)
07:00-08:59 (0/1)	0.025	(0.039)			-0.502***	(0.087)	-0.480***	(0.083)
09:00-11:59 (0/1)	0.071**	(0.033)			-0.406***	(0.080)	-0.386***	(0.076)
11:00-12:59 (0/1)	0.143***	(0.037)			-0.332***	(0.080)	-0.313***	(0.077)
13:00-14:59 (0/1)	0.164***	(0.039)			-0.296***	(0.080)	-0.278***	(0.077)
15:00-16:59 (0/1)	0.099**	(0.040)			-0.186***	(0.059)	-0.175***	(0.058)
17:00-18:59 (0/1)	-0.049	(0.037)			-0.033	(0.041)	-0.036	(0.041)
21:00-22:59 (0/1)	-0.129**	(0.053)			0.344***	(0.083)	0.320***	(0.079)
23:00-00:59 (0/1)	-0.058*	(0.034)			-0.238***	(0.049)	-0.232***	(0.048)
Sunday (0/1)	0.058*	(0.031)			0.189***	(0.038)	0.182***	(0.037)
Monday (0/1)	0.101***	(0.027)			0.070**	(0.028)	0.074***	(0.028)
Tuesday (0/1)	0.118***	(0.028)			0.056*	(0.030)	0.064**	(0.030)
Wednesday (0/1)	0.062**	(0.025)			-0.018	(0.028)	-0.011	(0.028)
Thursday (0/1)	0.066**	(0.027)			-0.004	(0.030)	0.001	(0.030)
Friday (0/1)	0.098***	(0.027)			0.035	(0.028)	0.041	(0.028)
January (0/1)	-0.379***	(0.053)			-0.419***	(0.055)	-0.445***	(0.056)
February (0/1)	-0.359***	(0.051)			-0.425***	(0.054)	-0.450***	(0.055)
March (0/1)	-0.357***	(0.048)			-0.470***	(0.054)	-0.432***	(0.052)
April (0/1)	-0.313***	(0.050)			-0.448***	(0.054)	-0.362***	(0.051)
May (0/1)	–0.314***	(0.051)			-0.470***	(0.054)	-0.344***	(0.054)
June (0/1)	-0.292***	(0.054)			-0.496***	(0.056)	-0.346***	(0.058)
July (0/1)	-0.176***	(0.065)			-0.306***	(0.057)	-0.148**	(0.069)
August (0/1)	-0.204***	(0.060)			-0.317***	(0.055)	-0.163***	(0.063)
September (0/1)	-0.202***	(0.060)			-0.206***	(0.059)	-0.072	(0.064)
October (0/1)	-0.210***	(0.059)			-0.154**	(0.062)	-0.053	(0.064)
November (0/1)	-0.010	(0.063)			0.049	(0.068)	0.076	(0.067)
Public Holiday (0/1)	-0.006	(0.056)			0.038	(0.058)	0.029	(0.059)
Super Bowl (0/1)	-0.236	(0.237)					0.437**	(0.211)
Blizzard (0/1)	-0.136	(0.110)					0.251*	(0.130)
Temperature deg. C	-0.006*	(0.004))					-0.015***	(0.004)
Temperature deg. C^2	0.000	(0.000)					0.000**	(0.000)
Constant	0.018	(0.064)	0.658***	(0.080)	1.577***	(0.223)	1.563***	(0.214)
F		16.000		27.559		16.164		14.458
R^2		0.072		0.0001		0.067		0.068
RMSE		0.751		0.827		0.792		10.787

Table 2 Estimation Results for High-Complexity Goods (Full Model - Results with and without Instrumental Variable)

3.5. Discussion of Potential Confounding Effects Regarding the IV

The use of IVs typically raises questions with regard to potential confounding effects that may bias the estimation results. In theory, we would need a perfectly orthogonal variable that is entirely uncorrelated with the error term of equation (1) and that fully influences the dependent variable through the endogenous variable TV viewership. While the impact on TV viewership can be tested in the first stage of a two-stage model, the orthogonality cannot be tested because it involves the unobserved error term. Therefore, it is necessary to discuss potential confounding effects and rely on theoretical arguments.

The question is thus whether the special broadcast on the Boston Marathon bombing influenced online shopping behavior in the New York City area above and beyond the distraction effect for which we argue. First, we can rule out anticipation effects: We can safely assume that no online shoppers in the New York City area changed their online shopping plans before the attack because this event was fully surprising and truly exogenous. Second, we can rule out reverse causality; online shopping sales did not influence the timing of the explosion. While there are certain advantages of the proposed instruments, there may still be some confounding effects, which we discuss in the following.

The main problem of the proposed IVs is that they could alter viewers' mood states. The Boston bombing event was certainly shocking and disturbing, and this could have had an impact on the mood of viewers and prospective online shoppers. In this case, not only would the number of TV viewers go up but their behavior might also change because of the nature of the event. From a theory perspective, the loss of life could result in grief or anger. While research associates grief with deactivation (Smith and Ellsworth 1985), which could lead to the avoidance of shopping decisions, anger has been argued to be associated with high activation levels (Berger and Milkman 2012, Feldman Barrett and Russell 1998, Smith and Ellsworth 1985). We expect that part of the population became angry while another part became sad, which could even out this potentially confounding effect. However, we cannot be 100% confident of this. To address this issue, we chose the following approach. We use only the manhunt for the suspect that ensued on April 19 as the IV (IV2 in equation (2)) and not the actual bombing (IV1 in equation (2)). While the manhunt was certainly thrilling and can be considered comparable to an attention-drawing crime thriller, we cannot directly associate it with the mood states that the first event (i.e., the Boston Marathon bombing) could have potentially generated. We still observe a negative effect of TV viewership on sales of high-complexity products (coef.=-.159, $p<.01$) and an insignificant positive effect on sales of low-complexity products (coef.=.934, $p \geq .1$). All IV-related statistics such as the weak IV test indicate that IV2 alone also constitutes a valid IV. This analysis confirms our results from the full models and makes mood states as a confounding effect rather unlikely. We can conclude that at that the least, high-complexity products suffer from viewers' focus on a TV program. Low-complexity products do not suffer from this attention effect and may even benefit from the second screen effect.

As a second way to approach the mentioned confounding effect, we can exploit the fact that we have two IVs for one endogenous variable. This allows us to test whether the instruments satisfy the orthogonality condition and whether they are appropriately uncorrelated with the disturbance process. Indeed, all tests indicate that our estimates deliver a valid causal interpretation.

4. Individual-level Data Analysis

While analysis at the aggregate level can yield valuable insights, we present in the following an analysis of individual-level behavior based on a unique dataset. We identify a sample of individuals who concurrently watch TV and engage in online shopping and compare their online shopping behavior to the same sample's online shopping behavior on a reference day.

4.1. Data and Identification Strategy

To supplement our findings, we investigate the effect of product complexity using data on individuals' online shopping activity during a television broadcast. We compare potential changes in online shopping behavior due to the influence of TV consumption compared to the typical online shopping behavior of these individuals. For individual shopping activity, we use data from the Bing Toolbar that provides anonymized browsing activity for users who have opted in to data collection. We look at their visits and purchases at Amazon.com, the leading e-commerce store in the United States. As in the previous study, we want to estimate the effect of watching television on buying products, and how this behavior changes with the complexity of products.

4.2. Inferring Product Visits and Purchases

The first task is to identify the products browsed by each individual using data on their page visits to Amazon.com. Fortunately, Amazon uses a small number of templates to generate web URLs for product pages. These templates have the form www.amazon.com/dp/B010OAQPJA/, where the code "dp", for instance, corresponds to a product webpage and the ten-character code following it provides an identifier for the product, known as an "ASIN ID", which uniquely identifies each product sold on the website. We use these templates to find page visits that corresponded to viewing a product on Amazon.

Estimating purchases is more complicated. Most purchases on Amazon are completed over a secure https web protocol and thus webpage URLs do not contain information about the products purchased. However, as a proxy for purchase, we use adding a product to a customer's cart, which is considered indicative of an actual purchase. That is, for each page visit that is generated after a user clicks on a product's "Add to Cart" button, we identify the product visited just before this page visit and label it as a purchased product. Note that this proxy does not work for all product categories; in particular, Amazon-specific products such as the Kindle or gift cards, or online media products such as Amazon Video do not employ standard add-to-cart buttons, and we thus exclude these categories from our analysis.

Combining these two methods, we obtain a timestamped dataset of products visited and purchased by each user. For each product in the dataset, we use the Amazon Product Advertising API to fetch additional information, such as the product name, manufacturer and category.

4.3. Television Viewing Activity

In principle, while we could examine the effect of TV viewership on Amazon activity for any TV show, we restrict our analysis to popular television broadcasts that include sports events and award shows to identify samples of TV viewers that are large enough for our purpose. Specifically, we analyze the 2017 and 2016 Super Bowls, 2017 and 2016 Oscars, and 2016 Grammy telecast. We do not consider the 2017 Grammys because its reference day (which we use for analysis below) was the day of another popular broadcast show. For each of these events, we collect broadcast dates and times.

To connect television viewing with shopping on Amazon, we also need to know, among the people who visited Amazon.com, who was also watching television during the event. Without access to individual-level TV viewership data, however, this is difficult to estimate. Therefore, we instead look for people who conveyed an interest in watching the event and thus would be more likely to watch the event on TV than others. We do so by using people's searches about the event on the Bing search engine a few days before the event. Specifically, we look for all users whose searches on Bing included phrases indicating their interest in watching a TV program, such as "when is superbowl". This provides a proxy for identifying individuals with a higher chance of watching the TV program compared to all other users. We restrict our analysis of purchases on Amazon to users from this set. In total, this leads to a dataset of thousands of products and and thousands of users over the five TV broadcast events.

4.4. Identifying the Causal Effects of TV Viewing

To obtain estimates of change in individuals' online shopping behavior due to television viewing, we compare activity on Amazon for each individual during the event to the same time a week before the event, which we call the "reference day". The identifying assumption is that everything else remains the same during identical hours a week before compared to the event day, except watching TV on the event day. Comparing the purchases made on the event day versus a week before provides us an estimate of the effect of watching TV.

As mentioned above, there are thousands of unique products that are visited both during the event and during the same hours on the reference day. This leads to a sparse distribution of visits and purchases per product. Therefore, we aggregate individual products to their product category and treat each category as a unit of analysis.

As with all observational studies, we need to guard against confounding factors. First, our strategy for including users in the analysis based on their Bing searches may potentially introduce a selection bias. The reason is that people who search for an event online before the event are also more likely than others to browse e-commerce websites, simply because they are active online during this time. This is known as activity bias, and it is commonly reported in studies of online behavior (Lewis et al. 2011). To counter this bias, we include only users who searched for the event a few days before its date, instead of on the day of the event. Such queries still indicate users' interest in the event but do not introduce bias towards users who have higher online activity on the event day. Still, there may be day-to-day differences in Amazon activity. To avoid bias due to these differences, we compare the relative fraction of purchases for each product category instead of the actual number of purchases on the event or reference day.

4.5. Descriptive Study

We first use complexity scores by Lovett et al. (2014) to analyze sales on Amazon. In general, however, Amazon categories such as Fashion, Televisions, or Computer Accessories, tend to be more specific than Lovett et al.'s classifications. Moreover, there are many other categories that do not directly correspond to their classifications. We thus combine categories from Amazon to match four relevant categories from Lovett et al. (2014). We analyze the Amazon categories Beauty, Consumables (Food and Beverages), Fashion (Clothing), and Electronics (Technology and Telecom). In Lovett et al. (2014), these categories are classified as they are listed in Table 3, with respect to their complexity. We rank Consumables and Beauty as low-complexity categories because their scores are among the lowest in the analyses by Lovett et al. (2014). The Fashion category has slightly higher than average complexity. Electronics is a good example of a high-complexity category. Lovett et al. (2014) list only Financial Products as more complex, but these are not (yet) sold by Amazon.

According to our findings in the previous section, we expect that sales in the low-complexity category should increase during TV consumption while they should decrease for the high-complexity category. We focus on the share or proportion of activity for the category of interest among all categories to address potential activity bias, mentioned above. Table 4 presents the share of the purchase basket for the different categories in our study, both on the event day and the reference day. We selected one event of each type, namely the Grammys, Oscars, and Super Bowl. The

	Complexity according to Lovett et al. (2014)	Our Classification
Beauty	1.55	Low Complexity Category
Consumables	1.55/1.62	Low Complexity Category
Average across All Categories	*1.87*	
Fashion (Clothing)	1.95	Medium Complexity Category
Electronics (Technology and Telecom)	2.04/2.14	High Complexity Category

Table 3 Classification of Amazon Categories by Complexity

observations with check marks are in line with our expectations while the cross indicates a result that is not in line with our expectations. The table illustrates, for example, that 16.7% of all sales initiated by people watching the 2016 Grammys came from the Beauty category. If we examine the purchases of the same people one week before the Grammys, this share is 0%. The shares of beauty products bought by people who watched the Oscars or the Super Bowl in 2017 are 4.3% and 6.5%, respectively, during the event while one week before this event, the shares of beauty products purchased by the same people were 1.8% and 1.6%, respectively. This increase during TV consumption could indicate that people are more likely to buy beauty products while consuming TV, which is in line with our hypothesis that TV consumption can increase sales of low-complexity goods. The same holds true for products from the Consumables category.

On the other side of the complexity spectrum, we analyze the Electronics category and observe that the share of this category decreases by -6.4, -1.3 and -4.6 percentage points when people watch TV (2016 Grammys, 2017 Oscars, and 2017 Super Bowl, respectively). In line with our theoretical arguments and the previous findings in Hinz et al. (2016), this observation indicates that people are less likely to buy high-complexity products while watching TV.

For the medium complexity category, we observe once that its share of purchases decreases and once that it increases. We observe an increase of fashion sales during the Super Bowl, which could be either driven by increased sales of sporting apparel or be a result driven by a special segment of consumers. We report a similar observation at the aggregate level in Section 3 and mention that the result could be driven by a subsample of the population that is not interested in the game and uses the time to make complex purchase decisions. This observation is not in line with our hypotheses and clearly shows that not only complexity but also other category features are of relevance and potentially influenced by the context of TV consumption.

Overall, this analysis of a highly unique dataset over thousands of products and thousands of consumers allows a meaningful descriptive analysis of individual behavior in the context of TV consumption for the first time. The results support our findings presented in Section 3. The subsequent section further develops our approach and suggest a new operationalization for complexity that can be calculated based on observable behavior and allows us to conduct a quantitative study on individual behavior.

	Grammys 2016				Oscars 2017				Super Bowl 2017			
	Before*	During*	Change**		Before*	During*	Change**		Before*	During*	Change**	
Beauty	0	16.7	+16.7	✓	1.8	4.3	+2.5	✓	1.6	6.5	+4.9	✓
Consumables	7.7	25.0	+17.3	✓	9.9	10.6	+0.7	✓	4.7	7.5	+2.8	✓
Fashion	0	0	0		27.9	21.3	-6.6	✓	10.9	20.4	+9.5	X
Electronics	23.1	16.7	-6.4	✓	4.5	3.2	-1.3	✓	5.7	1.1	-4.6	✓

Table 4 Share of Basket before and during Major TV Events (* in %, ** in Percentage Points)

4.6. Quantitative Study

4.6.1. Defining Complexity Since product categories on Amazon do not directly correspond to Lovett et al.'s classifications, we devise a novel strategy to estimate the complexity of products. Relating complexity to the attention required to purchase a product, we define complexity as the amount of browsing effort that people spend before purchasing a product. We quantify browsing effort for a product by the number of related search, category, and product pages viewed by an individual before they add a product to their cart. Complexity of a product category is defined by the percentage ratio of the number of products purchased and the total number of browsing visits leading to those purchases.

$$Complexity = 1 - \frac{Number\ of\ Add\text{-}to\text{-}carts}{Number\ of\ browsing\ visits} * 100 \tag{4}$$

We scale this complexity score by subtracting from it the minimum complexity score across all categories. With this definition, we use Amazon browsing data from an unrelated week in May 2017 to estimate complexity for all product categories and to avoid endogeneity between complexity scores and Amazon purchasing activity during the events.

Table 5 illustrates that this automatically computed complexity score is in line with Lovett et al.'s classification, which is based on survey data generated by numerous respondents. Thus, this alternative strategy to estimate the complexity of products may also be useful for other research questions.

4.6.2. Results We restrict our analysis to the 20 top-selling Amazon categories that contribute to more than 72% of all sales and again use the 2016 and 2017 Super Bowl, the 2016 and 2017 Oscars, and the 2016 Grammys as events. For each hour on the event day and the reference day, we compute the relative fraction of purchases for each category. We use these hourly observations for the fraction of category sales for the five events and thus arrive at a dataset that consists of 24(hours per day)*2 (reference and event day)*20(categories)*5(events)=4,800 observations.

We estimate fractional outcome regressions to account for the dependent variable–the fraction of sales for a category over a given hour–which can take on values between 0 and 1. We estimate a parsimonious Model (1) in Table 6 using the variables of interest, complexity of a category, and

Product Category	Complexity Score
Ink and Toner	0
Art and Craft Supplies	7.8
Groceries	9.2
Beauty	9.5
Health and Personal Care	9.5
Shirts	11.7
Shoes	12.7
Televisions	13.1
Notebook Computers	13.1

Table 5 Complexity Scores for a Sample of Product Categories on Amazon. Median complexity of a category is 10.6. Green cells show categories with complexity lower than the median, while red cells show categories with higher complexity than the median.

whether an hour corresponds to the time when a TV event was broadcast. We extend this model by adding control variables for time of the day and the particular day (Model (2) in Table 6).

Table 6 shows that during the event and thus during TV consumption, the fraction of sales increases significantly in our sample of categories (p<.05). This indicates that we observe more sales in the top categories during TV consumption and that people obviously purchase fewer obscure products from niche categories in this contextual situation. We also observe that categories of high complexity account for fewer sales than categories of lower complexity at this time (p<.05).

The effect of interest is captured by the interaction between the event (i.e., TV consumption) and complexity. The results show that at the time of the events, complexity significantly decreases sales (p<.05). Consumers are obviously not able to process complex information while watching TV and thus their online shopping focuses on categories of low complexity. Overall, this quantitative study–with an innovative data collection, a highly unique dataset, and a new operationalization of complexity–confirms the previous findings. We can triangulate the results that were generated at the aggregate level with this complementary analysis.

5. General Discussion

This paper is, to the best of our knowledge, the first that examines the impact of concurrent TV consumption on online shopping behavior. In two empirical studies, which use very different datasets and analysis strategies, we find that second screening consumes some part of the consumer's attention, who is then unable or unwilling to engage in complex decision making.

5.1. Research Contribution

This article thereby contributes to research in the area of the attention economy and delivers good evidence that people may not have sufficient stores of attention to effectively engage in several activities at the same time. In particular, we show that if TV consumes some of the available attention, consumers focus on buying low-complexity products.

	(1) Fractional Regression		(2) Fractional Regression with Controls	
Complexity	-6.138**	(2.871)	-6.160**	(2.854)
Event	1.687**	(0.787)	1.524**	(0.776)
Complexity*Event	-18.708**	(7.952)	-18.767**	(7.879)
Hour Controls	no		yes	
Day Controls	no		yes	
Constant	-3.032***	(0.272)	-2.903***	(0.391)
Wald Chi	16.59		359.71	
N	4,800		4,800	

Standard errors in parentheses, * p<.1, **p<.05, ***p<.01

Table 6 Impact of Complexity on Sales Fraction per Category

This study is also a very good example of the importance of context when it comes to economic decision making, an area that has so far been underresearched. Literature on the interplay between context and economic decision making is rather scarce and, at best, we find isolated and detached insights in scattered sources (e.g., Lussier and Olshavsky 1979, Klein and Yadav 1989).

The reason for this deficit can be attributed to two factors: First, in the last decades, context was hardly measurable and could not be easily integrated in economic analyses. A few exceptions are the influence of time or weather on sales (Spiekermann et al. 2011, Tian et al. 2016). Second, context is usually very difficult or nearly impossible to influence and thus does not constitute a manageable factor, so management disciplines naturally focus on factors that can be changed or altered (e.g., prices or budget allocation) by the focal manager. For this reason, context was, at best, a factor to be controlled for in empirical studies.

However, two recent technical developments have completely changed the picture. First, while computer science had already begun to understand the importance of context for system design in the 1990s, it was an academic question at that point. Stationary devices with only few sensors dominated the market at that time such that not only were the possibilities for measuring context limited, but it was also impossible to act on the basis of this scant information anyway (Abowd et al. 1999). However, it is currently possible to access much more information due to the plethora of various kinds of sensors that accompany and surround us. This information can help us to understand the influence of context on economic decision making. Second, we can now take actions based on the potential for new insights, which explains why context is becoming increasingly important for our discipline. With the improved possibilities of information and communications technology in general and data analysis in particular, managerial decisions can be much more fine-grained and personalized than in the past.

The development of context-aware systems, new data opportunities, and, finally, the results of this paper show that context is important and that research should pay more attention to its influence on economic decision making.

5.2. Managerial Implications

The promise of personalized advertisements and recommendations has always been that they can deliver on the idea of a customer-oriented marketing strategy and provide customers with more pertinent information with less effort (Montgomery and Smith 2009). Besides information about individuals and their preferences, contextual information may prove valuable for advertising and recommending items in particular circumstances (Adomavicius and Tuzhilin 2015). Such contextual information can consist of data on time, location, and/or an individual's state of mind or current activities. It can be used, for example, by a travel site recommender to suggest vacation packages for summer or winter. Context-aware recommenders sometimes ask individuals for input about context. For example, some music recommenders ask users to specify their current mood, which is then incorporated in the recommendation process. Netflix also uses contextual information when making recommendations. According to their CEO Reed Hasting, the performance of its recommendation algorithms increases by up to 3% when taking into account simple contextual information such as the time of the day or an individual's location (Adomavicius and Tuzhilin 2015).

Search advertising (SEA) is another example of the consideration of context in advertising and is certainly one of the most successful forms of online marketing. Ads are sold and delivered based on keywords, which not only can be used to identify individual preferences but also implicitly carry contextual information. SEA is so attractive because the keywords allow inferring an individual's stage in the conversion funnel.

While the data sources used in this paper are currently rather unique, further technological improvements will allow tracking individuals across different devices and capturing this important contextual information in the near future on a broader basis. An exemplary solution comes from Simulmedia, a television advertising company that provides a digital advertisement targeting platform for linear television. Its audience network reaches 110 million households across the US through partnerships with 84 national cable networks and top multi-channel video programming distributors (Simulmedia Inc. 2015). In addition, it recently declared a partnership with social media giant Facebook (Ellwanger 2016). Cross-media analyses are one of the company's priorities. Other examples of advances in the area of addressable TV are Invidi Technologies, a Google-funded company that was acquired by ATT Inc., Dish Network Corp., and WPP PLC in November 2016 (Armental 2016), and Television and Cross Media Services by comScore, which aims to help clients understand cross-channel and cross-platform consumption (Weisler 2015). As mentioned in the introduction, Alphonso is another player in this relatively new market. This firm initially created software to capitalize on ads through second-screen viewing, as people increasingly turned their attention to smartphones and tablets during TV breaks, but has now broadened its focus (Maheshwari 2017).

Our results indicate why access to contextual information may be valuable. Our insights about the influence of concurrent TV consumption on shopping behavior can be used to optimize personalization by incorporating contextual information. Concretely, our results indicate that recommendations or advertisements on websites should vary when individuals are second screening. This context, however, does not impact all types of products equally. In particular, firms should not advertise high-complexity products such as a laptop to consumers who concurrently watch TV; instead, they should recommend low-complexity products such as household goods and beauty products.

5.3. Limitations and Directions for Future Research

Our study is subject to some limitations that provide avenues for future research. First, we did not determine the influence of certain show types on online shopping behavior. Potentially, genres such as drama may affect online shopping behavior in a different way than comedy or sports shows. Second, while our study did not consider activity on social media while second screening, we expect a nuanced interplay between TV watching, social media activity, and online shopping behavior. Future research should examine this interplay in detail. Third, cultural differences may exist, and a repetition of the study in a different cultural context could yield interesting new insights. Finally, the increasingly prevalent practice of integrating software in apps that can detect sounds, even when a phone is not being actively used and is in a pocket, for example, is questionable from a privacy perspective. Nevertheless, personalized ads or recommendations may be beneficial for consumers if irrelevant pieces of information are filtered out and the remaining information is useful and not annoying. Therefore, there should be a vivid discussion on the opportunities and challenges that accompany technological advancements. Our research indicates that newly available data are valuable, and researchers, industry, consumers and policy makers need to find ways to leverage these opportunities in a fair manner.

References

Accenture. 2015. Digital video and the connected consumer. https://www.accenture.com/us-en/~/media/Accenture/Conversion-Assets/Microsites/Documents17/Accenture-Digital-Video-Connected-Consumer.pdf.

Adomavicius, G., A. Tuzhilin. 2015. Context aware recommender systems. Francesco Ricci, Lior Rokach, Bracha Shapira, Paul B Kantor, eds., *Recommender Systems Handbook*. Springer, 266–290.

Armental, Maria. 2016. At&t, dish, wpp to buy invidi technologies. https://www.wsj.com/articles/at-t-dish-wpp-to-buy-invidi-technologies-1479766965.

Babin, Barry J, William R Darden, Mitch Griffin. 1994. Work and/or fun: measuring hedonic and utilitarian shopping value. *Journal of consumer research* **20**(4) 644–656.

Baum, Schaffer, C. F., Stillman S. M. E. 2010. ivreg2: Stata module for extended instrumental variables/2sls, gmm and ac/hac, liml and k-class regression. ttp://ideas.repec.org/c/boc/bocode/s425401.html.

Belk, Russell W. 1975. Situational variables and consumer behavior. *Journal of Consumer research* **2**(3) 157–164.

Berger, Jonah, Katherine L Milkman. 2012. What makes online content viral? *Journal of marketing research* **49**(2) 192–205.

Davenport, Thomas H, John C Beck. 2001. *The attention economy: Understanding the new currency of business*. Harvard Business Press.

Dhar, Ravi, Klaus Wertenbroch. 2000. Consumer choice between hedonic and utilitarian goods. *Journal of marketing research* **37**(1) 60–71.

Dimmick, John, Yan Chen, Zhan Li. 2004. Competition between the internet and traditional news media: The gratification-opportunities niche dimension. *The Journal of Media Economics* **17**(1) 19–33.

Ellwanger, Steve. 2016. Simulmedia, facebook to advance reach and frequency coordination across tv and social. https://www.beet.tv/2016/04/morgan-simulmedia2.html.

Falkinger, Josef. 2007. Attention economies. *Journal of Economic Theory* **133**(1) 266–294.

Feldman Barrett, Lisa, James A Russell. 1998. Independence and bipolarity in the structure of current affect. *Journal of personality and social psychology* **74**(4) 967.

Goldhaber, Michael H. 1997. The attention economy and the net. *First Monday* **2**(4).

Grether, WF. 1973. Human performance at elevated environmental temperatures. *Aerospace Medicine* **44**(7) 747–755.

Hall, Alastair R, Glenn D Rudebusch, David W Wilcox. 1996. Judging instrument relevance in instrumental variables estimation. *International Economic Review* 283–298.

Hancock, PA, Ioannis Vasmatzidis. 2003. Effects of heat stress on cognitive performance: the current state of knowledge. *International Journal of Hyperthermia* **19**(3) 355–372.

Hinz, Oliver, Shawndra Hill, Ju-Young Kim. 2016. Tv's dirty little secret: The negative effect of popular tv on online auction sales. *MIS Quarterly* **40**(3) 623–644.

Kleibergen, Frank, Richard Paap. 2006. Generalized reduced rank tests using the singular value decomposition. *Journal of econometrics* **133**(1) 97–126.

Klein, Noreen M, Manjit S Yadav. 1989. Context effects on effort and accuracy in choice: An enquiry into adaptive decision making. *Journal of Consumer Research* **15**(4) 411–421.

Kroeber-Riel, Werner. 1987. Informationsüberlastung durch massenmedien und werbung in deutschland. *Die Betriebswirtschaft* **47**(3) 257–264.

Kushwaha, Tarun, Venkatesh Shankar. 2013. Are multichannel customers really more valuable? the moderating role of product category characteristics. American Marketing Association.

Lewis, Randall A, Justin M Rao, David H Reiley. 2011. Here, there, and everywhere: correlated online behaviors can lead to overestimates of the effects of advertising. *Proceedings of the 20th international conference on World wide web*. ACM, 157–166.

Liu, Yong. 2006. Word of mouth for movies: Its dynamics and impact on box office revenue. *Journal of marketing* **70**(3) 74–89.

Lodish, Leonard M, Magid Abraham, Stuart Kalmenson, Jeanne Livelsberger, Beth Lubetkin, Bruce Richardson, Mary Ellen Stevens. 1995. How tv advertising works: A meta-analysis of 389 real world split cable tv advertising experiments. *Journal of Marketing Research* 125–139.

Lovett, Mitchell, Renana Peres, Ron Shachar. 2014. A data set of brands and their characteristics. *Marketing Science* **33**(4) 609–617.

Lussier, Denis A, Richard W Olshavsky. 1979. Task complexity and contingent processing in brand choice. *Journal of Consumer Research* **6**(2) 154–165.

Maheshwari, Sapna. 2017. That game on your phone may be tracking what you're watching on tv. https://www.nytimes.com/2017/12/28/business/media/alphonso-app-tracking.html.

Moe, Wendy W. 2006. An empirical two-stage choice model with varying decision rules applied to internet clickstream data. *Journal of Marketing Research* **43**(4) 680–692.

Montgomery, Alan L, Michael D Smith. 2009. Prospects for personalization on the internet. *Journal of Interactive Marketing* **23**(2) 130–137.

Shapiro, Carl, Hal R Varian. 1998. *Information rules: a strategic guide to the network economy*. Harvard Business Press.

Simon, Herbert A. 1996. *The sciences of the artificial*. MIT press.

Simulmedia Inc. 2015. Simulmedia to guarantee better business outcomes than traditional tv buys using audience-targeted tv. http://www.simulmedia.com/news/2015/03/05/better-business-outcomes-using-audience-targeted-tv.

Smith, Craig A, Phoebe C Ellsworth. 1985. Patterns of cognitive appraisal in emotion. *Journal of personality and social psychology* **48**(4) 813.

Stock, J.H., M. Yogo. 2005. Testing for weak instruments in linear iv regression. D.W.K. Andrews, J.H. Stock, eds., *Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg*. Cambridge University Press, 80–108.

Weisler, Charlene. 2015. What are data companies contributing to the tv market? – part 1. https://www.mediavillage.com/article/what-are-data-companies-contributing-to-the-tv-market-part-1/.

Appendix. First-stage estimates

	(2)		(3)		(4)	
	IV results, Only IV		IV results, (2) & Time Controls		IV results, (3) & Weather Controls	
IV1 (Explosion)	3.464***	(0.960)	1.458***	(0.295)	1.442***	(0.296)
IV2 (Manhunt)	3.443***	(0.865)	1.586***	(0.182)	1.677***	(0.183)
01:00-02:59 (0/1)	-5.294***	(0.045)	-5.284***	(0.045)		
03:00-04:59 (0/1)	-6.575***	(0.045)	-6.564***	(0.044)		
05:00-06:59 (0/1)			-6.243***	(0.050)	-6.233***	(0.049)
07:00-08:59 (0/1)			-4.340***	(0.051)	-4.333***	(0.050)
09:00-11:59 (0/1)			-4.101***	(0.048)	-4.093***	(0.047)
11:00-12:59 (0/1)			-4.091***	(0.050)	-4.083***	(0.049)
13:00-14:59 (0/1)			-4.019***	(0.054)	-4.011***	(0.053)
15:00-16:59 (0/1)			-3.181***	(0.052)	-3.174***	(0.051)
17:00-18:59 (0/1)			-1.733***	(0.054)	-1.733***	(0.053)
21:00-22:59 (0/1)			0.465***	(0.057)	0.465***	(0.056)
23:00-00:59 (0/1)			-2.674***	(0.051)	-2.667***	(0.051)
Sunday (0/1)			0.628***	(0.037)	0.628***	(0.036)
Monday (0/1)			-0.157***	(0.032)	-0.137***	(0.031)
Tuesday (0/1)			-0.302***	(0.031)	-0.278***	(0.031)
Wednesday (0/1)			-0.390***	(0.030)	-0.370***	(0.029)
Thursday (0/1)			-0.343***	(0.030)	-0.326***	(0.030)
Friday (0/1)			-0.306***	(0.031)	-0.296***	(0.030)
January (0/1)			-0.252***	(0.038)	-0.312***	(0.039)
February (0/1)			-0.319***	(0.040)	-0.434***	(0.036)
March (0/1)			-0.469***	(0.036)	-0.397***	(0.035)
April (0/1)			-0.479***	(0.039)	-0.307***	(0.041)
May (0/1)			-0.460***	(0.041)	-0.209***	(0.047)
June (0/1)			-0.628***	(0.036)	-0.337***	(0.050)
July (0/1)			-0.224***	(0.044)	0.084 (0.066)	
August (0/1)			-0.163***	(0.037)	0.143***	(0.055)
September (0/1)			0.309***	(0.037)	0.587***	(0.047)
October (0/1)			0.512***	(0.046)	0.730***	(0.052)
November (0/1)			0.351***	(0.041)	0.418***	(0.042)
Public Holiday (0/1)			0.183***	(0.066)	0.167** (0.066)	
Super Bowl (0/1)					2.793***	(0.557)
Blizzard (0/1)					1.816***	(0.216)
Temperature deg. C					-0.031***	(0.003)
Temperature deg. C					0.001***	(0.000)
Constant	5.671***	(0.025)	9.425***	(0.054)	9.523***	(0.055)
F	14.41		2433.43		2196.33	
R^2	0.002		0.884		0.888	
RMSE	2.309		0.787		0.773	

Table 7 First-stage Estimates with Number of TV Viewers as Dependent Variable

Note: * $p < .1$, ** $p < .05$, *** $p < .01$, two-tailed significance levels, Estim.: Estimates, SE: Standard Errors, RMSE: Root Mean Squared Error.