INSTITUTE OF RETAIL ECONOMICS

DISTRACTION AND CONSUMER BEHAVIOR: EVIDENCE FROM A NATURAL FIELD EXPERIMENT

SVEN-OLOV DAUNFELDT, HAMPUS POPPIUS & NIKLAS RUDHOLM

HFI WORKING PAPER No7

Distraction and Consumer Behavior: Evidence From a Natural Field Experiment*

Daunfeldt, Sven-Olov†
Poppius, Hampus‡
Rudholm, Niklas§

December 19, 2019

Abstract

This paper explores how distraction from a consumer's surroundings may influence consumption. In a natural field experiment involving 16 fast-food restaurants over five months, we randomly varied the degree of familiarity of the background music. We find that playing familiar music reduces revenues and quantity sold by more than 4 % relative to playing similar but unfamiliar music. We conduct a complementary survey that suggests that the reason that familiar music reduces consumption is that it distracts consumers. We conclude that when consumers become distracted, they consider fewer consumption opportunities and therefore consume less. The results have implications for the literature on attention and framing as well as for marketing policy.

JEL-codes: D91, D89, M31, L83.

*We thank Hakan J. Holm and Alessandro Martinello for very insightful comments. We also thank seminar participants at Lund University; Arne Ryde Workshop on Experimental Methods in Study of Firms, Management, and Entrepreneurs; and the 2017 SEA Conference. We are grateful to Soundtrack Your Brand, especially Jasmine Moradi, for making this study possible and for genuine support during the process.
†Institute of Retail Economics, Stockholm; and Department of Economics, Dalarna University.
‡Department of Economics, Lund University; and Institute of Retail Economics, Stockholm. E-mail: hampus.poppius@nek.lu.se.
§Institute of Retail Economics, Stockholm.

1 Introduction

Traditional economic theory regards choice as a result of stable and well-defined preferences. However, a substantial body of literature recognizes that preferences are easily manipulated by the framing of the choice (Thaler, 1985; Kahneman *et al.*, 1990; Tversky & Thaler, 1990). Therefore, the policy designer should not solely focus on what alternatives to present but also on how they are presented. Understanding the full scope of policy design offers the policy designer a wider range of tools and opportunities to affect behavior without limiting freedom of choice.

One opportunity is to influence the decision maker's surrounding environment. Psychological theory predicts that the surroundings of the decision maker stimulate her emotions and thereby influence her behavior (Mehrabian & Russell, 1974; Donovan & Rossiter, 1982; Donovan *et al.*, 1994). For example, Donovan *et al.* (1994) show that consumers' emotions during a store visit correlate with purchasing behavior.

We provide causal estimates from a large-scale randomized field experiment of the effect of the decision environment on consumer choices. Specifically, we randomly varied the familiarity of the background music in 16 fast-food restaurants over 20 weeks. We find that familiar music decreases consumption by approximately 4 percent compared to novel music. Hence, we not only show that the surrounding environment influences consumers but also that a very subtle manipulation of the surrounding environment can substantially alter consumer behavior.

Based on the results from the experiment, we formulate a hypothesis about attention as the mechanism at work. Supplementing the estimated effect on sales, we use survey responses from over 2,000 customers to show that familiar music attracts more attention than novel music does. However, customers purchase more when the music is novel. Therefore, we propose that the negative effect of familiar music on consumption arises because familiar music distracts the customer. In other words, more novel music allows the customer to allocate more attention to available consumption possibilities. This explanation is not only useful for the specific situation we study but also has broader implications spanning several studies that relate to information processing and attention in decision making.

First, we contribute to the literature that recognizes the importance of attention in decision making and the fact that individuals often face more information than they can process (Simon, 1955; Sims, 2003; Tversky & Kahneman, 1975). Hefti & Heinke (2015) define the set of alternatives that a decision maker can consider as the attention set, which is limited by a cognitive restraint. They stress the importance of stimulus-driven attention allocation, where the relative salience of competing information sources determines the allocation of attention (Nothdurft, 2000). Instead of goal-driven attention allocation, rather than considering the most relevant information, the decision maker allocates attention to salient but irrelevant information (Jonaityte, 2016). We contribute to this literature by providing experimental evidence of the importance of irrelevant information cues.

Second, we contribute to the literature on focusing (Kőszegi & Szeidl, 2012). This literature studies how much the decision maker emphasizes a specific attribute of considered alternatives depending on the attribute's relative salience. However, it is unclear what determines which alternatives enter the set of considered alternatives. Similarly, Gabaix (2014) models consumer choice when attention is limited and allocated across attributes of the considered alternatives. The implications are important for many classical results. However, the model does not explain what external factors determine the limits of attention. Our contribution to this literature is to empirically demonstrate the importance of the total amount of attention allocated to the choice. Furthermore, we argue that the total amount of attention affects the size of the consideration set and thereby affects consumption.

Third, we contribute to the literature on background music and consumption. This literature includes a wide variety of research questions. For example, studies have investigated the choice of tempo (Milliman, 1982, 1986; Oakes, 2003), different music styles (Areni & Kim, 1993; North *et al.*, 1999; Wilson, 2003), and interactions with other sensory cues (Mattila & Wirtz, 2001). However, many aspects of music are difficult to evaluate because it is difficult to quantify an aspect of music when there is no objective measurement. We overcome this challenge by collaborating with Soundtrack Your Brand (henceforth SYB), the provider of Spotify Business. This means that we have data on the number of streams for each song and can therefore objectively determine the level of familiarity with the playlists used in the experiment.

Furthermore, most experimental studies on background music and consumption typically use only one store or restaurant, and the duration of the experiments tend to be limited to a few weeks at most.[1] Because of the lack of a control group, observing only one unit while varying music over time does not provide causal evidence. Without a control group, estimates may depend on time trends, and it is thus not possible to make causal inference. A short experimental duration is also problematic because it increases the risk of confounding factors coinciding with music variation and decreases statistical power. In contrast to previous studies, the collaboration with SYB makes it possible for us to conduct a large-scale randomized field experiment and thereby contribute to the literature by providing causal and robustly estimated effects of background music on consumer behavior.

The remainder of this paper is structured as follows. Section 2 describes the experimental setup, the data and how we estimate the effect of music. Section 3 presents our results. Finally, Section 4 discusses and concludes the paper.

[1] Garlin & Owen (2006) provides an overview.

2 The Experiment

2.1 Experimental Design

We conducted the experiment and data collection over the course of 20 weeks in the spring of 2016 (January through May). The experiment included 16 fast-food restaurants in the Stockholm metropolitan area belonging to the same restaurant chain. We randomly varied four different music treatments to measure how they affect sales. One of the treatments is no music, only silence. The three music treatments were constructed by experts at SYB. They followed two binary selection rules when designing the playlists used in the experiment. These are the same rules that they usually follow when designing playlists for their clients.

The first rule selects songs with a sound that suits the restaurant chain's brand values. This rule is based on findings in the business literature that sending congruent signals to customers enhances customer satisfaction and willingness to consume (Beverland *et al.*, 2006). To decide which songs are congruent with brand values, the restaurant chain provides SYB with value words that it wishes to be associated with, and these words are then used to classify songs as "brand-fit" or not. In our case, the restaurant chain wants to signal that they are "welcoming", "easy-going", "expressive", "youthful", and "humane".

Selecting songs that reflect brand values did not affect sales in the experiment, and we choose not to focus on this selection rule for two distinct reasons. First, the way in which songs are selected is highly subjective. Even if the music experts are highly skilled, it is not certain that their interpretation of the sound of songs is the same as that of the average customer. Second, it is impossible to disentangle the effect of congruence with the mere effect that this music would have in any other restaurant. These words are rather general and could induce the same effect among customers elsewhere. Therefore, while it is possible to estimate the difference in sales between a brand-fit playlist and another, it is not clear that the difference is due to the congruence per se.

The second rule is based on the familiarity of the songs included in the playlist. To determine a song's familiarity, the music experts use data on user behavior on the streaming service. Each time a user listens to a specific song, it counts as one stream of that song. Therefore, by selecting songs based on each song's number of streams, the music experts can determine the overall familiarity of a specific playlist.

We use three different music treatments, so-called playlists, in the experiment. The main difference among the playlists is the familiarity of the songs. The fourth treatment is silence, i.e., no music. The silent treatment tests whether the mere presence of music increases sales. All three playlists consist of 336 or 337 songs. We thus have large enough playlists to avoid considering the characteristics of specific songs and can focus instead on the general selection rules. All playlists exclude songs with explicit content.

Two playlists consist only of familiar songs. The first playlist (Familiar A) contains the songs with the most streams based on Spotify's Top 1000 Sweden playlist. The Familiar A playlist thus has a very high degree of familiarity. The second playlist (Familiar B) contains the most familiar songs from Spotify's Top 1000 Sweden playlist but excludes songs that are not in congruence with the brand values of the company. Most familiar songs satisfy the brand value alignment criteria, and 65 percent of the songs in the Familiar A playlist also appear in the Familiar B playlist. The remaining songs in these playlists thus differ in the first selection rule -– alignment of sound and brand values. Whereas Familiar B is completely aligned with brand values, only 65 percent of Familiar A is so aligned. The Familiar A playlist has the highest possible degree of familiarity, while Familiar B includes somewhat fewer familiar songs. However, the difference in familiarity is small since both playlists are based on Spotify's Top 1000 Sweden playlist.

The third playlist (Novel) includes songs that are much less familiar. Specifically, 74 percent of the songs are not among the top 1000 songs according to the number of streams. The Novel playlist thus differs from both Familiar A and Familiar B in terms of familiarity, and the difference is largest between Novel and Familiar A. However, Novel and Familiar A also differ to some extent in sound alignment. While 65 percent of the songs in Familiar A satisfy the sound alignment criteria, all songs in the Novel playlist satisfy do so.

To test whether sound alignment affects sales, we compare sales between Familiar A and Familiar B. Because we do not find any significant difference in sales between Familiar A and Familiar B, we conclude that sound alignment is not important for sales. Therefore, we argue that any observed difference in sales between Familiar A and Novel is due to the difference in familiarity and not sound alignment. We thus discard sound alignment as a relevant difference between Familiar A and Novel because our results show no effect of sound alignment on sales.

In summary, there are four treatments. Silence means that there is no music played in the restaurant. Familiar A includes the most popular songs and therefore has the highest degree of familiarity. Familiar B is also highly familiar, but SYB replaced 35 percent of the songs in the Familiar A playlist with somewhat less familiar songs. Therefore, while the content is still overall highly familiar to the average customer, it is not as familiar as Familiar A. Novel mainly includes songs that are not familiar, as they have not been played a great deal by the streaming service's users.

We have 16 restaurants in total. Eight restaurants are subject to treatment, and eight restaurants constitute the control group. The control restaurants receive the Novel treatment in all time periods. Because only eight restaurants are available for treatment assignment, randomly assigning one treatment to each restaurant and then comparing mean outcomes would mainly capture fundamental differences between the restaurants. Therefore, we randomly assign the restaurants four different treatment schedules. Each treatment schedule is a specific order of the four treatments. Because there are eight restaurants and four treatment schedules, there are two restaurants per treatment schedule. The treatment schedules are presented in Table 1.

We collect data on sales four weeks before and four weeks after the treatment period. All restaurants play the Novel playlist in the first and last four weeks. They played Novel because that is the playlist they played before the experiment. During the treatment period, the treatments vary within each restaurant. Because the treatment schedules start with different music treatments, there is variation in treatment across restaurants throughout the treatment period. We can thus control for both restaurant and time fixed effects. Because we expose the treatment restaurants to all four treatments but during different periods, the treatment restaurants act as controls for one another. In addition, collecting data from eight control restaurants provides pure control restaurants. The additional data from before and after the treatment period increases the number of control observations for both control and treated restaurants, which increases statistical power and precision.

Table 1: Treatment Schedules

	Weeks					
	Pre	Treatment Period				Post
Restaurants	1-4	5-7	8-10	11-13	14-16	17-20
1, 2	Novel	-	Fam A	Fam B	Novel	Novel
3, 4	Novel	Fam A	Fam B	Novel	-	Novel
5, 6	Novel	Fam B	Novel	-	Fam A	Novel
7, 8	Novel	Novel	-	Fam A	Fam B	Novel
Control	Novel	Novel	Novel	Novel	Novel	Novel

The music systems in the restaurants are controlled from the headquarters of SYB, and the staff in the restaurants cannot influence the music played during the experiment and control periods. Therefore, we are unaware of any risk of any non-compliance.

2.2 Sales outcomes

We use four different sales outcomes in our analysis: average purchase size, total revenues, total quantity, and number of purchases. By combining these four sales outcomes, we can provide a detailed account of the treatment effect on consumers' purchasing behavior.

The average purchase size is a common figure when analyzing sales. However, some customers return to the counter for an additional purchase after the main purchase. The total number of purchases is thus the sum of main purchases and additional purchases:

$$N = N^M + N^A$$

Because an additional purchase, x^A, generally involves smaller items, e.g., coffee and desserts, additional purchases are on average smaller than the main purchases, x^M. Therefore, the number of additional purchases decreases the average purchase size, $\bar{x}$:

$$x^A < x^M \iff \frac{\partial \bar{x}}{\partial N^A} < 0$$

On the other hand, if the number of additional purchases does not change but customers increase expenditures on the main purchase, average purchase size increases:

$$\frac{\partial \bar{x}}{\partial x^M} > 0$$

We assume that treatment does not affect the number of main purchases or the average size of the additional purchases. It is unlikely that the number of main purchases would change with the music treatments. That would require the music to be so disturbing that a customer leaves before placing her order. Relaxing the assumption that treatment does not affect the size of additional purchases does not change the conclusion if additional purchases do not grow by more than the main purchases, on average. If the music also does not change the sizes of additional purchases, treatment T can only affect average purchase size through the number of additional purchases and the average size of main purchases:

$$\frac{\mathrm{d}}{\mathrm{d}T}\bar{x} = \frac{\partial \bar{x}}{\partial N^A}\frac{\mathrm{d}N^A}{\mathrm{d}T} + \frac{\partial \bar{x}}{\partial x^M}\frac{\mathrm{d}x^M}{\mathrm{d}T}$$

The important assumption is that treatment does not affect the number of main purchases. Equivalently, we assume that the number of customers entering the restaurant for a meal is independent of our manipulation of background music. Therefore, if treatment increases (decreases) both the number of purchases and the average purchase size, treatment also increases (decreases) the average main purchase size. Furthermore, if treatment increases not only the average purchase size but also the number of additional purchases, the average main purchase size must also change. That is, the treatment must also increase the average main purchase size to the extent that it compensates for the decrease in average purchase size caused by the increase in the number of additional purchases:

$$\underset{(+)}{\frac{\partial \bar{x}}{\partial x^M}}\underset{(?)}{\frac{\mathrm{d}x^M}{\mathrm{d}T}} > -\underset{(-)}{\frac{\partial \bar{x}}{\partial N^A}}\underset{(?)}{\frac{\mathrm{d}N^A}{\mathrm{d}T}} \qquad \iff \qquad \frac{\mathrm{d}\bar{x}}{\mathrm{d}T} > 0$$

$$\implies \frac{\mathrm{d}\bar{x}}{\mathrm{d}T} > 0 \quad \text{and} \quad \frac{\mathrm{d}N^A}{\mathrm{d}T} > 0 \qquad \implies \qquad \frac{\mathrm{d}x^M}{\mathrm{d}T} > 0$$

Because we cannot separate the main purchases from additional purchases, we cannot directly estimate the effect on the average main purchase size. However, due to the arguments above, we can use the estimated effect on average purchase size with the effect on the number of additional purchases to infer how treatment affects the size of main purchases. As stated above, the number of customers is independent of treatment due to random assignment. Therefore, an effect of treatment on the number of purchases is an effect on the number of additional purchases:

$$\frac{\mathrm{d}N}{\mathrm{d}T} = \frac{\mathrm{d}N^A}{\mathrm{d}T}$$

Because a treatment effect on the number of purchases is an effect only on additional purchases, if treatment increases (decreases) both the number of purchases and the average purchase size, the treatment must also increase (decrease) the average main purchase size:

$$\frac{\mathrm{d}\bar{x}}{\mathrm{d}T} > 0 \quad \text{and} \quad \frac{\mathrm{d}N}{\mathrm{d}T} > 0 \qquad \Longrightarrow \qquad \frac{\mathrm{d}x^M}{\mathrm{d}T} > 0$$

$$\frac{\mathrm{d}\bar{x}}{\mathrm{d}T} < 0 \quad \text{and} \quad \frac{\mathrm{d}N}{\mathrm{d}T} < 0 \qquad \Longrightarrow \qquad \frac{\mathrm{d}x^M}{\mathrm{d}T} < 0$$

This result is important for the interpretation of our estimates. If treatment increases (decreases) both the number of purchases and average purchase size, we can conclude that it also increases (decreases) average main purchase size.

Total revenue is the sum of all purchases. Hence, it can increase (decrease) due to an increase (decrease) in the average purchase size, the number of purchases, or in both. Total revenues identify total changes in expenditure, while total quantity (number of units sold) identifies total changes in consumed units. The combination of these two measures reveals whether consumers substitute quantity for quality (e.g., price) or whether effects on revenues are solely driven by changes in quantity consumed. If quantity increases while revenues do not change, customers buy a larger quantity of less costly items. On the other hand, if revenues increase while quantity does not, customers substitute for costlier items. If revenues and quantity increase proportionally, the effect is solely driven by an increase in the number of consumed units. Hence, the relative sizes of treatment effects on revenues and quantity provide a deeper understanding of how purchases change with treatment.

2.3 Estimation strategy

During the experiment, two restaurants refused the Silence treatment. Because Novel was played before the experiment, these two restaurants received this treatment instead of Silence. In the presence of imperfect compliance, the conventional approach to estimate the treatment effect is to use the initial random assignment as an instrumental variable (IV) for the actual receipt of treatment. The IV estimates the local average treatment effect (LATE), which is the effect of treatment on the compliers. The higher the compliance rate, the more similar the LATE is to the intent to treat (ITT) estimate, which is the effect of the initial random assignment under perfect compliance. Because of imperfect compliance with the Silence treatment, the choice of treatment estimator matters for the estimation of the effect of Silence. However, the main purpose of this paper is to compare Novel to Familiar A. Because of the large number of observations of Novel, the compliance rate is almost exactly 1. Therefore, the ITT and LATE are almost identical when comparing Novel to Familiar A and B. Because the focus of this paper is to compare Novel to Familiar A, we only present ITT estimates.

We estimate the effect of the music treatments (ITT) with the following linear specification:

$$ln(S_{it}) = \alpha + \beta_1 FamA_{it} + \beta_2 FamB_{it} + \beta_3 SILENCE_{it} + \gamma R_i + \lambda WEEK_t + e_{it} \tag{1}$$

where $ln(S_{it})$ is the natural logarithm of a sales outcome S in restaurant i at time t, $FamA$ is an indicator variable equal to one for the Familiar A treatment, $FamB$ is an indicator variable equal to one for the Familiar B treatment, and $SILENCE$ is an indicator variable equal to one for the Silence treatment, in all cases zero otherwise. Because Novel is the baseline treatment, the interpretation of the β coefficients is the percentage difference in sales between the respective music treatment and the Novel playlist.[2] Hence, the β values measure the effect of the treatments in relation to Novel. R is a fixed effect for each restaurant, and $WEEK$ is a fixed effect for each week.

Because we use restaurant fixed effects, the ITT estimates are difference-in-difference (DD) estimates. The internal validity of DD estimates relies on the common-trends assumption, which means that in the absence of treatment, the outcome variable trends similarly across treatment and control groups. For corporate secrecy reasons, we are not permitted to disclose any descriptive statistics of the sales data. Therefore, we show the weekly sales trends in the pre-treatment period without value labels in Figure 1. Because the trends are parallel, the DD estimates provide causal ITT estimates.

We observe 16 restaurants for 20 weeks, which should result in 2,240 daily observations. However, due to construction work, some restaurants were closed for a total of eight days. Therefore, the total number of daily observations is 2,232. All missing days are in the post-treatment period.

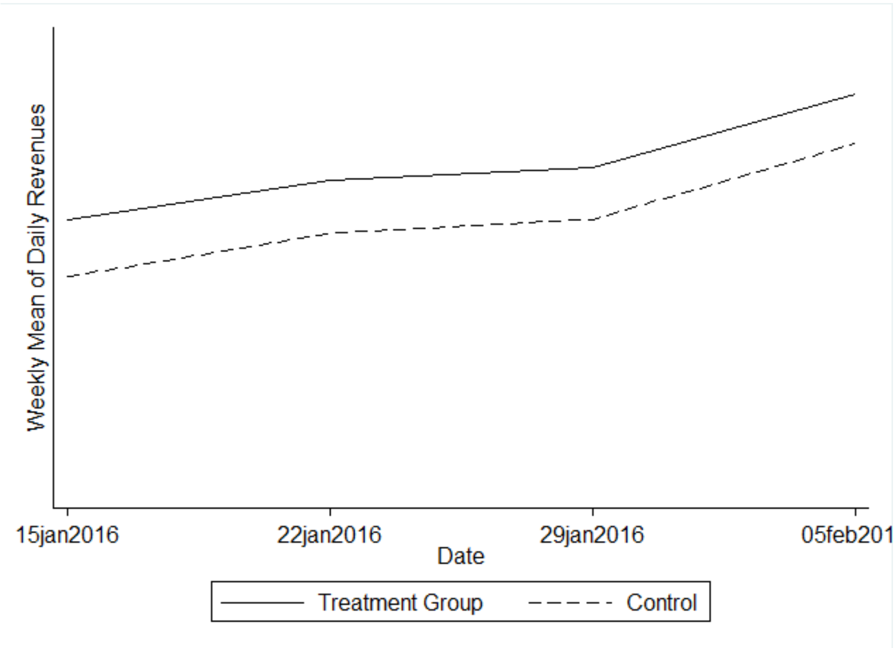
We combine the sales data with survey responses from 2,101 customers. The surveys were collected simultaneously in all eight treatment restaurants over six days in two different treatment periods. All respondents were approached outside the restaurant after their visit without any mention of the purpose of the study.

The key survey questions for this paper are "I noticed the music played in the restaurant" (notice), "I liked the music" (like), and "I recognized the

[2]The exact effect in percentage terms of a parameter estimate β can be calculated using the formula $100 \times [exp(\beta)\breve{\ }1]$. However, since the parameter estimates in our setting are small, the differences are negligible.

music" (recognize), respectively. The notice question indicates whether the music attracted the respondent's attention, while the like question measures how much they liked the music. The recognize question assesses the difference in familiarity between the playlists. The notice question is a binary YES or NO response, while the like and recognize questions ask the respondent to indicate agreement on a Likert scale from one to seven. If the answer to the notice question is NO, the respondent does not answer the like and recognize questions.

Figure 1: Revenues - Pre-experiment Period



Note: We are not allowed to present sales values.

We propose that customers returning to the counter for an additional purchase drives the effect of music on the number of purchases. We test this mechanism by asking respondents whether they made an additional purchase.

Because the responses to the survey questions are binary or on a scale from one to seven, proper inference requires non-linear estimators. However, OLS yields very similar results to logit regressions, and for simplicity, we therefore present the OLS estimates. Because the surveys are not collected in all time periods, the inclusion of restaurant fixed effects substantially reduces variation in treatment. We therefore employ a careful control strategy, using the following questions as control variables: 5 "How often do you visit a [brand name] restaurant?", 6.1 "What meal did you have today?", 7 "I visited the restaurant today accompanied by ___, age, and gender.

We report the means and standard errors of each variable for each treatment group in Table 2, and the results show that we have balance across treatments with respect to the control variables. We do not report the standard errors for binary responses. The differences across treatments are negligible for almost all variables. One notable difference is that respondents in groups of friends are more common than lone respondents in Familiar B than in the other treatments. We also observe that the Familiar B respondents are a few years younger on average. However, because our focus is on the difference between the Novel and Familiar A playlists, we do not regard these minor imbalances as a threat to identification.

For proper statistical inference with clustered data, we must consider the correlation within the clusters. Each restaurant is a cluster within which we cannot assume independent residuals. The common practice to handle within-cluster correlation is cluster-robust standard errors. However, we cannot rely on the asymptotic properties of this approach when the number of clusters is as small as 16. With only 16 restaurants, cluster-robust standard errors over-reject the null hypothesis.

For reliable hypothesis testing, we use two bootstrap procedures. The Wild cluster bootstrap (Cameron *et al.*, 2008) provides reliable hypothesis testing, but it does not estimate standard errors for the coefficients. Therefore, we also use a bootstrap that resamples clusters to obtain standard errors. We present inference based on both bootstraps for the effect on sales outcomes presented in Table 3. However, because the different bootstraps provide very similar inference, we omit Wild bootstrap p-values for the remaining estimates. For robustness, we also perform randomization inference (Rosenbaum *et al.*, 2002).

Finally, to show that treatment does not coincide with heterogeneous seasonal trends, we use sales data from the previous year. We apply the real treatment variables as if the experiment had been conducted one year earlier. If treatment does not coincide with seasonal trends, these estimated placebo effects should be close to zero.

3 Results

3.1 Main results

The main results are presented in Table 3. Each column shows the estimated effects on the corresponding sales outcome. Because the outcome variables are in logs and Novel is the baseline treatment, the coefficients represent the percentage difference in outcome between the corresponding music treatment and the Novel treatment. For inference, we present results from both types of bootstraps. We discuss statistical inference and robustness in further detail in Section 3.2.

All coefficients in Table 3 are negative, which implies that the Novel playlist

Table 2: Survey Data – Control Variables

Question	Novel	Familiar B	Familiar A	Silence
Visit frequency	3.980	4.050	3.856	4.016
	(1.438)	(1.452)	(1.413)	(1.432)
Which meal?				
Breakfast	.099	.102	.134	.122
Lunch	.526	.453	.476	.497
Snack	.163	.208	.146	.144
Dinner	.208	.235	.242	.236
Accompanied by?				
Partner	.129	.116	.098	.108
My children	.178	.158	.203	.238
Friends	.225	.312	.189	.150
Siblings	.011	.016	.014	.012
Alone	.395	.280	.423	.435
Other	.062	.118	.073	.058
Age	41.659	35.871	39.266	41.792
	(17.409)	(17.455)	(15.540)	(15.146)
Male	.577	.554	.646	.613
Number of observations	534	558	508	501

generates the highest value for all sales outcomes. The differences in music familiarity are largest between Novel and Familiar A, and we therefore focus on the coefficients for Familiar A.

The effect on average purchase size is -0.8, which suggests that the average purchase size is 0.8 percent lower when music is familiar than when it is novel. The difference is larger when we consider revenues, namely -4.1 percent. Because revenues decrease more than average purchase size, the number of purchases must also be decreasing. This is supported by the results in Column 4, showing that the effect of the Familiar A playlist on the number of purchases is -3.2 percent. Thus, the effect of playing familiar music on the number of purchases is driven by its effect on additional purchases.

As we describe in Section 2.2, if the treatment effect on both average purchase size and the number of purchases is negative, the effect on the size of main purchases must also be negative. This result means that while some customers revise their orders immediately, some customers return for an additional purchase instead. From the negative estimated effect of Familiar A on the number of purchases, we know that fewer customers return for an additional purchase when the music is familiar. However, this effect is only compatible with a simultaneous decrease in average purchase size if some customers decrease the size of their main purchase. Hence, familiar music makes fewer customers return for an additional purchase and means that some customers has already decreased their consumption in the main purchase.

Because the effects on revenues and quantity are roughly equal, the change in sales is driven by a change in the number of units sold. However, the marginal units that the familiarity of the music affects are similarly priced to the average item sold. If the marginal units were higher priced, revenues would change more than quantity and vice versa. In the sample, the average price of a sold unit is approximately 20 Swedish Kronor ($ 2). 20 Swedish Kronor is the typical price of milkshakes, smoothies, cappuccinos and lattes, and some types of ice cream.

The difference in sales between Familiar A and Familiar B is not statistically significant. This is the difference that measures the effect of aligning the music with the brand values of the company. We therefore conclude that this effect is negligible. There is also a very small difference in familiarity between these treatments. However, this difference is not substantial enough to show any effect on sales.

The effect of silence is of limited interest because it is not clear whether silence is distracting. On the one hand, customers likely expect there to be background music in the restaurant. Therefore, silence might be surprising and distracting. On the other hand, silence does not provide any additional audio that might distract.

Next, to understand the mechanisms by which familiar music affects sales, we exploit our complementary survey data. In the survey results in Table 4, Column 1 reveals when customers notice the music. Because the effect of Familiar A (and B) is positive and significant, more customers notice the music when it is familiar. Column 2 shows what music customers recognize, pro-

Table 3: Main Results

	(1) Purchase size	(2) Revenues	(3) Quantity	(4) # Purchases
Familiar A	$-.008^{*}$	$-.041^{***}$	$-.044^{***}$	$-.032^{**}$
	(.004)	(.015)	(.016)	(.015)
	[.047]	[.005]	[.025]	[.021]
Familiar B	$-.013^{**}$	$-.032^{*}$	$-.026^{*}$	$-.019$
	(.005)	(.017)	(.016)	(.014)
	[.008]	[.011]	[.031]	[.075]
Silence	$-.006$	$-.019$	$-.022$	$-.012$
	(.007)	(.014)	(.013)	(.014)
	[.190]	[.028]	[.005]	[.122]
Observations	2,232	2,232	2,232	2,232
R^2	.084	.178	.159	.159

Bootstrapped standard errors in parentheses, clustered at the restaurant level. ***: p<0.01, **: p<0.05, *: p<0.1. 1,000 replications. Wild Bootstrap p-values in brackets. 1,000 replications. Fixed effects are at the restaurant and week levels.

vided that they notice it. The results show that customers recognize the music more when it is familiar, which ensures us that the Familiar A (and B) treatment actually contains more familiar music. Column 3 shows that customers on average prefer familiar music to novel music.

We can thus conclude that while customers notice, recognize, and prefer familiar music, the familiar music makes them purchase less. Therefore, we suggest that familiar music decreases sales because it attracts attention. When a customer is distracted by the familiar music, she allocates less attention to potential consumption. As she considers fewer items, she also purchases less.

The results in column 4 also indicate that 4.1 percent fewer customers return to the counter for an additional purchase during the Familiar A playlist than in the Novel treatment. This effect of familiarity is consistent with the effect on the number of purchases in Table 3. Therefore, we conclude that the additional purchases drive the difference in the total number of purchases. As this result bolsters the assumption of no treatment effect on the total number of costumers, we are confident in the reasoning behind and conclusion of which types of purchases depend on the music. Specifically, familiarity with the music decreases both the number of additional purchases and the size of main purchases.

Table 4: Survey Results

	(1) Notice	(2) Recognize	(3) Like	(4) Additional purchase
Familiar A	.111***	.444**	.287*	−.041*
	(.030)	(.175)	(.154)	(.022)
Familiar B	.104***	.435***	.368**	.002
	(.029)	(.162)	(.152)	(.024)
Silence	−.046	−.350*	−.097	−.010
	(.031)	(.197)	(.171)	(.023)
Constant	.853***	6.863***	6.886***	.300***
	(.063)	(.351)	(.309)	(.051)
Observations	2,101	1,110	1,118	2,101
R^2	.085	.170	.123	.022
Outcome	0/1	1-6	1-6	0/1

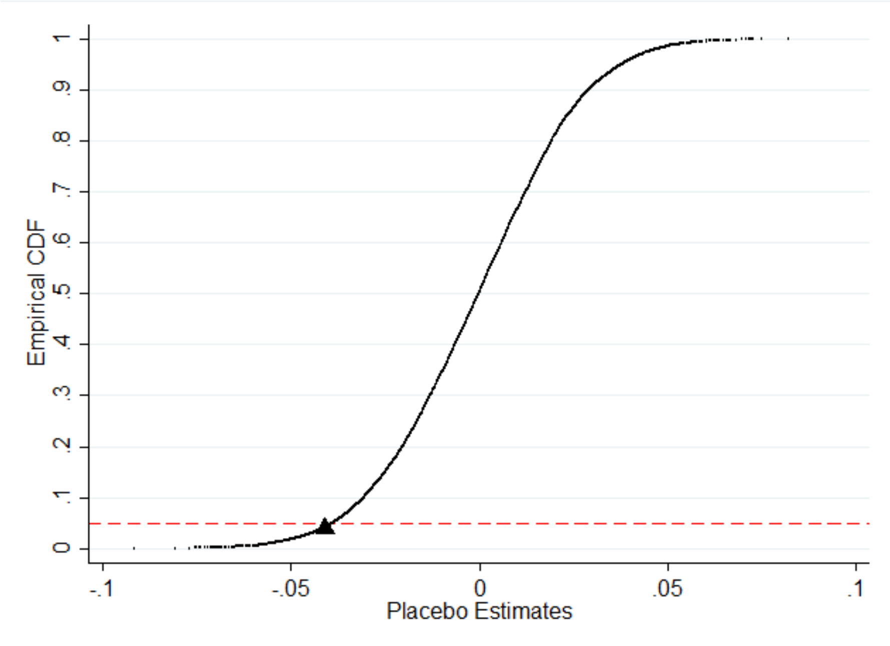
Linear estimates. Robust standard errors in parentheses. ***: p<0.01, **: p<0.05, *: p<0.1. All regressions control for questions 5, 6.1, 7, age, and gender.

3.2 Randomization inference and placebo estimations

This section discusses the inference and potential coincidence of the assigned treatment with heterogeneous seasonal trends. In addition to bootstrapping, we use randomization inference for robustness. To address questions about seasonal trends coinciding with treatment assignment, we use sales data from the corresponding time period but in the previous year.

We perform randomization inference as follows. First, we randomly assign treatment according to the same procedure as in the original treatment assignment. Then, we estimate the effect of these new, placebo, treatment variables on revenues. We repeat this procedure 10,000 times, resulting in 10,000 estimated placebo treatment effects. Because we assign these placebo treatment variables to restaurants and time periods at random, the estimated treatment effects should on average be zero.

Figure 2: Randomization Inference



Empirical CDF of placebo treatment effects, 10,000 iterations. Each dot is one placebo estimate of revenues. The triangle is the true ITT estimate of revenues.

Figure 2 plots the cumulative distribution of the placebo estimates. The dashed line denotes 0.05. As expected, the distribution is centered on zero. The triangle is the true ITT estimate from the actual treatment assignment. Because it is below the dashed line, less than 5 percent of the placebo estimates are lower than the true estimated ITT. Therefore, we conclude that the estimated effect is unlikely to arise by chance.

To address the potential issue of heterogeneous trends coinciding with treatment, we use data from the year before the experiment. We match the treatment variables to these data as if we had conducted the experiment in that year. In this test, we have the actual combination of treatment variables, restaurants, and weeks. However, the sales data are from the year before the experiment. Because the treatment cannot affect sales in the previous year, these treatment variables are placebos. With the same specification as in the main results in Table 3, we estimate the effect of these placebo treatment variables. If heterogeneous seasonal trends do not coincide with the treatment assignment, the placebo treatments should not correlate with sales.

Table 5 provides the placebo estimates. Because no coefficient is significantly different from zero, we conclude that the treatment assignment did not coincide with heterogeneous seasonal trends.

Table 5: Placebo Estimation

	(1) Purchase size	(2) Revenues	(3) Quantity	(4) #Purchases
Familiar A	.005	−.020	−.030	−.025
	(.005)	(.018)	(.021)	(.020)
Familiar B	−.007	−.022	−.025	−.015
	(.008)	(.024)	(.023)	(.020)
Silence	−.001	−.010	−.015	−.009
	(.004)	(.017)	(.020)	(.018)
Observations	2,217	2,217	2,217	2,217
R^2	.068	.137	.141	.130
Number of restaurants	16	16	16	16

Bootstrapped standard errors in parentheses, clustered at the restaurant level. ***: p<0.01, **: p<0.05, *: p<0.1. 1,000 replications. Fixed effects are at the restaurant and week levels. Constant not presented.

4 Conclusion

Most previous studies assume that the environment sends out signals that affect consumers' emotional states, which leads them to either avoid or approach the environment. We propose another mechanism that instead depends on the consumer's attention, where the presence of distracting surroundings reduces attention paid to available choices and thereby decreases consumption. We call this the attention hypothesis.

By conducting a randomized field experiment on the effects of background music in 16 fast-food restaurants over 20 weeks, we found that both average purchase size and the number of purchases decreased with the familiarity of the music played in the restaurant. A survey of over 2,000 customers also showed that they noticed the music more often when the music was familiar. Because the customers noticed the music, the music competed for their attention. Therefore, we suggest that familiar music decreases sales because it attracts attention. When a customer is distracted by familiar music, she allocates less attention to potential consumption. As she considers fewer items, she also purchases less.

This paper suggests that policy planners and decision researchers need to increase the scope of decision design. If the environment affects the decision maker, the policy design must also account for the environment. Moreover, as the environment may influence the decision maker in a predictable way, the environment gives the policy designer new and useful tools.

We have shown that distracting information reduces consumption and that even a small change in the consumer's surrounding environment can influence sales. Our results have important managerial implications for retailers and restaurants since they are constantly searching for new techniques to influence customers. This paper provides both specific and general techniques. Specifically, stores and restaurants should avoid playing familiar music because it distracts consumers. Generally, competing informational cues should be avoided because distracted consumers tend to purchase less.

The attention hypothesis needs formalization to provide a general theory with clear predictions. Furthermore, the hypothesis requires additional testing. However, in this paper, we have taken a first step towards a new theory of attention and consumption. The theory does not consider the attention paid to different alternatives or features of alternatives but purely the amount of attention allocated to the choice. As increasing numbers of informational cues compete for consumer attention, this question takes on greater importance.

References

Areni, Charles S, & Kim, David. 1993. The influence of background music on shopping behavior: classical versus top-forty music in a wine store. *ACR North American Advances*.

Beverland, Michael, Lim, Elison Ai Ching, Morrison, Michael, & Terziovski, Milé. 2006. In-store music and consumer–brand relationships: Relational transformation following experiences of (mis) fit. *Journal of Business Research*, **59**(9), 982–989.

Cameron, A Colin, Gelbach, Jonah B, & Miller, Douglas L. 2008. Bootstrap-based improvements for inference with clustered errors. *The Review of Economics and Statistics*, **90**(3), 414–427.

Donovan, Robert J, & Rossiter, John R. 1982. Store atmosphere: an environmental psychology approach. *Journal of retailing*, **58**(1), 34–57.

Donovan, Robert J, Rossiter, John R, Marcoolyn, Gilian, & Nesdale, Andrew. 1994. Store atmosphere and purchasing behavior. *Journal of retailing*, **70**(3), 283–294.

Gabaix, Xavier. 2014. A sparsity-based model of bounded rationality. *The Quarterly Journal of Economics*, **129**(4), 1661–1710.

Garlin, Francine V, & Owen, Katherine. 2006. Setting the tone with the tune: A meta-analytic review of the effects of background music in retail settings. *Journal of Business Research*, **59**(6), 755–764.

Hefti, Andreas, & Heinke, Steve. 2015. On the economics of superabundant information and scarce attention. *Œconomia. History, Methodology, Philosophy*, 37–76.

Jonaityte, Inga. 2016. Salience, Selective Attention and Learning with Information-Overload. *Department of Management, Università Ca' Foscari Venezia Working Paper*.

Kahneman, Daniel, Knetsch, Jack L, & Thaler, Richard H. 1990. Experimental tests of the endowment effect and the Coase theorem. *Journal of political Economy*, **98**(6), 1325–1348.

Kőszegi, Botond, & Szeidl, Adam. 2012. A model of focusing in economic choice. *The Quarterly journal of economics*, **128**(1), 53–104.

Mattila, Anna S, & Wirtz, Jochen. 2001. Congruency of scent and music as a driver of in-store evaluations and behavior. *Journal of retailing*, **77**(2), 273–289.

Mehrabian, Albert, & Russell, James A. 1974. *An approach to environmental psychology*. The MIT Press.

Milliman, Ronald E. 1982. Using background music to affect the behavior of supermarket shoppers. *The journal of Marketing*, 86–91.

Milliman, Ronald E. 1986. The influence of background music on the behavior of restaurant patrons. *Journal of consumer research*, **13**(2), 286–289.

North, Adrian C, Hargreaves, David J, & McKendrick, Jennifer. 1999. The influence of in-store music on wine selections. *Journal of Applied psychology*, **84**(2), 271.

Nothdurft, Hans-Christoph. 2000. Salience from feature contrast: additivity across dimensions. *Vision research*, **40**(10), 1183–1201.

Oakes, Steve. 2003. Musical tempo and waiting perceptions. *Psychology & Marketing*, **20**(8), 685–705.

Rosenbaum, Paul R, *et al.* 2002. Covariance adjustment in randomized experiments and observational studies. *Statistical Science*, **17**(3), 286–327.

Simon, Herbert A. 1955. A behavioral model of rational choice. *The quarterly journal of economics*, **69**(1), 99–118.

Sims, Christopher A. 2003. Implications of rational inattention. *Journal of monetary Economics*, **50**(3), 665–690.

Thaler, Richard. 1985. Mental accounting and consumer choice. *Marketing science*, **4**(3), 199–214.

Tversky, Amos, & Kahneman, Daniel. 1975. Judgment under uncertainty: Heuristics and biases. *Pages 141–162 of: Utility, probability, and human decision making*. Springer.

Tversky, Amos, & Thaler, Richard H. 1990. Anomalies: preference reversals. *The Journal of Economic Perspectives*, **4**(2), 201–211.

Wilson, Stephanie. 2003. The effect of music on perceived atmosphere and purchase intentions in a restaurant. *Psychology of music*, **31**(1), 93–112.