CESifo WORKING PAPERS

7475
2019

January 2019

To Buy or not to Buy? Shrouding and Partitioning of Prices in an Online Shopping Field Experiment

Markus Dertwinkel-Kalt, Mats Köster, Matthias Sutter

CESifo

Impressum:

CESifo Working Papers
ISSN 2364-1428 (electronic version)
Publisher and distributor: Munich Society for the Promotion of Economic Research - CESifo GmbH
The international platform of Ludwigs-Maximilians University's Center for Economic Studies and the ifo Institute
Poschingerstr. 5, 81679 Munich, Germany
Telephone +49 (0)89 2180-2740, Telefax +49 (0)89 2180-17845, email office@cesifo.de
Editor: Clemens Fuest
www.cesifo-group.org/wp

An electronic version of the paper may be downloaded
· from the SSRN website: www.SSRN.com
· from the RePEc website: www.RePEc.org
· from the CESifo website: www.CESifo-group.org/wp

CESifo *Working Paper No. 7475*
Category 13: Behavioural Economics

To Buy or not to Buy? Shrouding and Partitioning of Prices in an Online Shopping Field Experiment

Abstract

We examine whether shrouding surcharges or partitioning prices raises demand in online shopping where consumers have very low costs of cancelling an initiated purchase process. In a field experiment with more than 34,000 consumers, we find that consumers in the online shop of a large German cinema initiate a purchase process more often when surcharges are shrouded or indicated separately, but they also drop out more often when the overall price becomes known at the check-out. In sum, the demand distribution is independent of the price presentation. This result qualifies previous findings on the effectiveness of such pricing practices and can be rationalized through low cancellation costs.

JEL-Codes: D810, C930.

Keywords: salience, inattention, shrouding, price partitioning, field experiment.

Markus Dertwinkel-Kalt
Frankfurt School of Finance & Management
Adickesallee 32-34
Germany - 60332 Frankfurt am Main
m.dertwinkel-kalt@fs.de

Mats Köster
University of Düsseldorf
Universitätsstr. 1
Germany - 40225 Düsseldorf
mats.koester@hhu.de

Matthias Sutter
Max Planck Institute (MPI) for Research on Collective Goods
Kurt-Schumacher-Str. 10
Germany - 53113 Bonn
matthias.sutter@coll.mpg.de

January 17, 2019
We thank Arno Apffelstaedt, David Birke, Katie Coffmann, Nicola Gennaioli, Paul Heidhues, Johannes Johnen, David Laibson, Axel Ockenfels, Gerhard Riener, Dmitry Taubinsky, and Achim Zeileis for very helpful comments and suggestions. Financial support through the German Science Foundation (Project number 404416232 and Excellence Cluster "ECONtribute: Markets and Public Policy), through the Graduate Programme Competition Economics (GRK 1974; University of Düsseldorf) of the German Science Foundation, and through C-SEB (University of Cologne) is gratefully acknowledged. Part of the research was carried out while Mats Köster was visiting UC Berkeley and he is grateful for their hospitality. Judith Epping provided excellent research assistance.

1. Introduction

It is well-known that consumers have limited attention and therefore often act myopically rather than in a sophisticated way, for which reason the presentation and salience of prices can affect their purchasing behavior. On the one hand, there are numerous examples on how companies try to increase their profits by making prices less salient, in particular by shrouding price attributes (like surcharges or add-ons) or by partitioning the overall price into several components (see Heidhues and Kőszegi, 2018, for an overview). On the other hand, also policy makers have adopted the presumption that consumers' attention is limited and try to design policies that improve decision making of inattentive consumers (for an overview see Bernheim and Taubinsky, 2018).[1] In order to design optimal policies in the presence of inattentive consumers, however, it is necessary to know how *and* under which circumstances inattention and price (non-)salience matters for consumption behavior.

An important and widely acknowledged example of salience effects is that of consumers under-reacting to non-salient taxes. In a seminal paper, Chetty et al. (2009) report that displaying sales taxes on the price tags reduces demand for cosmetic and beauty products by 8%, meaning that tax-exclusive prices lead to higher revenues in their field experiment in a grocery store. Experiments by Feldman and Ruffle (2015) or Taubinsky and Rees-Jones (2018) also find that shrouding taxes, when displaying prices, increases average demand. Similarly, tolls can be made non-salient by adopting an electronic toll collection system, thereby raising toll rates (Finkelstein, 2009). Closely related, experiments on partitioned pricing have shown that a decrease in the base price and a likewise increase in the add-on price raises demand as consumers tend to underweight additional price components. The effect of price partitioning has been shown to hold with shipping charges on eBay (Hossain and Morgan, 2006; Brown et al., 2010) and in hypothetical experiments on various goods and surcharges (Morwitz et al., 1998; Carlson and Weathers, 2008).

One important feature of the above-mentioned studies is that the consumers were either not presented the full price prior to the purchase (Hossain and Morgan, 2006; Brown et al., 2010; Morwitz et al., 1998), or the full price was only presented when consumers were practically *locked in* in the sense that it was relatively costly to cancel the purchase process. For example, in the field experiment of Chetty et al. (2009), rescinding from the purchase at the cashier, when

[1] Several countries – including the U.S., the U.K., and Germany – invest in consumer protection to prevent firms from exploiting unsophisticated consumers, and some countries (such as the U.S. or the U.K.) have even implemented "behavioral insights" teams to improve government policies based on psychologically more realistic views of consumption behavior.

being negatively surprised by the overall price, costs time and feels awkward. Other consumers might be annoyed by the delay and may get the impression that a consumer cancelling the purchase process just could not afford what she put in her shopping basket, thus negatively affecting the consumer's social image. Indeed, it might be exactly this interaction of salience and cancellation costs because of which shrouding or partitioning prices affects actual demand.

In our field experiment, we investigate whether varying the salience of surcharges (by shrouding them or presenting them separately) also affects demand when it is very cheap for the consumer to drop out after she has learned the full price. In other words, we examine whether presenting the full price before the confirmation of the purchase can de-bias inattentive consumers who are susceptible to shrouding or price partitioning. According to the studies mentioned above, inattentive consumers might be biased in the sense that they show higher demand when some price attributes are shrouded or separately displayed. But, if the full price is shown quickly after consumers put an item in their shopping basket and if the costs of cancelling an initiated purchase process are low, the bias from shrouding or price partitioning may not have any effect on actual purchases. An environment with low costs of cancelling a purchase process is different from the settings in the above-mentioned papers, but it is typical for many online stores. Arguably, telling a cashier to cancel the entire purchase as a consequence of an unexpectedly high tax amount represents a higher hurdle (i.e., gives rise to *social cancellation costs*) than just abandoning the confirmation screen of a purchase in an online store. Put differently, in the moment the full price is presented the consumer is typically locked in offline, but not necessarily online, because online cancellations avoid any kind of social image concerns and are only a matter of a few mouse-clicks and thus seconds, so that de-biasing may indeed work better online. We thus ask whether the effects of price salience on demand carry over to situations where the costs of cancellation are small.

We consider this an important question in light of the rapidly growing literature on *Behavioral Public Economics* that derives optimal policy adjustments in response to consumer inattention (see, e.g., Farhi and Gabaix, 2018, for a general framework to study optimal taxation with inattentive consumers). For that, it is important to understand under which circumstances (and to which degree) demand responds to salience effects and in which not. In times of digitization,[2] we regard it as particularly important to study the role of salience effects when (unlike in the case of offline purchases) social cancellation costs are zero and overall

[2] Online sales have been steadily increasing, amounting to $395 billion (11.7% of overall sales) in the United States and $1.9 trillion (8.7% of total retail spending) worldwide in 2016. See https://www.census.gov/retail/mrts/www/data/pdf/ec_current.pdf for the official report by the U.S. Commerce Department, and also see the e-commerce report by Statista, https://www.statista.com/statistics/379046/worldwide-retail-e-commerce-sales (downloaded October 14, 2018).

cancellation costs are also very low. By examining the effects of different price presentations on purchasing behavior in such a low-cancellation-costs environment, our study is not only informative for regulatory authorities that are dealing with the question on how online price presentations should be regulated, but it also points to a *context-dependency* of salience effects that has been understudied so far.

We present a field experiment run in collaboration with the online store of a large German multiplex cinema from April 2017 to January 2018. During this period, we manipulated the presentation of prices for 3D movies. These prices consist of a base price, which varies across movies and days, and an additional 3D surcharge of 3 Euro. We implemented three treatments by presenting either (i) the full price including the 3D surcharge, (ii) the base price with a small footnote indicating that an additional 3D surcharge has to be paid (i.e., shrouding), or (iii) the base price and the 3D surcharge separately (i.e., partitioning). In this setup, we can observe when consumers initiate a purchasing process, and can then examine whether the treatments, i.e., the different price frames, have an impact on a consumer's likelihood to (1) proceed to the check-out where the full price is presented before actual purchase, and (2) actually buy the product. Examining both parts of a purchasing process separately allows us to study whether different price frames affect the initiation and the actual completion of a purchase to the same extent or whether consumers react to price frames only in one part, but not in the other.

The feature of studying both parts of the purchase process sheds light on how consumers react to learning the full price at different stages of the process and allows us to test whether consumers can be de-biased. This distinguishes our study from recent experiments by Feldman and Ruffle (2015) and Taubinsky and Rees-Jones (2018) where participants could not choose to cancel a purchase process, but had to make choices (under different frames with respect to the salience of taxation) that were immediately implemented. Another novelty of our paper is that we study shrouding and partitioning of prices in a unified framework, thus making it possible to compare the relative effects of both pricing practices under *ceteris paribus* conditions. To the best of our knowledge, we are also the first incentivized experiment where partitioning of prices (one of our treatments) is examined in the cleanest possible way, namely by only splitting up the price, holding everything else constant. Beside splitting up the total price, previous studies on partitioned pricing (see Greenleaf et al., 2016, for a survey) also presented surcharges less prominently than the base price, for instance by using different locations or smaller font sizes for surcharges. In contrast, when partitioning the overall price, we present both price components at an equally prominent location using the same font size.

Tracking more than 34,000 consumers over a period of 9 months, we find that shrouding the 3D surcharge significantly increases the likelihood that a consumer places tickets for a 3D movie in the shopping basket. More specifically, under the assumption that inattention to shrouded surcharges is independent of the valuation for watching a 3D movie, we estimate that around 10% of our consumers neglect the shrouded fee at the initial screen. While inattention to the shrouded fee is in line with the existing empirical literature, we also find that shrouding does *not* affect actual purchases, which is contrary to findings in Chetty et al. (2009) and other papers discussed above. This null-result arises from consumers in the shrouding treatment dropping out much more likely once they see the overall price at the check-out. We conclude that – at least in an environment with low cancellation costs – consumers can be easily de-biased by presenting them the full price before they complete the purchase, which suggests that shrouding effects depend very much on the context in which consumers make their purchase.

Regarding the effects of partitioning prices (i.e., showing the base price and the 3D surcharge separately) we find that the probability to proceed to the check-out is only slightly higher than in case the full price is shown alone, but again we see no effect on the likelihood to actually buy movie tickets. This suggests that previous findings about substantial effects of partitioned pricing (e.g., Hossian and Morgan, 2006) depended upon a combination of partitioning prices and using different locations and font sizes for surcharges, while the splitting up of prices under *ceteris paribus* conditions does not seem to generate a difference in purchasing behavior.

The rest of the paper is organized as follows. In Section 2 we introduce our experimental design. Sections 3 and 4 present the results. Section 5 discusses our main findings and concludes the paper.

2. Experimental Design

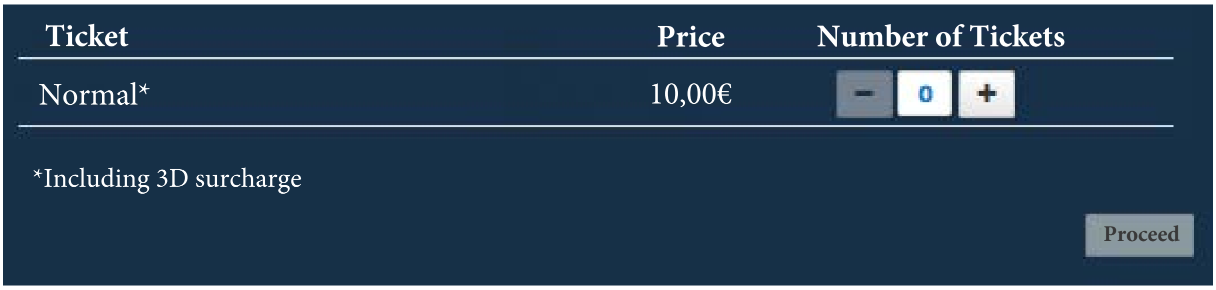
We conducted a natural field experiment (List and Rasul, 2011) in cooperation with a large German multiplex cinema that allowed us to vary the presentation of prices for 3D movies in its online store that has on average more than 10,000 online bookings per month. In Germany, the price of a ticket for a 3D movie always includes the base price for a movie ticket plus a 3D surcharge, which amounts in the case of our cinema – as it is typical for German cinemas – to 3 Euro. Typically, the same movie is also shown in a 2D variant for which the surcharge does not apply. These 2D shows take place in a different room within the same multiplex at a potentially different time.

Purchase Process. The purchase process in the cinema's online store consists of three steps: on a first screen, the consumer selects the number of tickets she would like to buy for a particular movie. On a second screen, the full price is presented and the consumer has to enter her payment details. On a third screen, all relevant information is summarized and the consumer has to finally confirm the purchase. Our treatment variation concerns only the first screen on which we vary the presentation of prices. The second and third screen are identical across treatments.

Treatments. The price presentation on the initial screen varies across three treatments:

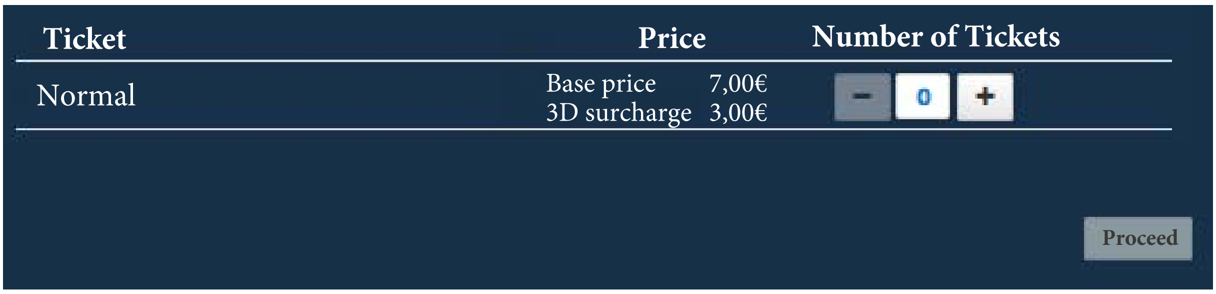
- ***Inclusive***: In the first treatment, we present the overall ticket price, including the 3D surcharge, and add a footnote stating that the surcharge is already included (for an illustration see Figure 1 and for the actual screen see Figure E.1 in the Appendix).

Figure 1. *Stylized price presentation on the initial screen in Inclusive.*



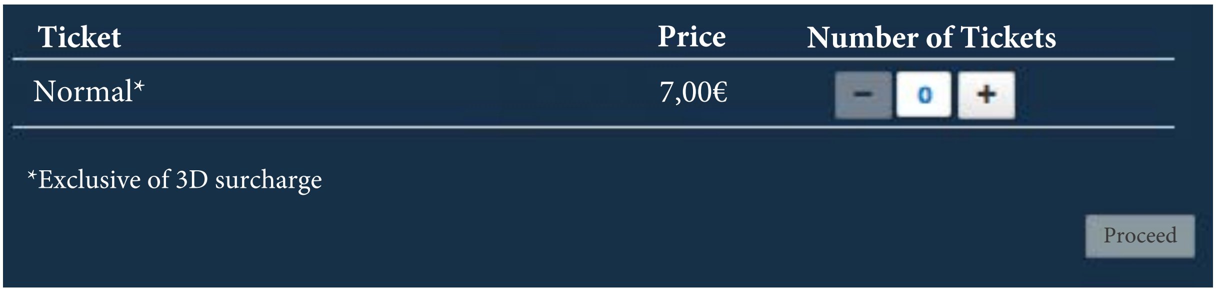
- ***Partitioned***: In this treatment, we split up the full price by presenting the two price components – the base price and the 3D surcharge – in separate lines, but identical font and font size (see Figure 2 and the actual screen in Figure E.2 in the Appendix).

Figure 2. *Stylized price presentation on the initial screen in Partitioned.*



- ***Shrouded***: In the third treatment, we "shroud" the 3D surcharge by presenting the base price and mentioning the additional surcharge (but not the exact amount) only in a footnote (see Figure 3 and the actual screen in Figure E.3 in the Appendix).[3]

Figure 3. *Stylized price presentation on the initial screen in Shrouded.*



Randomization and Identifying Assumption. In order to buy tickets for a certain movie via the cinema's online store, a consumer has to browse the cinema's schedule on the homepage, then she has to click on a particular show of this movie, and afterwards she has to login with her email address and a password. Only after logging in, a consumer sees the initial screen of the purchase process (i.e., the price of a ticket as presented above) and chooses how many tickets she would like to have. Each consumer has a unique user ID, based on which we randomized our treatment assignment. This implies that each consumer is assigned the same treatment over the entire duration of the experiment. Our identifying assumption then is that the random assignment of the treatment worked properly.

Hypotheses. In Appendix A, we derive the following two hypotheses on behavior at different stages of the purchase process from a model that builds on the idea of focusing by Kőszegi and Szeidl (2013), but extends their model to (i) cover the case of shrouded price attributes and (ii) incorporate the costs of cancelling a purchase process. Under the assumption that consumers have very low costs to cancel a purchase process, the model yields both hypotheses.

Our first hypothesis refers to the likelihood of initiating a purchase process for a 3D movie. We expect this likelihood to be lowest in treatment *Inclusive*, intermediate in *Partitioned*, and highest in *Shrouded*. According to the well-known contrast effect (e.g., Schkade and

[3] Since 3D surcharges are almost the same across cinemas all over Germany, the typical consumer is not only aware of the fact that such a surcharge applies, but can be assumed to have a good knowledge of its size even before the first purchase in our cinema (and, for certain, after the first purchase). As Bernheim and Taubinsky (2018) argue, if consumers were used to see the price exclusive of the surcharge (e.g., consumers being used to tax-exclusive prices as in Chetty et al., 2009), good knowledge about the surcharge might be problematic, because consumers could misinterpret the surcharge-inclusive price as an increase in the base price. This should be no concern in our setup as surcharge-inclusive prices were used prior to our intervention.

Kahneman, 1998; Dunn et al., 2003), partitioning the overall price into two price components diverts attention away from the overall price, so that the likelihood to initiate a purchase process for a 3D movie should be higher in *Partitioned* than in *Inclusive*. Hiding the 3D surcharge in a footnote should make the overall price even less salient, so that the likelihood to initiate a purchase process for a 3D movie should be highest in *Shrouded*.

Hypothesis 1. *The likelihood to initiate a purchase process for a 3D movie is lowest in Inclusive, at an intermediate level in Partitioned, and highest in Shrouded.*

Our second hypothesis refers to the likelihood of purchasing tickets for a 3D movie. We expect this likelihood to be independent of the price presentation. Given that the costs of cancelling an initiated purchase process on the second or third screen are arguably low, consumers in *Partitioned* and *Shrouded* should be fully de-biased when observing the full price, so that purchase rates should *not* differ across the three treatments.

Hypothesis 2. *The likelihood to buy tickets for a 3D movie does not vary across treatments*.

Discussion of our Design. We conclude this section with a brief discussion of the design. We have found a unique setting that allows us to test for the demand effects of different pricing frames in an environment with low costs of cancelling an initiated purchase process. In addition, as we observe decisions at all steps of the purchase process, we are able to compare a consumer's behavior inside a given *price frame* (on the initial screen) to her behavior outside of the frame (on the second and third screen). In previous studies, which documented the importance of related framing effects, the consumers were *either* not presented the full price prior to the purchase (i.e., the final decision was made *within* the price frame; e.g., Brown et al., 2010), *or* the full price was only presented when consumers were *locked in* in the sense that it was quite costly to cancel the purchase process at this stage (e.g., Chetty et al., 2009).

From a methodological perspective, our experimental design has two further advantages compared to previous studies such as Chetty et al. (2009). On the one hand, we assigned treatments randomly based on a unique user ID, so that our treatment effects are identified as accurately as in a laboratory setting. On the other hand, we observe not only aggregate revenues, but individual decisions throughout the whole purchase process, which provides us with information on how a specific consumer behaves inside and outside of a given price frame. Moreover, we tracked the consumers' purchase history over the course of the experiment, so that we can also analyze long-term framing effects.

3. Main Empirical Analysis: Average Treatment Effects

3.1. Data, Descriptive Statistics, and Empirical Strategy

Data. Our intervention ran from April 24, 2017 until January 14, 2018. During this treatment period, we tracked all clicks in the cinema's online store at the level of an individual consumer, whereby each click refers to a different purchase process. In the following analysis, we consider all 34,902 consumers who have clicked at least once on a 3D movie during our intervention, thereby being treated. For some randomization checks, we also take into account when these consumers were interested in 2D movies. Yet, we exclude further consumers who were only interested in 2D movies and never clicked on a show of a 3D movie.

Descriptive Statistics. Over the course of our experiment, the 34,902 consumers, who clicked on a 3D movie at least once, started 173,695 purchase processes for both 3D and 2D movies, and around 30% (i.e., in total 53,357) of these purchase processes were actually completed. The majority of these purchase processes refer to 3D movies (i.e., 100,300 out of 173,695 purchase processes, see Panel B of Table 1 for 3D movies and Panel C for 2D movies), suggesting that the treated consumers have a preference for 3D movies.

As we can observe from Panel B in Table 1, the average purchase rate for 3D movies, when taking all clicks during our intervention into account, does not vary much across treatments, ranging only from 30.27% to 30.96%, which is a first hint in favor of Hypothesis 2. Notably, the results remain basically the same when taking for each consumer only her first click on a 3D movie during the treatment period into account (see Panel A in Table 1).

Consistent with Hypothesis 1, when considering 3D movies, the drop-out rate on the initial screen is highest in *Inclusive*, at an intermediate level in *Partitioned*, and lowest in *Shrouded.* When we look at 2D movies (in Panel C of Table 1) – where all consumers faced an identical initial screen – we see that drop-out rates on the initial screen and actual purchase rates are very similar across all treatment groups, which we consider a first indication of successful randomization.

Randomization Check. In order to test our identifying assumption of random treatment allocation in more detail, we performed a randomization check. The last row of Panel A in Table 1 shows the number of consumers assigned to each of the three treatments. Using a X^2-test, we cannot reject the null-hypothesis of a uniform distribution of consumers across treatments (p-value = 0.707). Also, when taking observables such as the month of the first click on a 3D show during our intervention (p-value = 1.000, X^2-test) or the 3D movie first clicked on (p-value = 0.933, X^2-test) into account, we cannot reject the null-hypothesis of random treatment allocation. This suggests that our identifying assumption is fulfilled.

Table 1. *Descriptive statistics for 3D movies and 2D movies across treatments.*

Panel A: 3D movies First Clicks	Absolute Frequencies			Relative Frequencies		
	Inclusive	Partitioned	Shrouded	Inclusive	Partitioned	Shrouded
Drop-out first screen	6,295	6,260	5,747	54.40%	53.81%	49.13%
Cancel later screen	1,352	1,535	2,025	11.68%	13.20%	17.31%
Purchase	3,924	3,838	3,926	33.91%	32.99%	33.56%
# Consumers	11,571	11,633	11,698	-	-	-
Panel B: 3D movies All Clicks	Absolute Frequencies			Relative Frequencies		
	Inclusive	Partitioned	Shrouded	Inclusive	Partitioned	Shrouded
Drop-out first screen	20,210	19,022	18,354	60.14%	58.23%	53.94%
Cancel later screen	3,158	3,530	5,373	9.40%	10.81%	15.79%
Purchase	10,238	10,115	10,300	30.46%	30.96%	30.27%
# Clicks in total	33,606	32,667	34,027	-	-	-
Panel C: 2D movies All Clicks	Absolute Frequencies			Relative Frequencies		
	Inclusive	Partitioned	Shrouded	Inclusive	Partitioned	Shrouded
Drop-out first screen	15,177	14,648	14,763	61.72%	60.54%	59.99%
Cancel later screen	1,917	2,030	2,156	7.80%	8.39%	8.76%
Purchase	7,493	7,519	7,692	30.48%	31.07%	31.25%
# Clicks in total	24,587	24,197	24,611	-	-	-

Notes to Table 1: *For 2D-movies, there were no treatments and thus also no first clicks. We split up the data for 2D movies according to the consumers' respective 3D-treatment.*

Empirical Strategy. Our empirical analysis is divided into two parts. In the first part, we will consider the subset of observations that includes for each consumer only her first click on a 3D show within the treatment period that lasted for 9 months (Panel A in Table 1). Here, we will test for treatment effects on the probability to drop out on the initial screen (Hypothesis 1) and on the probability to complete the purchase process (Hypothesis 2). In a second part, we will aggregate all observations over the entire duration of the experiment at the individual level (Panel B in Table 1) and we will test whether the treatments affect certain statistics such as the average number of drop-outs on the initial screen (Hypothesis 1) or the average number of purchases (Hypothesis 2). This allows us to test for long-term framing effects. As we discuss below, we have to deal with selection issues when testing our predictions on drop-out rates by using the average number of drop-outs within the treatment period. In the following, we briefly delineate our general empirical strategy and subsequently apply it to each part separately.

Average Treatment Effects. For each consumer i, denote as y_i the variable of interest (e.g., whether she buys tickets conditional on clicking for the first time on a 3D movie) and let T_i indicate her treatment. In addition, let X_i be observable controls (e.g., movie or time dummies).

Using *Inclusive* (henceforth: $incl$) as the default, we are interested in the effects of either partitioning the total price into its two components or shrouding the 3D surcharge. Formally, for each treatment $k \in \{Partitioned, Shrouded\}$, we want to determine the *average treatment effect* (ATE) relative to *Inclusive*,

$$\mathbb{E}[y_i|T_i = k, X_i] - \mathbb{E}[y_i|T_i = incl, X_i],$$

whereby for each consumer only one of the above terms is observed.

Using potential outcome notation, for an individual i assigned to treatment k, we denote as y_i^k the actual outcome of interest and as y_i^{incl} the potential outcome given that she was assigned to *Inclusive* instead. Then, we can rewrite the average treatment effect in terms of the *average treatment effect on the treated* (i.e., the first term in the second line below) and a selection bias (i.e., the second and third terms in the second line below):

$$\begin{aligned}&\mathbb{E}[y_i|T_i = k, X_i] - \mathbb{E}[y_i|T_i = incl, X_i]\\ &\qquad = \mathbb{E}\left[y_i^k - y_i^{incl}\middle|T_i = k, X_i\right] + \mathbb{E}\left[y_i^{incl}\middle|T_i = k, X_i\right] - \mathbb{E}\left[y_i^{incl}\middle|T_i = incl, X_i\right].\end{aligned}$$

Now, if treatment assignment is random, the last two terms cancel out (i.e., there is no selection bias) and the first term simplifies to $\mathbb{E}[y_i^k - y_i^{incl}|X_i]$, which can be consistently estimated from our sample.

Part 1: Subsample of First Clicks. In order to test our hypotheses on treatment effects, conditional on seeing the treatment for the first time, we run simple OLS regressions,

$$y_i = \alpha + \beta \cdot T_i + \mu \cdot X_i + \epsilon_i,$$

where y_i is *either* a binary indicator of whether consumer i drops out on the initial screen in the purchase process *or* a binary indicator of whether consumer i actually buys tickets, and ϵ_i is zero-mean noise. As we have seen above, given that we use for each consumer only the first purchase process for a 3D movie during our intervention, treatment allocation is indeed random, so that we are able to identify ATEs using OLS.[4]

Part 2: Aggregate Statistics. In order to test for the long-term effects of our different price frames, we aggregate all observations at the individual level. On the one hand, we will estimate count models with *either* the overall number of purchases (for both 3D and 2D movies and for each movie type separately) *or* the overall number of purchased tickets (again for all movies and for each movie type separately) as the dependent variable. In addition, we will look more directly into revenues and regress the per-customer revenue over the whole experiment on treatment indicators using OLS. On the other hand, we will try to understand whether the average number of drop-outs on the initial screen in the process of buying tickets for a 3D movie varies across treatments when taking all available information into account.

As long as we use the number of purchases or the number of purchased tickets as the dependent variable, we do not have to deal with selection issues when aggregating all observations at the individual level, since not entering the online store could be interpreted as not buying any tickets.[5] In other words, differential attrition is a crucial part of the potential treatment effects we are interested in and not a threat to identification. Formally, the corresponding treatment effects are identified due to the random treatment assignment conditional on the first click on a 3D movie, which we have verified above.

If we use the overall number of drop-outs on the first screen in the process of buying tickets for a 3D movie as the dependent variable, however, selection is an important issue. As we have seen in Table 1, our treatments affect the overall number of initiated purchase processes for 3D movies (see # Clicks in total in Panel B of Table 1). More specifically, regressing the number of purchase processes for 3D movies on treatment indicators shows that consumers in *Partitioned* start significantly less purchase processes than consumers in *Inclusive* (see Table F.2 in the Appendix). Hence, it is not innocuous to simply compare the average number of drop-outs on the first screen across treatments, since we do not know whether consumers in

[4] When testing for the ATEs on the probability to buy conditional on the first click on a 3D movie, we also take into account whether the consumer comes back later in order to buy tickets for the exact same show (after having cancelled the purchase process initially); that is, in this case y_i takes a value of one if and only if consumer i at some point in time buys tickets for the 3D show that she clicked on first during the treatment period. Notice, however, that our qualitative results do not change if we take also here only the very first click into account.

[5] This takes it for granted that there are no differential treatment effects on the probability to buy tickets directly at the cinema.

Partitioned would have dropped out more often on the first screen if they had started the same number of purchase processes as consumers in *Inclusive*. More formally, due to this differential attrition, random treatment assignment conditional on the first click on a 3D movie during the treatment period is not sufficient to identify treatment effects on the average number of drop-outs. In order to address this problem, we will first naively estimate a count model as if selection was not an issue, but then impose a worst-case scenario that qualifies our estimated treatment effects by providing a lower bound.

3.2. Treatment Effects on the Probability to Drop out on the Initial Screen

In this subsection, we want to test whether the different price frames systematically affect the probability to drop out on the initial screen in the process of buying tickets for a 3D movie.

Part 1: Subsample of First Clicks. In order to avoid selection issues and to rigorously test for Hypothesis 1, we first rely on the subsample of observations that includes for each consumer only her first click on a 3D movie during the treatment period. Using these observations, we regress a binary indicator of whether a consumer drops out on the initial screen in the purchase process on treatment indicators and controls. The regression results are presented in Table 2.

We observe that, relative to *Inclusive*, the average probability to drop out on the initial screen of the purchase process significantly decreases by around 5 percentage points in *Shrouded*, which is consistent with Hypothesis 1. The average probability to drop out on the initial screen in *Partitioned*, however, does not change significantly compared to *Inclusive*. Although the point estimate is negative, as expected, its insignificance is inconsistent with Hypothesis 1. The estimated treatment effects are robust to adding movie and time dummies as well as controls for whether a 2D show of the same movie is running at broadly the same time (i.e., within 1 hour) and for the time of the day at which the show runs. In the Appendix, we further verify that our results do not change when adding a control for whether the same 3D movie runs at a different cinema in the same city at broadly the same time. Moreover, the results are robust to interacting treatment indicators with an indicator of whether a 2D or 3D show of the same movie runs at the same or another cinema, respectively, at broadly the same time (Table C.1). We also show that treatment effects do *neither* differ for blockbuster movies (Table C.3) *nor* for movie shows running at the weekend (Table C.5).

Table 2. *Likelihood to drop out on initial screen during first click on a 3D movie.*

Parameter	Drop-out	Drop-out
Partitioned	-0.007	-0.007
	(0.007)	(0.006)
Shrouded	-0.054***	-0.053***
	(0.007)	(0.007)
2D Substitute	-	0.009
		(0.007)
Noon	-	-0.022**
		(0.010)
Afternoon	-	-0.026***
		(0.007)
Night	-	0.072***
		(0.009)
Movie FE	no	yes
Time FE	no	yes
# Observations	34,902	34,902

Notes to Table 2: *Results of an OLS-regression with a binary indicator of whether a consumer drops out on the initial screen conditional on clicking on a 3D movie for the first time as dependent variable and treatment indicators as independent variables (whereby Inclusive serves as base category). In the second column, we include movie, year, month, and day of the week fixed effects as well as controls for whether a 2D show of the same movie is running within +/- 1 hour, and for the time of the day at which the show runs (whereby Evening serves as base category). Standard errors are provided in parenthesis. Significance level: *: 10%, **: 5%, ***: 1%.*

Using the subsample of first clicks, we can also estimate the share of subjects who neglect shrouded fees. Under the assumption that a consumer who does not drop out in *Inclusive* does also not drop out in *Shrouded*, we estimate that around (0.544 – 0.4913)/0.544 = 9.69% of the subjects who drop out in *Inclusive* would *not* drop out in *Shrouded*, thereby revealing inattention to the shrouded surcharge. If we assume that inattention to shrouded surcharges is independent of the valuation for movie tickets and thus independent of the decision to drop out in *Inclusive*, then 9.69% gives a point estimate of the overall share of inattentive consumers.

Part 2: Aggregate Statistics. Next, we aggregate data over the entire intervention period and ask whether the treatments affect a consumer's overall number of drop-outs on the initial screen. As mentioned before, a selection issue arises when we naively compare the average number of drop-outs on the initial screen across treatments. Hence, we will proceed in two steps. First, we ignore the selection issue and estimate a count model as if selection was not a problem. Second, we impose a *worst-case scenario* that qualifies our initial estimates of ATEs.

Given the structure of our data, we estimate Negative Binomial models with the number of drop-outs on the initial screen in the process of buying tickets for a 3D movie as the dependent variable and treatment indicators as independent variables.[6] Table 3 presents the corresponding regression results, where the model estimated in the second column takes also the number of clicks on 3D movies into account (i.e., a model with exposure).

Table 3. *Average number of drop-outs on the initial screen.*

Parameter	**# Drop-outs**	**# Drop-outs**
Partitioned	-0.066***	-0.029***
	(0.018)	(0.010)
Shrouded	-0.107***	-0.107***
	(0.018)	(0.010)
Exposure	no	yes
# Observations	34,902	34,902

Notes to Table 3: *Results of regressing the number of drop-outs on the initial screen in purchase processes for 3D movies on treatment indicators (whereby Inclusive serves as base category), using a Negative Binomial model. Standard errors are provided in parenthesis. Significance level: *: 10%, **: 5%, ***: 1%.*

Here, we observe that, relative to *Inclusive*, the average number of drop-outs significantly decreases both in *Partitioned* and in *Shrouded*, which is consistent with Hypothesis 1. Yet, as we have discussed in the previous subsection, consumers in *Partitioned* clicked significantly less often on 3D movies than consumers in *Inclusive*, so that we may overestimate the ATE of partitioning the total price. That selection indeed matters is illustrated by the fact that the estimated treatment effect of splitting up the total price is much smaller in the model with exposure – which takes also the number of clicks on 3D movies into account – than in the model without exposure – which neglects the number of clicks on 3D movies. Hence, in the worst case, the estimated ATE might be fully driven by fewer clicks in *Partitioned*.

[6] For each of the three treatments, the conditional variance of the number of drop-outs largely exceeds the conditional mean (see Table B.1 in the Appendix). Also in formal tests, we can reject the null-hypotheses of mean-variance equivalence against the alternatives of overdispersion. Given these patterns, a Negative Binomial model is more appropriate than a Poisson model.

In order to address this selection issue, we impose a worst-case scenario by assuming that all "missing" clicks in *Partitioned* go against Hypothesis 1. As we have seen in Panel B of Table 1, consumers in *Partitioned* have started 939 purchase processes less for 3D movies than consumers in *Inclusive.* The most conservative way to test for the ATE of partitioning the total price is to assume that *all* missing clicks in *Partitioned* would be drop-outs on the first screen.[7] Then we add these missing drop-outs to those consumers with the lowest drop-out rates and the smallest numbers of clicks on 3D movies[8], as this maximizes the increase in the average drop-out rate.[9] Given these assumptions, we estimate Negative Binomial models – with and without exposure – with the accordingly adjusted number of drop-outs as the dependent variable and treatment indicators as independent variables to obtain a lower bound on the ATE. Table 4 presents the corresponding regression results.

Table 4. *Lower-bound estimation of drop-outs in Partitioned (worst-case scenario).*

Parameter	**# Drop-outs**	**# Drop-outs**
Partitioned	-0.018	-0.010
	(0.017)	(0.010)
Exposure	no	yes
# Observations	23,204	23,204

Notes to Table 4: *Results of a worst-case scenario in which we regress the adjusted number of drop-outs on the initial screen for 3D movies on treatment indicators (whereby Inclusive serves as base category and consumers in Shrouded are left out), using a Negative Binomial model and the adjusted number of clicks on 3D movies as exposure variable. Standard errors are provided in parenthesis. Significance level: *: 10%, **: 5%, ***: 1%.*

We observe that the point estimate of the lower bound on the ATE is negative (which is consistent with Hypothesis 1), but not significantly different from zero. Taking our highly significant, naive estimate of the ATE (see Table 3) together with this lower bound on the ATE, we would argue that partitioning the total price into its two components might have an effect in

[7] But recall from Table 1 that Partitioned has a drop-out rate from the first screen of about 58% over all clicks, so our worst-case scenario to estimate a lower bound of the treatment effect (i.e., assuming a drop-out rate among the missing subjects of 100% instead of 58%) may be overly conservative.

[8] There are more than 939 consumers in *Partitioned* with only a single click on a 3D movie and no drop-out on the initial screen (i.e., they either purchased or cancelled the process on consecutive screens). Among these consumers, we chose randomly and increased both the number of clicks on 3D movies and the corresponding number of drop-outs on the first screen by one to end up with the adjusted number of clicks and drop-outs, respectively, that we use to estimate a lower bound on the ATE.

[9] For illustrative purposes, denote as D_i the number of drop-outs and as N_i the number of clicks on 3D movies by consumer i. In addition, let $s_i := D_i/N_i$ be her drop-out rate. Now, increasing both D_i and N_i by one results in an increase of the drop-out rate by $(1 - s_i)/(N_i + 1)$, which decreases in s_i and N_i.

line with Hypothesis 1, although a potentially small one.Finally, we look more closely into the ATE of shrouding the 3D surcharge. As we have seen in Panel B of Table 1, consumers in *Shrouded* have started 421 purchase processes more (for 3D movies) than consumers in *Inclusive*. In order to obtain a lower bound on the ATE of shrouding on drop-out rates, we assume that consumers in *Inclusive* had started 421 additional purchase processes with no additional drop-outs on the initial screen. The most conservative way to allocate these "missing" clicks in *Inclusive* is to add them to those consumers with the highest drop-out rates and the smallest numbers of clicks on 3D movies[10], as this maximizes the decrease in the average drop-out rate.[11] Given these assumptions, we estimate Negative Binomial models – with and without exposure – with the number of drop-outs as the dependent variable and treatment indicators as independent variables to obtain a lower bound on the ATE. Table 5 reports the corresponding regression results.

Table 5. *Lower-bound estimation of drop-outs in Shrouded (worst-case scenario).*

Parameter	**# Drop-outs**	**# Drop-outs**
Shrouding	-0.107***	-0.094***
	(0.018)	(0.010)
Exposure	no	yes
# Observations	23,269	23,269

Notes to Table 5: *Results of a worst-case scenario in which we regress the number of drop-outs on the initial screen for 3D movies on treatment indicators (whereby Inclusive serves as base category and consumers in Partitioned are left out), using a Negative Binomial model and the adjusted number of clicks on 3D movies as exposure variable. Standard errors are provided in parenthesis. Significance level: *: 10%, **: 5%, ***: 1%.*

Obviously, the estimated ATE when using a model without exposure is exactly the same as before (see Table 3). If we instead take also the (adjusted) number of clicks on 3D movies into account, the ATE is slightly smaller than before (9.4% instead of 10.7%), but still highly significant and of economically relevant size. This suggests that also in the long-run shrouding of 3D surcharges decreases significantly the probability that a consumer drops out on the initial screen in the process of buying tickets for a 3D movie.

[10] There are more than 421 consumers in *Inclusive* with only a single click on a 3D movie and one drop-out on the initial screen. Among these consumers, we chose randomly and increased the number of clicks on 3D movies by one to end up with the adjusted number of clicks that we use to estimate a lower bound on the ATE.

[11] Using the notation from above, we conclude that increasing the number of clicks on 3D movies, N_i, by one decreases the drop-out rate by $s_i/(N_i + 1)$, which increases in s_i and decreases in N_i.

Summary on Drop-out Rates. Shrouding the 3D surcharge decreases the probability to drop out on the initial screen significantly, not only conditional on the first click on a 3D movie, but also when taking into account all information over the course of our 9-months-intervention. This finding is consistent with Hypothesis 1. Partitioning the total price into its two components seems to have at most a small effect on the probability to drop out on the initial screen. While there is no significant effect of partitioning on the drop-out probability when considering the subsample of first clicks, our results on aggregate drop-out rates suggest that there might be an effect in line with Hypothesis 1, although a potentially small one. Notably, the small effect size of partitioning the total price could be explained by the fact that some consumers do not perceive the two price components as different attributes, which would be in line with experimental findings from the lab (e.g., Dertwinkel-Kalt et al., 2017).

3.3. Treatment Effects on the Purchase Probability

In this subsection, we want to test whether our treatment manipulations systematically affect the probability to buy tickets for 3D movies.

Part 1: Subsample of First Clicks. In a first step, we again rely on the subsample of observations that includes for each consumer only her first click on a 3D movie during the intervention period. This allows us to rigorously test for Hypothesis 2, which states that the purchase rates should not differ significantly across treatments. Table 6 provides the results of regressing a binary indicator of whether a consumer purchases tickets for the 3D show that she clicked on first during our intervention on treatment indicators and controls.

We observe that, conditional on clicking for the first time on a 3D movie, the average probability to purchase tickets does not vary significantly across treatments. This observation is consistent with Hypothesis 2. These null-results are also robust to adding movie, year, month, and day of the week fixed effects as well as further controls for whether a 2D show of the same movie is running at roughly the same time and for the time of the day at which the show runs.[12] Interestingly, the average probability to buy decreases significantly by around 4 percentage points if a 2D show of the same movie is running at a similar time.[13] This finding suggests that

[12] In the Appendix, we further verify that our results are robust to adding a control for whether the same 3D movie runs at a different cinema in the same city at broadly the same time. Moreover, our results do not change when interacting treatment indicators with an indicator of whether a 2D or 3D show of the same movie runs at the same or another cinema, respectively, at roughly the same time (Table C.2). We also show that treatment effects do *neither* differ for blockbuster movies (Table C.4) *nor* for movie shows running at the weekend (Table C.6).

[13] Any interaction term of treatment and the indicator of a 2D substitute is insignificant (see Table C.2 in the Appendix), meaning that the substitution effect of 2D shows does not depend on the treatment.

at least some consumers consider the corresponding 2D show as a viable substitute to watching a 3D show of a given movie.

Table 6. *Likelihood to buy tickets conditional on the first click on a 3D movie.*

Parameter	**Purchase**	**Purchase**
Partitioned	-0.006	-0.006
	(0.007)	(0.007)
Shrouded	0.000	-0.000
	(0.007)	(0.006)
2D Substitute	-	-0.041***
		(0.007)
Noon	-	0.019*
		(0.010)
Afternoon	-	0.015**
		(0.007)
Night	-	-0.083***
		(0.009)
Movie FE	no	yes
Time FE	no	yes
# Observations	34,902	34,902

Notes to Table 6: *Results of an OLS-regression with a binary indicator of whether a consumer buys ticket(s) for the first 3D show that she clicked on during our intervention as dependent variable and treatment indicators as independent variables (whereby Inclusive serves as base category). In the second column, we include movie, year, month, and day of the week fixed effects as well as controls for whether a 2D show of the same movie is running within +/- 1 hour and for the time of the day at which the show runs (whereby Evening serves as base category). Standard errors are provided in parenthesis. Significance level: *: 10%, **: 5%, ***: 1%.*

Part 2: Aggregate Statistics. In a second step, we aggregate all observations at the individual level (i.e., all purchase processes that refer to 3D movies),[14] and ask whether our treatment manipulations affect (i) the overall number of purchases for 3D movies, and (ii) the overall number of purchased tickets for 3D movies. In addition, we look (iii) into revenues more directly and regress the overall per-customer revenue (i.e., for 2D and 3D movies together) on

[14] If we consider instead the samples of purchase processes that either refer to all movies (i.e., 2D and 3D movies) or only to 2D movies, the qualitative results remain the same (see Tables C.8 and C.9 in the Appendix).

treatment indicators. While we estimate count models in order to answer the first and the second question, respectively, we estimate a simple OLS regression to answer the last one.

Again, Negative Binomial models are appropriate given our data.[15] Table 7 presents the regression results regarding the overall number of purchases and purchased tickets for 3D movies, where the models estimated in the second and fourth column take also the number of clicks on 3D movies into account (i.e., we show models with exposure).

Table 7. *Average number of purchases or bought tickets for 3D over intervention period.*

Parameter	**# Purchases**	**# Purchases**	**# Tickets**	**# Tickets**
Partitioned	-0.017	0.003	0.007	0.011
	(0.014)	(0.015)	(0.016)	(0.016)
Shrouded	-0.005	-0.009	0.000	0.004
	(0.014)	(0.015)	(0.016)	(0.016)
Exposure	no	yes	no	yes
# Observations	34,902	34,902	34,902	34,902

Notes to Table 7: *Results of regressing the number of 3D purchases and the number of purchased ticket(s) for 3D movies on treatment indicators (with Inclusive as base category), using a Negative Binomial model. Standard errors are provided in parenthesis. Significance level: *: 10%, **: 5%, ***: 1%.*

We observe that neither the average number of purchases nor the average number of purchased tickets for 3D movies varies significantly across treatments. This implies that neither partitioning the total price into its two components nor shrouding the 3D surcharge affects the average demand for 3D movies.

Finally, we look into revenues more directly. Table 8 presents the results of regressing the per-customer revenue (for both 2D and 3D movies) over the intervention period on treatment indicators. Also, when using this direct measure of revenues (including both movie types), we do not find any significant treatment effects, which suggests that the final demand for movie tickets as well as the firm's profits (at least for given prices) are independent of the price frame.

[15] An overview of conditional means and variances is provided in Table B.2 in the Appendix.

Table 8. *Average per-customer revenue.*

Parameter	Revenue
Partitioned	0.514
	(0.759)
Shrouded	0.591
	(0.758)
# Observations	34,902

Notes to Table 8: *Results of an OLS-regression with the per-customer revenue as dependent variable and treatment indicators as independent variables (whereby Inclusive serves as base category). Standard errors are provided in parenthesis. Significance level: *: 10%, **: 5%, ***: 1%.*

Summary on Purchases. Altogether, neither partitioning the total price into its two components nor shrouding the 3D surcharge affects the average demand for 3D movies, which is consistent with Hypothesis 2. This finding is robust to using either the subsample of first clicks on a 3D movie or the aggregated data over the nine months of our intervention, suggesting that the framing of prices neither affects short-run nor long-run demand for movie tickets.

4. Price Frames do not Affect the Distribution of Purchases

We have shown in the previous section that our treatments do neither affect the average number of purchases nor the average number of purchased tickets for 3D movies. Nevertheless, shrouding or partitioning of prices may be attractive for sellers if both features affect the distribution of purchases or purchased tickets (and thus allow for exploiting customers' willingness to pay more than in the baseline condition of *Inclusive*). Therefore, we want to take a closer look at the overall distribution – in particular, the variance – of purchases and purchased tickets for 3D movies and test for systematic differences across treatments.

Table 9 presents the first two moments of the empirical distributions of the number of purchases and the number of purchased tickets for 3D movies. We observe that the variance of the number of purchases and purchased tickets is very similar across treatments. Formally, since the distributions of purchases and purchased tickets are highly skewed, we use the robust version of Levene's test proposed by Brown and Forsythe (1974) to test for significant differences in variances. Indeed, only for one out of six pairwise comparisons there is a statistically significant (but still small) difference in variances: the variance in the number of purchased tickets in *Partitioned* is significantly larger than the corresponding variance in *Shrouded* (p-value = 0.023, Robust Levene test). In particular, regarding the number of purchases, there are no significant differences in variances across our three treatments.[16] Altogether, we would argue that there are no economically relevant differences in the first two moments of the empirical distributions of purchases and purchased tickets for 3D movies.[17]

Table 9. *Empirical moments of the distributions of the number of purchases and purchased ticket(s) for 3D movies during our intervention period.*

	# Purchases		**# Tickets**		**# Observations**
	mean	std. dev.	mean	std. dev.	
Inclusive	0.885	1.001	2.213	2.657	11,571
Partitioned	0.870	0.974	2.229	2.829	11,633
Shrouded	0.880	0.974	2.213	2.684	11,698

In order to investigate whether higher moments of these distributions are affected by our treatment manipulations, we take a closer look at the distributions in the different treatments. Figure 4 depicts the empirical cumulative distribution functions (CDFs) of the number of purchases (upper panel) and the number of purchased tickets (lower panel) separately for each treatment. Formally, regarding the number of purchases, we cannot reject the null-hypothesis that the distribution functions in the three treatments are identical (p-value = 0.403, X^2-test).[18] While we can reject the null-hypothesis that the distributions of purchased tickets are identical across treatments (p-value = 0.007, X^2-test),[19] an inspection of Figure 4 suggests that these differences – although, statistically significant – are rather small. Based on these results, we would argue that there are no economically relevant treatment effects on the distributions of purchases and purchased tickets for 3D movies.

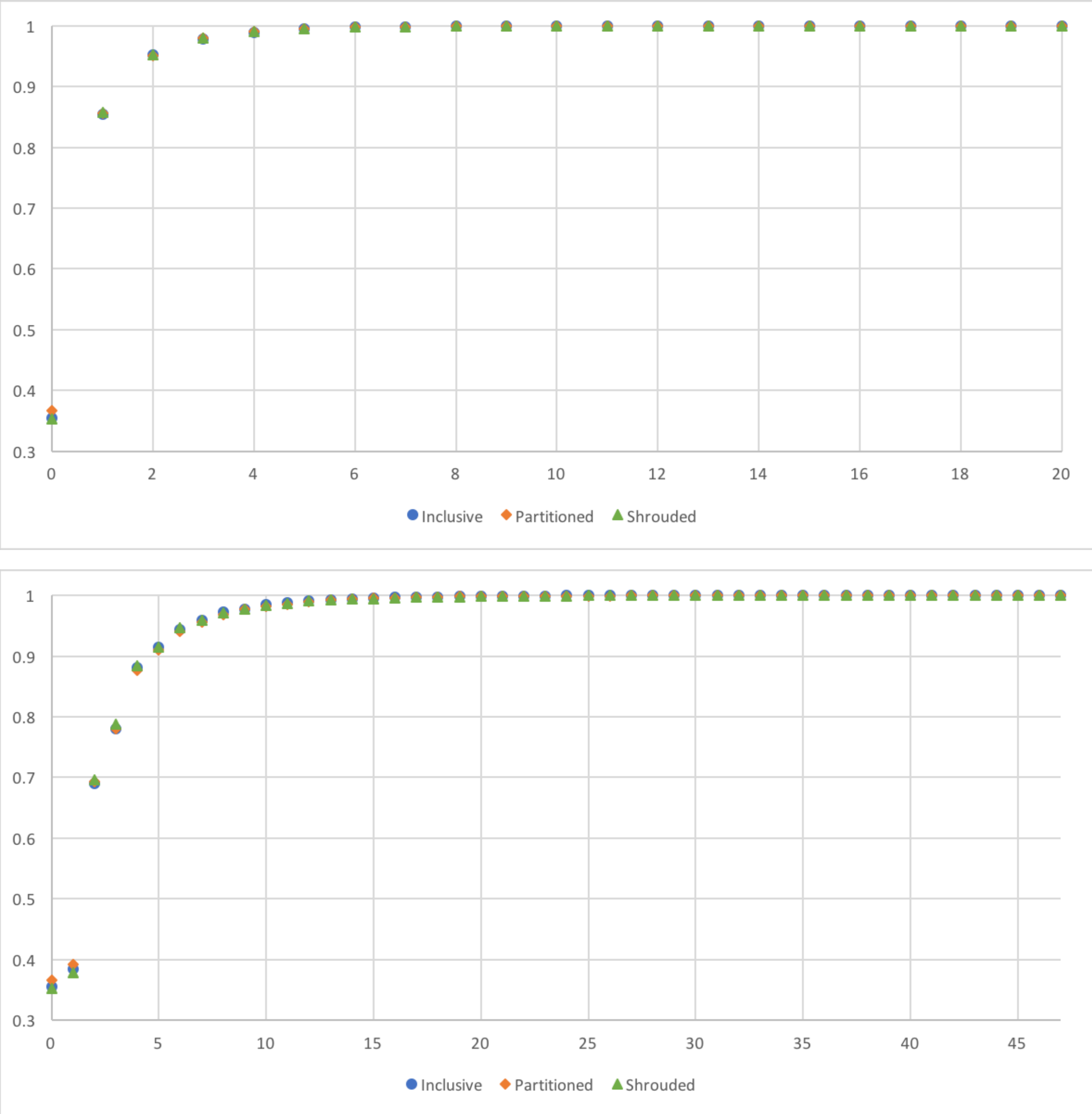
[16] Instead of using a non-parametric test, we could estimate Generalized Negative Binomial models, which allow not only the mean, but also the variance of the data-generating process to *directly* depend on the treatment (Rigby and Stasinopoulos, 2005). Table D.1 in the Appendix shows that the results are basically the same.

[17] Also, if we calculate the averages and variances separately for each month of our intervention period, we do not observe large differences across treatments (see Figure D.1 and Figure D.2 in the Appendix).

[18] We cluster all consumers with at least 5 purchases to increase the power of the test.

[19] We cluster all consumers with at least 15 purchased tickets to increase the power of the test.

Figure 4. *CDFs of the number of purchases and purchased ticket(s) for 3D movies.*



Notes to Figure 4: *The figure in the upper panel depicts the empirical cumulative distribution functions of the number of purchases for 3D movies (x-axis) for each treatment separately. The figure in the lower panel depicts the empirical cumulative distribution functions of the number of purchased tickets for 3D movies (x-axis) for each treatment separately. We truncate the distribution of the number of purchased tickets at 47, while there is one consumer in Partitioned who has actually bought 77 tickets for 3D movies during our intervention period.*

5. Discussion and Conclusion

We present the results of a field experiment with more than 34,000 consumers of a large German cinema that allows us to test for the effects of price salience on online shopping. We investigate the effects of partitioning the overall price and of shrouding certain price attributes, two practices that are frequently applied by companies to increase sales (Heidhues and Kőszegi, 2018). Our experimental design has allowed us to disentangle the effects of price partitioning or shrouding on the likelihood to initiate a purchasing process and the likelihood to actually buy a product. We find that shrouding a 3D surcharge substantially increases the probability that a consumer continues in the process of buying tickets for a 3D movie, compared to a presentation where the surcharge is already included in the displayed price. This shrouding effect is sizeable, as we estimate that around 10% of the consumers neglect shrouded surcharges. Partitioning the overall price into several components has, *ceteris paribus*, at most a small effect on the likelihood to initiate a purchase. However, when we look at the likelihood with which consumers actually buy movie tickets, we find no treatment differences at all, i.e., neither partitioning nor shrouding have a positive effect on purchase rates and thus on the firm's profits.

Our experimental results can be reconciled with the previous literature on shrouding and price partitioning (e.g. Chetty et al., 2009; Finkelstein, 2009; Taubinsky and Rees-Jones, 2018) under the assumptions that the effectiveness of de-biasing consumers by presenting them the full price prior to the purchase depends on the costs of cancelling an already initiated purchase process, and that these cancellation costs are very low just in our experiment. Arguably, in our field experiment the costs of cancelling the purchase on a later screen are, indeed, very low as the whole purchase process requires only three clicks from initiation to completion. Moreover, the anonymity of the online purchasing process avoids any kind of social costs that could arise in offline shopping when a consumer might feel embarrassed in front of the cashier or other consumers when cancelling a purchase at the check-out. The latter social costs might induce a kind of lock-in effect in offline environments that is absent in our setting.

We regard our results as an important complement to the existing empirical literature on salience effects. We show that shrouding and partitioning prices, two strategies often used to obfuscate consumers (e.g., Ellison and Ellison, 2009), can be inadequate instruments to trick consumers into buying more when the costs become fully transparent before completing the purchase *and* when the costs of cancelling an already initiated purchase process are low. If these two conditions are not fulfilled – as in the studies discussed in the introduction – then shrouding and partitioning are likely to affect overall demand. Put differently, while we were able to fully de-bias consumers by presenting the full price on the second screen of the purchase process, learning the full price, for instance, at the cashier in a supermarket as in Chetty et al. (2009) does not make consumers behave as if they had known the full price from the beginning. We suspect that the difference in the effectiveness of de-biasing – presenting the full price offsets all shrouding effects in our experiment, but not in Chetty et al. (2009) – mainly hinges on the difference in cancellation costs across the two setups. In this sense, our experimental findings also provide a rationale for why many companies make it time consuming to complete a purchase process after initiation in online shopping, namely, in order to make it costlier for consumers to drop out (which seems to be common practice for instance for booking flight tickets).

To conclude, our results can also speak to regulators who are interested in how to improve consumer protection on the Internet. In particular, our findings lend support to policy measures that require firms to present the overall price right from the beginning or to keep cancellation costs as low as possible.[20] Both instruments may help protect consumers with limited attention who might otherwise fall prey to price salience effects.

[20] The EU has followed several steps in this direction, in particular by tightening regulations on pricing strategies. For instance, for travel tickets, the European Union Article 23 of Regulation (EC) No 1008/2008 requires that "the final price to be paid shall at all times be indicated and shall include the applicable air fare or air rate as well as all applicable taxes, and charges, surcharges and fees which are unavoidable and foreseeable at the time of publication." An overview of how firms' pricing strategies in general are regulated in the EU can be found, for instance, on https://europa.eu/youreurope/citizens/consumers/shopping/pricing-payments/index_en.htm.

References

Bernheim, Douglas, and Dmitry Taubinsky, "Behavioral Public Economics," in D. Bernheim, S. DellaVigna and D. Laibson (eds.): *Handbook of Behavioral Economics: Applications and Foundations 1*, Vol. 1, Chapter 5, 2018, 381-516.

Birnbaum, Michael, and Juan Navarrete, "Testing Descriptive Utility Theories: Violations of Stochastic Dominance and Cumulative Independence," *Journal of Risk and Uncertainty*, 17, 1998, 49-78.

Bordalo, Pedro, Nicola Gennaioli, and Andrei Shleifer, "Salience and Consumer Choice," *Journal of Political Economy,* 121 (5), 2013, 803-843.

Brown, Jennifer, Tanjim Hossain, and John Morgan, "Shrouded Attributes and Information Suppression: Evidence from the Field," *Quarterly Journal of Economics,* 125 (2), 2010, 859-876.

Brown, Morton, and Alan Forsythe, "Robust Tests for the Equality of Variances," *Journal of the American Statistical Association*, 69 (346), 1974, 364-367.

Carlson, Jay, and Danny Weathers, "Examining Differences in Consumer Reactions to Partitioned Prices with a Variable Number of Price Components," *Journal of Business Research,* 61 (7), 2008, 724-731.

Chetty, Raj, Adam Looney, and Kory Kroft, "Salience and Taxation: Theory and Evidence," *American Economic Review,* 99 (4), 2009, 1145-1177.

Dertwinkel-Kalt, Markus, Holger Gerhardt, Gerhard Riener, Frederik Schwerter, and Louis Strang, "Concentration Bias in Intertemporal Choice," *Working Paper*, 2017.

Dertwinkel-Kalt, Markus, and Mats Köster, "Violations of First-order Stochastic Dominance as Salience Effects," *Journal of Behavioral and Experimental Economics,* 59, 2015, 42-46.

Dertwinkel-Kalt, Markus, and Mats Köster, "Salience and Skewness Preferences," *CESifo Working Paper no. 7416*, 2018.

Dunn, Elizabeth, Timothy Wilson, and Daniel Gilbert, "Location, Location, Location: The Misprediction of Satisfaction in Housing Lotteries," *Personality and Social Psychology Bulletin*, 29 (11), 2003, 1421-1432.

Ellison, Glenn, and Sara Fisher Ellison, "Search, obfuscation, and price elasticities on the internet," *Econometrica*, 77 (2), 2009, 427-452.

Feldman, Naomi E., and Bradley J. Ruffle, "The Impact of Including, Adding, and Subtracting a Tax on Demand," *American Economic Journal: Economic Policy,* 7 (1). 2015, 95-118.

Farhi, Emmanuel, and Xavier Gabaix, "Optimal Taxation with Behavioral Agents," *Working Paper*, 2018.

Finkelstein, Amy, "E-ztax: Tax Salience and Tax Rates," *Quarterly Journal of Economics,* 124 (3), 2009, 969-1010.

Gabaix, Xavier, and David Laibson, "Shrouded Attributes, Consumer Myopia, and Information Suppression in Competitive Markets," *Quarterly Journal of Economics,* 121 (2), 2006, 505-540.

Greenleaf, Eric, Eric Johnson, Vicki Morwitz, and Edith Shalev, "The Price does not Include Additional Taxes, Fees, and Surcharges: A Review of Research on Partitioned Pricing," *Journal of Consumer Psychology* 26 (1), 2016, 105-124.

Heidhues, Paul, and Botond Kőszegi, "Behavioral Industrial Organization," in D. Bernheim, S. DellaVigna and D. Laibson (eds.): *Handbook of Behavioral Economics: Applications and Foundations 1*, Vol. 1, Chapter 6, 2018, 517-612.

Hossain, Tanjim, and John Morgan, "... Plus Shipping and Handling: Revenue (Non) Equivalence in Field Experiments on Ebay," *Advances in Economic Analysis & Policy,* 6 (2), 2006.

Kőszegi, Botond, and Adam Szeidl, "A Model of Focusing in Economic Choice," *Quarterly Journal of Economics,* 128 (1), 2012, 53-104.

List, John A., and Imran Rasul, Field Experiments in Labor Economics. In: O. Ashenfelter, D. Card (eds.) *Handbook of Labor Economics*, 2011, 4A, 103-228. Amsterdam: Elsevier.

Morwitz, Vicki, Eric Greenleaf, and Eric Johnson, "Divide and Prosper: Consumers' Reactions to Partitioned Prices," *Journal of Marketing Research*, 35 (4), 1998, 453-463.

Olden, Andreas, "What Do You Buy When No One's Watching? The Effect of Self-Service Checkouts on the Composition of Sales in Retail," *Working Paper*, 2018.

Rigby, Robert, and Mikis Stasinopoulos, "Generalized Additive Models for Location, Scale and Shape," *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 54 (3), 2005, 507-554.

Rubinstein, Ariel, "Similarity and Decision-Making Under Risk (Is There a Utility Theory Resolution to the Allais Paradox?)," *Journal of Economic Theory*, 46 (1), 1988, 145-153.

Seim, Katja, Maria Vitorino, and David Muir, "Drip Pricing When Consumers Have Limited Foresight: Evidence from Driving School Fees," *Working Paper*, 2016.

Schkade, David, and Daniel Kahneman, "Does Living in California Make People Happy? A Focusing Illusion in Judgments of Life Satisfaction," *Psychological Science,* 9 (5), 1998, 340-346.

Taubinsky, Dmitry, and Alex Rees-Jones. "Attention Variation and Welfare: Theory and Evidence from a Tax Salience Experiment," *Review of Economic Studies*, 2018, forthcoming.

Tversky, Amos, "Intransitivity of Preferences," *Psychological Review*, 76 (1), 1969, 31-48.

Appendix A: A Simple Model on Salience Effects

A.1 Model Setup

Suppose that there is a mass of consumers facing the decision whether to buy a certain product or not; that is, the consumers' choice set includes two options: *buy* and *don't buy*. The product is characterized by several attributes that refer to either quality or price components. Without loss of generality, suppose that there is a single quality attribute, which gives rise to a consumption value of v, and $n \in \mathbb{N}$ price components, $p_k \in \mathbb{R}_+$, adding up to a total price of p.

We further assume that consumers differ with respect to their consumption values; formally, let v be drawn from some continuous probability distribution over $\mathbb{R}_+$. In addition, we assume that a consumer's utility is additively separable and linear in the different price components. Finally, we set the outside option of not buying the product to zero utility.

Building on Kőszegi and Szeidl (2013) and Bordalo et al. (2013), we assume that the consumers assign to each attribute a weight that depends on how much the available options—*buy* and *don't buy*—differ along this choice dimension. More precisely, we adopt the notion in Kőszegi and Szeidl (2013) by assuming that the range of attainable utility along a choice dimension determines the weight assigned to this dimension.

In the following, we extend the model of Kőszegi and Seidel (2013) by allowing certain attributes to be actively hidden from consumers ("shrouded"). This extension helps us to derive predictions for salience effects (through shrouding or partitioning) in our field experiment. But before we can define a consumer's decision rule, we need to specify the space of attributes.

Assumption 1 (Attribute Space)**.** *Each price component p_k represents a different attribute. In addition, the product's consumption value v also represents a distinct attribute.*

Given that we have specified a choice set and a set of attributes, we can define the weight a consumer assigns to each of the different attributes.

Assumption 2 (Focusing Function)**.** *Denote as Δ_k the range of attainable utility along price dimension k. In addition, let s_k be a binary indicator, which takes a value of one if price component p_k is shrouded, and a value of zero otherwise. Then, the weight on price component p_k is given by $g(\Delta_k, s_k)$, whereby the function g satisfies the following properties:*

$$\text{(i) } g(x,0) > g(x,1), \quad \text{and} \quad \text{(ii) } \frac{\partial}{\partial x} g(x,0) > 0, \quad \text{and} \quad \text{(iii) } \frac{\partial}{\partial x} g(x,1) \geq 0.$$

Any function that satisfies these properties, we call a focusing function.

The weight $g(\Delta_k, s_k)$ assigned to a given price component p_k depends on how much the available options differ along this attribute (i.e., $\Delta_k \in \mathbb{R}_+$), and on whether this price component is shrouded (i.e., $s_k = 1$) or not (i.e., $s_k = 0$). Property (i) of the focusing function captures the idea that a consumer pays less attention to hidden price components (see, e.g., Chetty et al., 2009, or Seim et al., 2016, for empirical support). Properties (ii) and (iii) reflect the well-known *contrast effect* (e.g., Schkade and Kahneman, 1998) whereby the attention directed to a choice dimension increases with the difference in attainable utility within this attribute. Given our assumption of linear utility, when deciding whether or not to buy the product, the utility difference across the two options in price dimension k equals $\Delta_k = p_k - 0$, as not buying the product is associated with paying a price of zero, and thus the difference increases in the price p_k. Finally, the weight on the product's consumption value v is defined analogously, with the one exception that we assume the product's consumption value not to be shrouded.[21]

Adapted to our experimental setup in the cinema's online shop, we analyze a purchase process that is structured in two stages: In a first stage, consumers observe the *price frame*, that is, the presentation of the product's price components $(p_k, s_k)_{k=1,\ldots,n}$, and decide whether to initiate the purchase process. In a second stage, consumers are confronted with the product's total price and decide whether to confirm the purchase, whereby we impose the following assumption on second-stage behavior.

Assumption 3 (Cancellation Costs)**.** *If a consumer cancels the purchase process at the second stage, she incurs cancellation costs $c \geq 0$.*

We discuss the potential nature of such cancellation costs in detail below (see Section 2.3).

A.2 Predictions on First- and Second-Stage Behavior

Stage 1. The consumer observes $(p_k, s_k)_{k=1,\ldots,n}$ and decides whether to initiate the purchase process. At this stage, the consumer values the product at

$$g(v, 0) \cdot v - \sum_{k=1}^{n} g(p_k, s_k) \cdot p_k. \tag{1}$$

By Assumptions 1 and 2, both partitioning prices and shrouding certain price components increases the value given in (1) and therefore makes the product to appear more attractive.

[21] Note that a profit-maximizing firm has no incentive to hide attributes that are of positive value to a consumer.

Prediction 1. *Partitioning a price component* p_k *into* p_{k_1} *and* p_{k_2} *with* $p_k = p_{k_1} + p_{k_2}$ *strictly increases the share of consumers initiating the purchase process.*

Prediction 2. *For any fixed number of price components, shrouding one of these price components strictly increases the share of consumers initiating the purchase process.*

Stage 2. Given that a consumer has initiated the purchase process, she is confronted with the total price p and decides whether to confirm the purchase or not. At this stage, the consumer values the product at

$$g(v,0)\cdot v - g(p,0)\cdot p \leq g(v,0)\cdot v - \sum_{k=1}^{n} g(p_k, s_k)\cdot p_k,$$

whereby the inequality is strict if either $n \geq 2$ or $s_k = 1$ for at least some price dimension k.

Since we assume that cancelling the purchase process goes along with some cancellation costs $c \geq 0$, a consumer actually confirms the purchase if and only if

$$g(v,0)\cdot v - g(p,0)\cdot p \geq -c. \quad (2)$$

Depending on the magnitude of cancellation costs, a consumer might be prone to cancel the purchase process, once she is confronted with the full price.

Prediction 3. *There is a weakly positive demand effect of partitioning prices or shrouding a price component that increases with cancellation costs* c*, but vanishes as* c *approaches zero.*

A.3 Discussion of Modelling Assumptions

We conclude the model section with a discussion of our central assumptions on the choice set, the attribute space, the contrast effect and shrouding, and the cancellation costs.

Choice Set. First, we assume that consumers decide whether or not to buy a certain product, but that there are no substitute products available. As we argue in Appendix A, however, our formal predictions above do not change if there are also substitute products available with price components $p'_k \leq p_k$ for any $k \in \{1, \ldots, n\}$. In the context of our experiment, natural substitute products, which satisfy this condition, are the corresponding 2D show to a given 3D movie or the same 3D movie running at broadly the same time in a different cinema.

Attribute Space. Second, we assume that each price component defines a different attribute (Assumption 1). While this assumption is crucial for parts of our predictions, the economic literature has not taken a clear stance on whether partitioning the total price into several price components indeed affects how consumers evaluate the total price.

On the one hand, Morwitz et al. (1998) find that partitioning an option's total price makes consumers reason as if this option has several smaller price attributes, thereby underweighting the total costs of the option. Relatedly, Birnbaum and Navarette (1998) observe that artificially increasing the state space in choices under risk by splitting up a lottery's payoffs affects behavior in a way that is consistent with assuming that the perceived state space becomes larger (see Dertwinkel-Kalt and Köster, 2015, for a formalization of this interpretation).

On the other hand, in an experimental study on intertemporal choice, Dertwinkel-Kalt et al. (2017) find that simply splitting up payments to be received within one day has little effect on behavior. If payments are dispersed over time (i.e., across several days), however, subjects in their experiment behave as if the number of attributes has increased, which suggests that the dates at which choice alternatives yield payments (and not necessarily the payments themselves) represent different choice dimensions.

We follow the first strand of the literature and derive our predictions under the assumption that different price components represent different attributes. If this was not the case, then Prediction 1 would no longer hold. In contrast, Prediction 2 as well as the respective part of Prediction 3 are robust to assuming that there is only a single price attribute.

Contrast Effect and Shrouding. Third, following Kőszegi and Szeidl (2013), we assume that a consumer's decision weight assigned to a given attribute is determined by a so-called focusing function. This entails two assumptions. On the one hand, a consumer attaches a lower weight to a certain attribute if this attribute is shrouded (Property (i) of Assumption 2). On the other hand, the larger the contrast in attainable utility along an attribute is the larger is the decision weight assigned to this attribute (Properties (ii) and (iii) of Assumption 2).

The fact that shrouding of price components affects behavior of at least some consumers is well-established in the behavioral economics literature (see Chetty et al., 2009, for empirical support, and Heidhues and Kőszegi, 2018, for an overview of the theoretical literature deriving implications for firms' pricing strategies). While we assume that *all* consumers neglect hidden prices, one may argue that only a few consumers neglect shrouded price components while others perfectly account also for hidden fees – an idea first modeled by Gabaix and Laibson (2006). As long as the consumers who neglect shrouded price components have on average a weakly lower valuation for the product than those consumers who are attentive also to hidden fees, this alternative assumption yields the same qualitative predictions on average behavior, since shrouding of price components strictly increases the probability that inattentive, low-value consumers initiate the purchase process. Thus, such an alternative approach is not distinguishable from our model based on aggregated data.

The contrast effect is the central ingredient of recent models on attentional focusing by Kőszegi and Szeidl (2013) and Bordalo et al. (2013b), but the underlying idea that large contrasts attract attention has been modeled already by Tversky (1969) and Rubinstein (1988). The contrast effect is also in line with various empirical observations (e.g., Schkade and Kahneman, 1998; Dunn et al., 2003) and has been supported by recent lab experiments (e.g., Dertwinkel-Kalt et al., 2017; Dertwinkel-Kalt and Köster, 2018). For instance, Dertwinkel-Kalt et al. (2017) find in a laboratory experiment on intertemporal choice that subjects prefer a single concentrated payment over several smaller payments that are dispersed over time.

Cancellation Costs. Fourth, we assume that consumers incur non-negative cancellation costs when cancelling the purchase process at the second stage (Assumption 3). These costs are likely to be of a psychological nature: for instance, a consumer may feel bad to cancel a purchase after having invested effort into buying in the first place (i.e., sunk cost fallacy), or if the consumer's decision is observed by others (e.g., at the cashier in the supermarket), it may feel awkward to cancel the purchase (i.e., a cancellation imposes social costs). In the context of grocery shopping, Olden (2018) finds that shopping behavior crucially depends on how observable a purchase is to others, suggesting that observability comes along with a social cost. Arguably, these costs are smaller online than offline – social costs can hardly play a role online – and decrease with the amount of time spent to work through the purchase process (as sunk costs are smaller when a consumer needs only a few clicks to work through the purchase process).

A.4 Formal Predictions for Larger Choice Sets

We expand the choice context as follows. Suppose that the consumer chooses from the set

$$\mathcal{C} = \{(v^1, p_1^1, \dots, p_n^1), (v^2, p_1^2, \dots, p_n^2), (0, 0, \dots, 0)\},$$

where $0 < v^2 \leq v^1$ and $0 \leq p_k^2 \leq p_k^1$ for any price component k. In addition, denote the products' total prices as p^1 and p^2, respectively. As before, the ranges of attainable utilities in the different attributes are given by $\Delta_v = v^1 - 0$, $\Delta_k = p_k^1 - 0$, and $\Delta_p = p^1 - 0$, respectively; that is, the presence of the second product does not change decision weights. For simplicity, let $s_k^i = s_k$ (i.e., price component k is shrouded either for both or no product).

1. Stage. For the sake of argument, consider Product 1. The consumer is confronted with the price frame $(p_k^i, s_k)_{k=1,\dots,n;\, i=1,2}$ and decides whether to initiate the purchase process for

Product 1. Since the decision weights are not affected by the presence of the second product, the consumer initiates a purchase process for at least one product if the value in (1) is positive. In addition, at this stage, the difference in valuations for the two products is given by

$$g(\Delta_v, 0) \cdot (v^1 - v^2) - \sum_{k=1}^{n} g(\Delta_k, s_k) \cdot (p_k^1 - p_k^2).$$

Thus, Assumptions 1 and 2 immediately imply that both splitting up a price component p_k^1 or shrouding certain price components makes Product 1 to appear more attractive both relative to Product 2 and relative to the outside option of not buying. In this sense, Predictions 1 and 2 are robust to extending the choice set by a low-price option (e.g., a 2D variant of a 3D movie).

2. Stage. Given that the consumer has initiated the purchase process for Product 1, she is confronted with the total price p^1 and decides whether to actually purchase the product or not. At this stage, the difference in valuations for the two products is equal to

$$\begin{aligned} &g(\Delta_v, 0) \cdot (v^1 - v^2) - g(\Delta_p, 0) \cdot (p^1 - p^2) \\ &\qquad < g(\Delta_v, 0) \cdot (v^1 - v^2) - \sum_{k=1}^{n} g(\Delta_k, s_k) \cdot (p_k^1 - p_k^2) \end{aligned}$$

for all $n > 2$. Accordingly, Product 1 becomes less attractive relative to Product 2 and relative to the outside option, once its total price is presented. In this sense, also Prediction 3 is robust to extending the choice set by a substitute product with weakly lower prices.

Appendix B: Testing for Overdispersion

Table B.1 presents conditional means and conditional variances of the number of drop-outs on the initial screen across the different treatments. For each of the three treatments, conditional variances largely exceed conditional means. Also in formal tests, we can reject the null-hypotheses of mean-variance equivalence against the alternatives of overdispersion (all p-values are smaller than 0.01). Given these patterns, a Negative Binomial model is more appropriate than a Poisson model.

Table B.1. *Conditional means and variances of the number of drop-outs for 3D movies.*

	# Drop-outs		**# Observations**
	mean	variance	
Inclusive	1.747	18.241	11,571
Partitioned	1.635	7.322	11,633
Shrouded	1.569	9.285	11,698

Table B.2 presents conditional means and variances of the number of purchases and purchased tickets for 3D movies, respectively, across the different treatments. For each of the three treatments, the conditional variances largely exceed conditional means. Also in formal tests, we can reject the null-hypotheses of mean-variance equivalence against the alternatives of overdispersion (all p-values are smaller than 0.01). Given these patterns, a Negative Binomial model is more appropriate than a Poisson model.

Table B.2. *Conditional means and variances of the number of purchases and the number of purchased ticket(s) for 3D movies.*

	# Purchases		**# Tickets**		**# Observations**
	mean	variance	mean	variance	
Inclusive	0.885	1.002	2.213	7.058	11,571
Partitioned	0.870	0.948	2.229	8.006	11,633
Shrouded	0.880	0.948	2.213	7.204	11,698

Appendix C: Additional Results and Robustness

Part 1: First Clicks. Table C.1 presents the results of regressing a binary indicator of whether a consumer drops out at the initial screen, conditional on clicking on a 3D movie for the first time, on treatment indicators interacted with indicators of whether a 2D or 3D show of the same movie runs in the same or another multiplex (in the same city) at broadly the same time.[22]

[22] Notice that the number of observations is reduced compared to our baseline regressions, as we do not have information on the schedules of other cinemas in the same city for each day of our intervention period.

Table C.1. *Likelihood to drop out on initial screen depending on substitute products.*

Paramater	**Drop-out**	**Drop-out**
Partitioned	-0.005	-0.005
	(0.012)	(0.012)
Shrouded	-0.056***	-0.053***
	(0.012)	(0.012)
2D Substitute	0.062***	0.013
	(0.011)	(0.011)
3D Substitute	0.079***	0.009
	(0.010)	(0.011)
2D Sub x Partitioned	-0.004	-0.006
	(0.015)	(0.015)
2D Sub x Shrouded	0.006	0.002
	(0.015)	(0.015)
3D Sub x Partitioned	0.005	0.006
	(0.014)	(0.014)
3D Sub x Shrouded	0.003	0.000
	(0.014)	(0.014)
Noon	-	-0.025**
		(0.010)
Afternoon	-	-0.023***
		(0.008)
Night	-	0.081***
		(0.009)
Movie FE	no	yes
Time FE	no	yes
# Observations	31,101	31,101

Notes to Table C.1: *Results of an OLS-regression with a binary indicator of whether a consumer drops out on the initial screen conditional on clicking on a 3D movie for the first time as dependent variable and treatment indicators interacted with indicators of available substitute shows as independent variables (whereby Inclusive and no substitute serves as base category). In the second column, we include movie, year, month, and day of the week fixed effects as well as a control for the time of the day at which the show runs (whereby Evening serves as base category). Standard errors are provided in parenthesis. Significance level: *: 10%, **: 5%, ***: 1%.*

Table C.2 presents the corresponding results for the probability to purchase ticket(s) for the 3D show that a consumer clicked on first during our intervention period. The treatment effects do not depend on whether a substitute is available or not.

Table C.2. *Likelihood to purchase tickets depending on substitute products.*

Paramater	**Purchase**	**Purchase**
Partitioned	0.001	0.003
	(0.012)	(0.012)
Shrouded	-0.002	-0.005
	(0.012)	(0.012)
2D Substitute	-0.066***	-0.039***
	(0.011)	(0.011)
3D Substitute	-0.044***	-0.000
	(0.010)	(0.012)
2D Sub x Partitioned	-0.016	-0.014
	(0.015)	(0.015)
2D Sub x Shrouded	-0.003	0.001
	(0.015)	(0.015)
3D Sub x Partitioned	-0.010	-0.012
	(0.015)	(0.014)
3D Sub x Shrouded	0.006	0.009
	(0.015)	(0.014)
Noon	-	0.022**
		(0.011)
Afternoon	-	0.008
		(0.008)
Night	-	-0.086***
		(0.009)
Movie FE	no	yes
Time FE	no	yes
# Observations	31,101	31,101

Notes to Table C.2: *Results of an OLS-regression with a binary indicator of whether a consumer buys ticket(s) for the first 3D show that she clicked on during our intervention as dependent variable and treatment indicators interacted with indicators of available substitute shows as independent variables (whereby Inclusive and no substitute serves as base category). In the second column, we include movie, year, month, and day of the week fixed effects as well as a control for the time of the day at which the show runs (whereby Evening serves as base category). Standard errors are provided in parenthesis. Significance level: *: 10%, **: 5%, ***: 1%.*

Table C.3 presents the results of regressing a binary indicator of whether a consumer drops out at the initial screen, conditional on clicking on a 3D movie for the first time, on treatment indicators interacted with an indicator of blockbuster movies. We classified a movie as a blockbuster if it belongs to the top 25% of movies in our sample in terms of worldwide revenue (revenue data is collected from http://www.boxofficemojo.com, accessed on July, 18 2018).

Table C.3. *Likelihood to drop out on initial screen for blockbuster movies.*

Paramater	**Drop-out**	**Drop-out**
Partitioned	-0.004	-0.004
	(0.011)	(0.011)
Shrouded	-0.057***	-0.056***
	(0.011)	(0.011)
Blockbuster	-0.085***	-0.159
	(0.010)	(0.134)
Blockbuster x Partitioned	-0.004	-0.004
	(0.014)	(0.013)
Blockbuster x Shrouded	0.007	0.006
	(0.014)	(0.013)
2D Substitute	-	0.009
		(0.007)
Noon	-	-0.022**
		(0.010)
Afternoon	-	-0.026***
		(0.007)
Night	-	0.072***
		(0.009)
Movie FE	no	Yes
Time FE	no	yes
# Observations	34,902	34,902

Notes to Table C.3: *Results of an OLS-regression with a binary indicator of whether a consumer drops out on the initial screen conditional on clicking on a 3D movie for the first time as dependent variable and treatment indicators interacted with an indicator of blockbuster movies as independent variables (whereby Inclusive and no blockbuster serves as base category). In the second column, we include movie, year, month, and day of the week fixed effects as well as controls for whether a 2D show of the same movie is running within +/- 1 hour and for the time of the day at which the show runs (whereby Evening serves as base category). Standard errors are provided in parenthesis. Significance level: *: 10%, **: 5%, ***: 1%.*

Table C.4 presents the corresponding results for the probability to purchase ticket(s) for the 3D show that a consumer clicked on first during our intervention period.

Table C.4. *Likelihood to purchase tickets for blockbuster movies.*

Paramater	**Purchase**	**Purchase**
Partitioned	-0.001	-0.001
	(0.011)	(0.011)
Shrouded	0.007	0.006
	(0.011)	(0.011)
Blockbuster	-0.038***	0.033
	(0.010)	(0.135)
Blockbuster x Partitioned	-0.009	-0.009
	(0.014)	(0.014)
Blockbuster x Shrouded	-0.013	-0.011
	(0.014)	(0.013)
2D Substitute	-	-0.041***
		(0.007)
Noon	-	0.019*
		(0.010)
Afternoon	-	0.015**
		(0.007)
Night	-	-0.083***
		(0.009)
Movie FE	no	yes
Time FE	no	yes
# Observations	34,902	34,902

Notes to Table C.4: *Results of an OLS-regression with a binary indicator of whether a consumer buys ticket(s) for the first 3D show that she clicked on during our intervention as dependent variable and treatment indicators interacted with an indicator of blockbuster movies as independent variables (whereby Inclusive and no blockbuster serves as base category). In the second column, we include movie, year, month, and day of the week fixed effects as well as controls for whether a 2D show of the same movie is running within +/- 1 hour and for the time of the day at which the show runs (whereby Evening serves as base category). Standard errors are provided in parenthesis. Significance level: *: 10%, **: 5%, ***: 1%.*

Again, we do not find differential treatment effects. In addition, we observe that neither the average probability to drop out at the initial screen nor the average probability to buy differs for blockbuster movies, once we control for observable characteristics. Our qualitative results are robust to choosing a different classification threshold for blockbuster movies, for instance, the top 10% or top 50% of movies in our sample in terms of worldwide revenue.

Table C.5 presents the results of regressing a binary indicator of whether a consumer drops out at the initial screen, conditional on clicking on a 3D movie for the first time, on treatment indicators interacted with an indicator of shows running at the weekend (i.e., shows running on Friday, Saturday, or Sunday).

Table C.5. *Likelihood to drop out on initial screen at weekends.*

Paramater	**Drop-out**	**Drop-out**
Partitioned	-0.007	-0.007
	(0.009)	(0.009)
Shrouded	-0.051***	-0.051***
	(0.009)	(0.009)
Weekend	-0.033***	0.001
	(0.009)	(0.013)
Weekend x Partitioned	0.002	0.001
	(0.013)	(0.013)
Weekend x Shrouded	-0.002	-0.002
	(0.013)	(0.013)
2D Substitute	-	0.009
		(0.007)
Noon	-	-0.022**
		(0.010)
Afternoon	-	-0.026***
		(0.007)
Night	-	0.072***
		(0.009)
Movie FE	no	yes
Time FE	no	yes
# Observations	34,902	34,902

Notes to Table C.5: *Results of an OLS-regression with a binary indicator of whether a consumer drops out on the initial screen conditional on clicking on a 3D movie for the first time as dependent variable and treatment indicators interacted with an indicator of weekends as independent variables (whereby Inclusive and no weekend serves as base category). In the second column, we include movie, year, month, and day of the week fixed effects as well as controls for whether a 2D show of the same movie is running within +/- 1 hour and for the time of the day at which the show runs (whereby Evening serves as base category). Standard errors are provided in parenthesis. Significance level: *: 10%, **: 5%, ***: 1%.*

Table C.6 presents the corresponding results for the probability to purchase ticket(s) for the 3D show that a consumer clicked on first during our intervention period.

Table C.6. *Likelihood to purchase tickets at weekends.*

Paramater	**Purchase**	**Purchase**
Partitioned	-0.009	-0.010
	(0.009)	(0.009)
Shrouded	0.001	0.001
	(0.009)	(0.009)
Weekend	0.015	-0.005
	(0.009)	(0.014)
Weekend x Partitioned	0.005	0.007
	(0.013)	(0.013)
Weekend x Shrouded	-0.004	-0.003
	(0.013)	(0.013)
2D Substitute	-	-0.041***
		(0.007)
Noon	-	0.019*
		(0.010)
Afternoon	-	0.015**
		(0.007)
Night	-	-0.083***
		(0.009)
Movie FE	no	yes
Time FE	no	yes
# Observations	34,902	34,902

Notes to Table C.6: *Results of an OLS-regression with a binary indicator of whether a consumer buys ticket(s) for the first 3D show that she clicked on during our intervention as dependent variable and treatment indicators interacted with an indicator of weekends as independent variables (whereby Inclusive and no weekend serves as base category). In the second column, we include movie, year, month, and day of the week fixed effects as well as controls for whether a 2D show of the same movie is running within +/- 1 hour and for the time of the day at which the show runs (whereby Evening serves as base category). Standard errors are provided in parenthesis. Significance level: *: 10%, **: 5%, ***: 1%.*

Altogether, we do not find differential treatment effects for movie shows running at the weekend. In addition, we observe that neither the average probability to drop out at the initial screen nor the average probability to buy differs for movie shows running at the weekend.

Part 2: Aggregate Statistics. Finally, we want to analyze treatment effects on the overall number of purchases and the overall number of purchased ticket(s) for either 3D and 2D movies together or only 2D movies. Again, we need to determine first which count model fits our data best. Table C.7 gives an overview of conditional means and variances across treatments.

Table C.7. *Conditional means and variances of the number of purchases and the number of purchased ticket(s) for all movies and for 2D movies only.*

Panel A:	**# Purchases**		**# Tickets**		**# Observations**
All movies	mean	variance	mean	variance	
Inclusive	1.532	3.459	3.815	21.305	11,571
Partitioned	1.516	3.725	3.857	29.386	11,633
Shrouded	1.538	3.180	3.868	21.844	11,698

Panel B:	**# Purchases**		**# Tickets**		**# Observations**
2D movies	mean	variance	mean	variance	
Inclusive	0.648	1.706	1.602	10.529	11,571
Partitioned	0.646	2.079	1.628	15.031	11,633
Shrouded	0.658	1.604	1.654	11.126	11,698

We observe that on both samples for all treatments the conditional variances are larger than the conditional means. Also in formal tests, we can reject the null-hypotheses of mean-variance equivalence against the alternatives of overdispersion (all p-values are smaller than 0.01). This suggests that Negative Binomial models are a more appropriate choice than Poisson models. Hence, we will estimate Negative Binomial models with either the number of purchases or the number of purchased ticket(s) as the dependent variable. Table C.8 reports the corresponding regression results for all movies and Table C.9 reports the results for 2D movies.[23]

[23] Since half of the consumers have never clicked on a 2D show during our intervention, we cannot estimate a model with exposure when using the subsample of observations that refer to 2D movies.

Table C.8. *Average number of purchases or bought tickets over entire intervention period.*

Parameter	**# Purchases**	**# Purchases**	**# Tickets**	**# Tickets**
Partitioned	-0.011	0.001	0.011	0.018
	(0.014)	(0.013)	(0.015)	(0.014)
Shrouded	0.004	0.001	0.014	0.008
	(0.014)	(0.013)	(0.015)	(0.014)
Exposure	no	yes	no	yes
# Observations	34,902	34,902	34,902	34,902

Notes to Table C.8: *Results of regressing the number of purchases and the number of purchased ticket(s), respectively, on treatment indicators (whereby Inclusive serves as base category), using a Negative Binomial model. Standard errors are provided in parenthesis. Significance level: *: 10%, **: 5%, ***: 1%.*

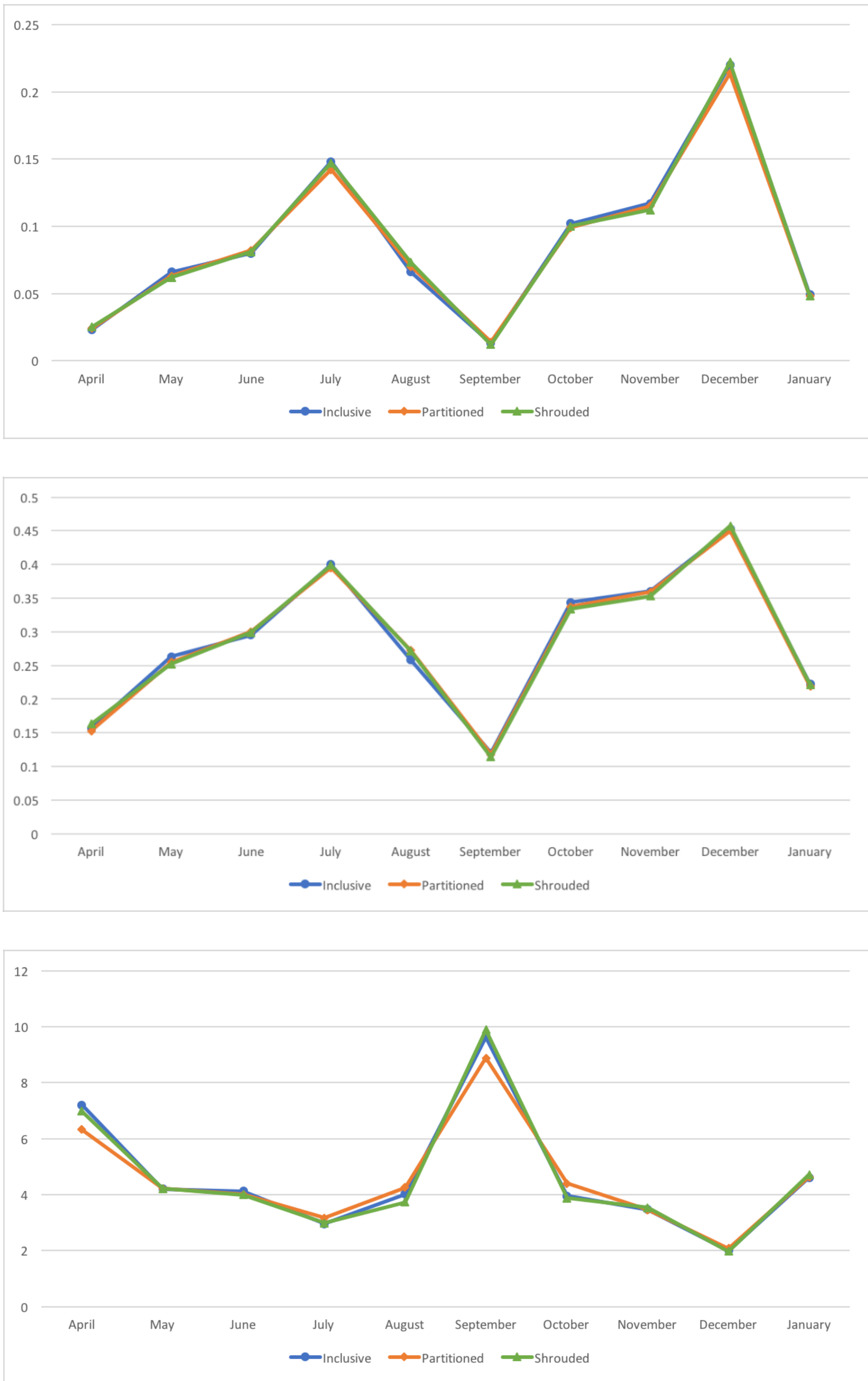
Table C.9. *Average number of purchases or bought tickets for 2D over intervention period.*

Parameter	**# Purchases**	**# Tickets**
Partitioned	0.005	0.016
	(0.020)	(0.030)
Shrouded	0.012	0.032
	(0.020)	(0.030)
Exposure	no	no
# Observations	34,902	34,902

Notes to Table C.9: *Results of regressing the number of 2D purchases and the number of purchased ticket(s) for 2D movies on treatment indicators (with Inclusive as base category), using a Negative Binomial model. We cannot estimate a model with exposure as some consumers have never clicked on a 2D show. Standard errors are provided in parenthesis. Significance level: *: 10%, **: 5%, ***: 1%.*

Appendix D: Heterogeneous Treatment Effects

Figure D.1. *Empirical moments of the number of purchases for 3D movies.*



Notes to Figure D.1: *The figure in the upper panel depicts the average number of purchases for 3D movies in each month for each treatment separately. The figure in the middle panel depicts the variance in the number of purchases for 3D movies in each month for each treatment separately. The figure in the lower panel depicts the skewness of the number of purchases for 3D movies in each month for each treatment separately.*

Figure D.2. *Empirical moments of the number of purchased tickets for 3D movies.*

Notes to Figure D.2: *The figure in the upper panel depicts the average number of purchased tickets for 3D movies in each month for each treatment separately. The figure in the middle panel depicts the variance in the number of purchased tickets for 3D movies in each month for each treatment separately. The figure in the lower panel depicts the skewness of the number of purchased tickets for 3D movies in each month for each treatment separately.*

For a given consumer i, we denote as y_i the variable of interest (i.e., the number of purchases or the number of purchased tickets for 3D movies), and let T_i be her treatment. Recall that the Negative Binomial models estimated in Section 4 assume that $\mathbb{E}[y_i|\mu_i, \alpha] = \mu_i$ and $Var[y_i|\mu_i, \alpha] = \mu_i + \alpha\mu_i^2$, where $\mu_i = \exp(\beta T_i)$ and variance-parameter $\alpha > 0$ is a constant. In contrast, a Generalized Negative Binomial model further allows the variance-parameter to vary across treatments (Rigby and Stasinopoulos, 2005, p. 529): specifically, we assume that $\alpha_i = \exp(\gamma T_i)$. Table D.1 reports the corresponding regression results.

Table D.1. *Mean and variance of the number of purchases and purchased tickets for 3D movies over our intervention period.*

	# Purchases	**# Purchases**	**# Tickets**	**# Tickets**
Mean				
Partitioned	-0.017	0.011	0.007	0.031
	(0.014)	(0.019)	(0.016)	(0.019)
Shrouded	-0.005	0.002	0.000	0.009
	(0.014)	(0.019)	(0.016)	(0.019)
Variance				
Partitioned	-0.201	0.066	0.099***	0.066**
	(0.293)	(0.071)	(0.032)	(0.029)
Shrouded	-0.463	0.062	-0.010	0.020
	(0.332)	(0.070)	(0.033)	(0.028)
Exposure	no	yes	no	yes
# Observations	34,902	34,902	34,902	34,902

Notes to Table D.1: *Results of regressing the number of 3D purchases and the number of purchased ticket(s) for 3D movies on treatment indicators (with Inclusive as base category), using a Generalized Negative Binomial model. Standard errors are provided in parenthesis. Significance level: *: 10%, **: 5%, ***: 1%.*

As before, we do not observe any statistically significant differences in the average number of purchases or purchased tickets for 3D movies. In line with the non-parametric tests presented in Section 5, however, we observe that the variance in the number of purchased tickets is significantly larger in *Partitioned* compared to *Inclusive* and *Shrouded*. But, although statistically significant, these differences in variances are rather small.

Appendix E: Decision Screens in the Different Treatments

Figure E.1. *Price presentation on the initial screen in Inclusive.*

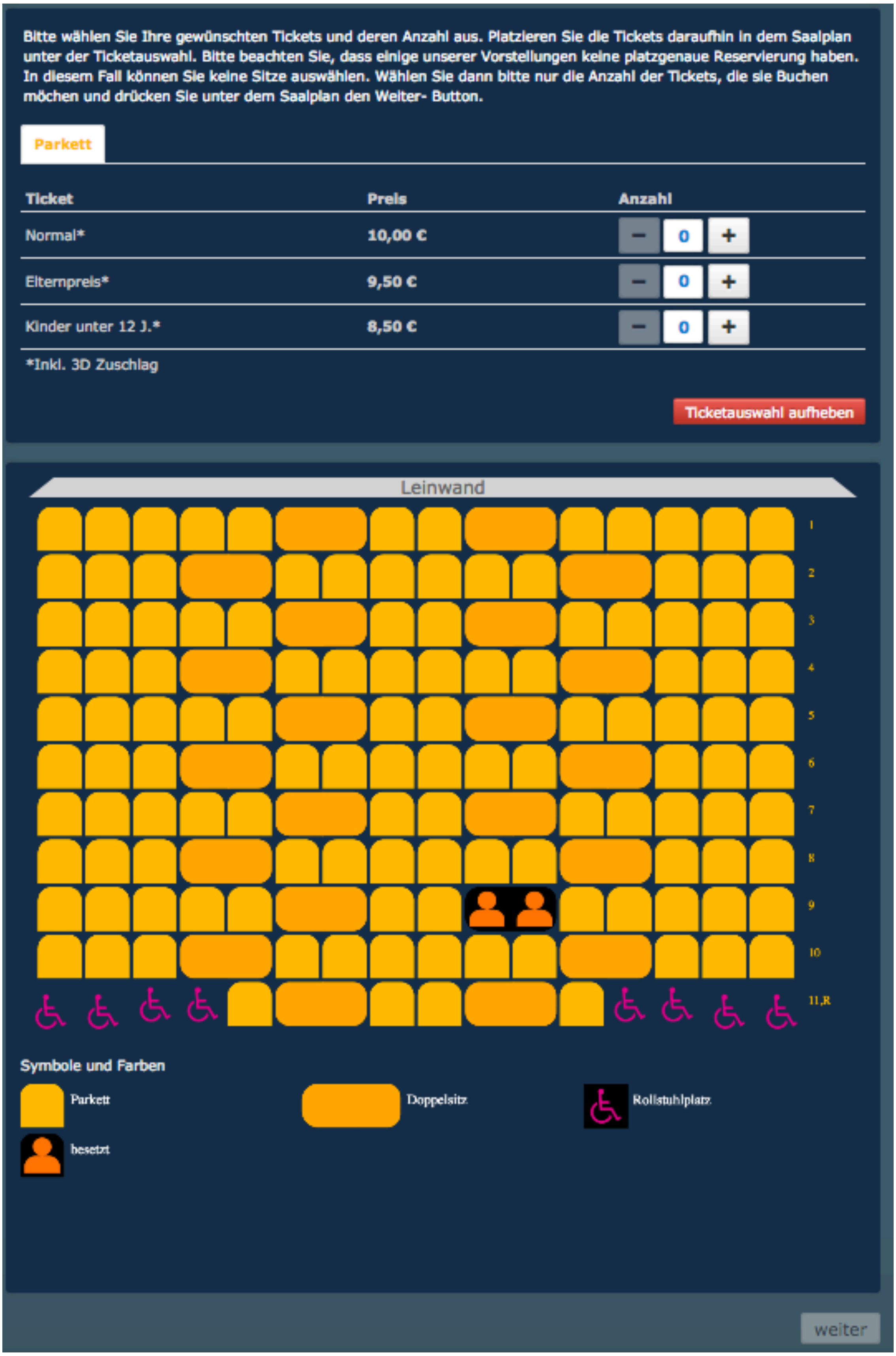


Figure E.2. *Price presentation on the initial screen in Partitioned.*

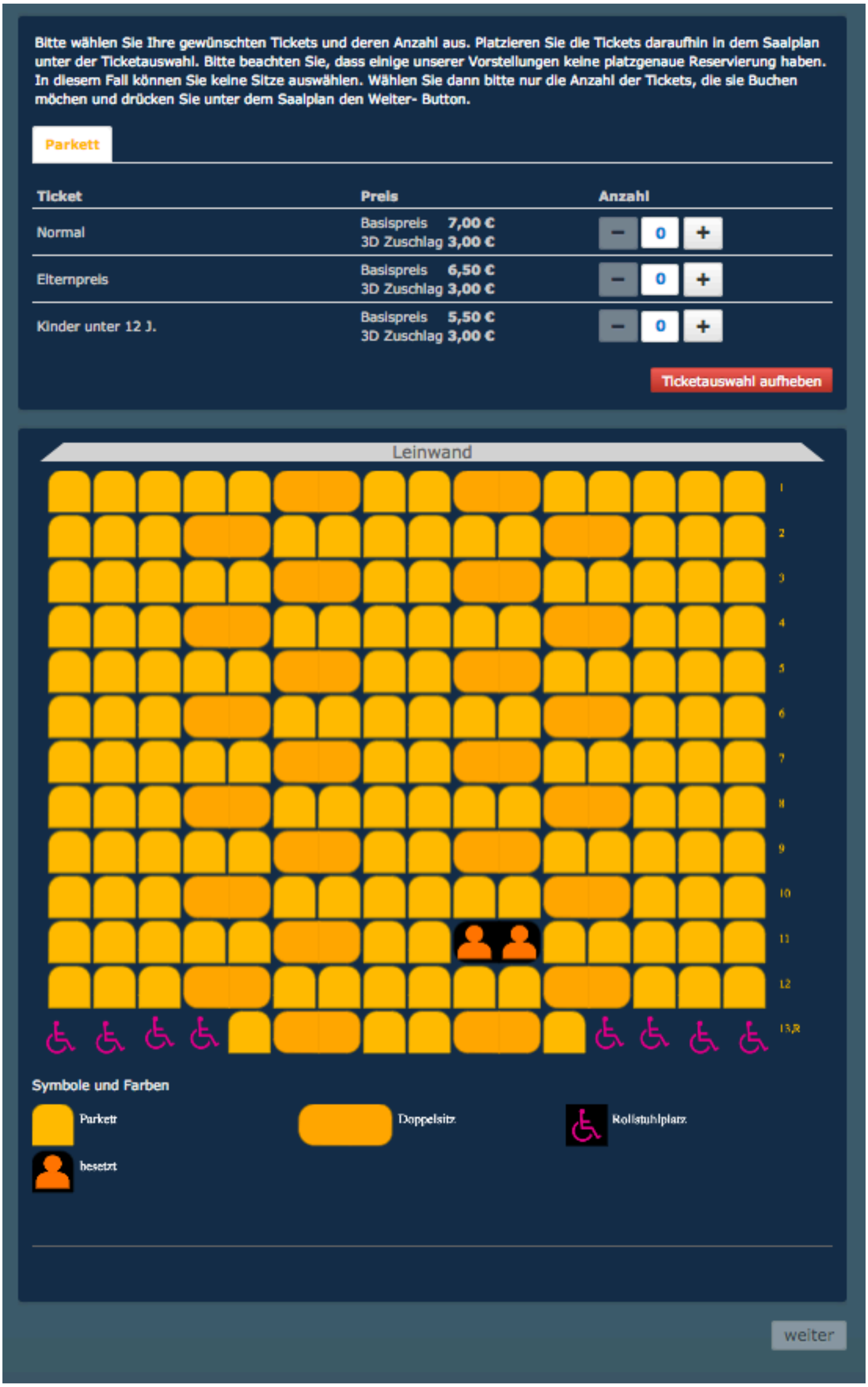
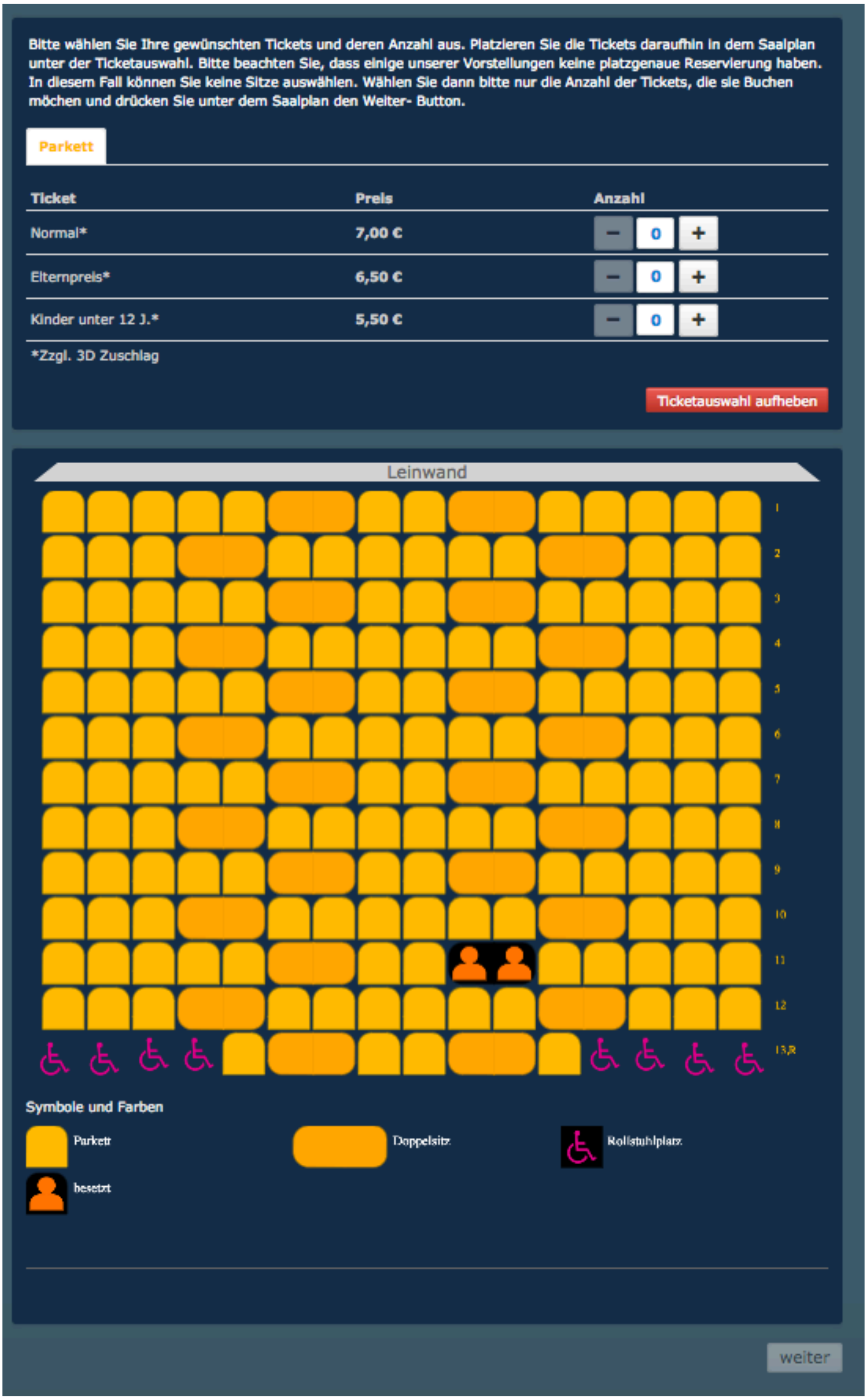


Figure E.3. *Price presentation on the initial screen in Exclusive.*



Appendix F: Selection Issues When Comparing Drop-Out Rates

In order to analyze treatment effects on the overall number of clicks on 3D movies, we need to determine first which count model fits our data best.

Table F.1. *Conditional means and variances of the number of clicks on 3D movies.*

	# Clicks 3D		**# Observations**
	mean	variance	
Inclusive	2.904	22.017	11,571
Partitioned	2.808	10.195	11,633
Shrouded	2.909	13.340	11,698

Table F.1 shows that for all treatments the conditional variances are much larger than the conditional means. Also in formal tests, we can reject the null-hypotheses of mean-variance equivalence against the alternatives of overdispersion. This suggests that a Negative Binomial model is a more appropriate choice than a Poisson model.

Hence, given the structure of our data, we estimate a Negative Binomial model with the number of clicks on 3D movies as dependent variable and treatment indicators as independent variables. Table F.2 reports the corresponding regression results.

Table F.2. *Average number of clicks on 3D movies.*

Parameter	**# Clicks 3D**
Partitioned	-0.033***
	(0.012)
Shrouded	0.002
	(0.012)
# Observations	34,902

Notes to Table F.2: *Results of regressing the number clicks on 3D movies on treatment indicators (whereby Inclusive serves as base category), using a Negative Binomial model. Standard errors are provided in parenthesis. Significance level: *: 10%, **: 5%, ***: 1%.*

We observe that, relative to *Inclusive*, consumers in *Partitioned* click significantly less often on 3D movies during the treatment period. This implies that there is a selection problem, and that using the overall number of drop-outs at the first screen in the process of buying ticket(s) for a 3D movie as a dependent variable may result in biased estimates of the ATEs.

Finally, we ask whether our treatments affect the average number of *different* 3D movies a consumer clicks on during the treatment period. Again, in order to answer this question, we need to determine first which count model fits our data best. As before, we look at the conditional means and variances across treatments (see Table F.3).

Table F.3. *Conditional means and variances of number of clicks on different 3D movies.*

	# Clicks Different 3D		**# Observations**
	mean	variance	
Inclusive	1.997	6.038	11,571
Partitioned	1.945	3.778	11,633
Shrouded	1.935	3.547	11,698

We observe that for all treatments the conditional variances are much larger than the conditional means. Also in formal tests, we can reject the null-hypotheses of mean-variance equivalence against the alternatives of overdispersion. This suggests that a Negative Binomial model is a more appropriate choice than a Poisson model.

Given the structure of our data, we estimate a Negative Binomial model with the number of clicks on different 3D movies as dependent variable and treatment indicators as independent variables. Table F.4 reports the corresponding regression results.

Table F.4. *Average number of clicks on different 3D movies.*

Parameter	**# Clicks Different 3D**
Partitioned	-0.026**
	(0.011)
Shrouded	-0.032***
	(0.011)
# Observations	34,902

Notes to Table F.4: *Results of regressing the number of clicks on different 3D movies on treatment indicators (whereby Inclusive serves as base category), using a Negative Binomial model. Standard errors are provided in parenthesis. Significance level: *: 10%, **: 5%, ***: 1%.*

We observe that, relative to *Inclusive*, both consumers in *Partitioned* and consumers in *Shrouded* click on significantly less different 3D movies on average. Hence, selection was still an issue if we would use for each consumer-film combination only the first recorded click during the treatment period.