# DISCUSSION PAPER SERIES

DP12948
(v. 4)

## Advertising as a reminder: Evidence from the Dutch State Lottery

Chen He and Tobias Klein

**INDUSTRIAL ORGANIZATION**

ISSN 0265-8003

# Advertising as a reminder: Evidence from the Dutch State Lottery

*Chen He and Tobias Klein*

Discussion Paper DP12948
First Published 24 May 2018
This Revision 29 March 2022

Centre for Economic Policy Research
33 Great Sutton Street, London EC1V 0DX, UK
Tel: +44 (0)20 7183 8801
www.cepr.org

This Discussion Paper is issued under the auspices of the Centre's research programmes:

- Industrial Organization

# Advertising as a reminder: Evidence from the Dutch State Lottery


## Abstract

Consumers who intend to buy a product may forget to do so because they suffer from limited attention. Therefore, they may value being reminded by an advertisement. This phenomenon could be important in many markets, but is usually difficult to document. We study it in the context of buying a product that has existed for almost 300 years: a ticket for the Dutch State Lottery. This context is particularly suitable for our analysis, because the product is simple, it is very well-known, and there are multiple fixed and known purchase cycles per year. Moreover, TV and radio advertisements are designed to explicitly remind consumers to buy a lottery ticket before the draw. This can conveniently be done online. We develop an approach to distinguish reminder effects of advertising from other effects, such as conveying information about the size of the jackpot. We use minute-level advertising and online sales data and find that the reminder effect of advertising is strong. Reaching one percent of the population leads to an increase in online sales of 1.7 percent in the first hour after the advertisement is aired. We also provide direct evidence that reminding consumers does not only affect the timing of purchases, but also leads to market expansion. Finally, we estimate a model of consumer behavior under limited attention to quantify the effect on total sales. We find that total sales would be 15.7 percent lower without the reminder effect of advertising and that shifting advertising to the week of the draw would lead to a 10.8 percent increase in sales.



JEL Classification: M37, D12, D83

Keywords: reminder advertising, limited attention, adoption model



Chen He - hechen@shanghaitech.edu.cn
*ShanghaiTech University*

Tobias Klein - t.j.klein@uvt.nl
*C.E.P.R. and Tilburg University and CEPR*



Acknowledgements
We are grateful to Blue Mango Interactive and the Dutch State Lottery for providing us with the data. We thank Menno Zevenbergen and Pieter van Geel for stimulating discussions and useful comments. Without them, this project would not have been possible. The views expressed in this paper, however, do not necessarily reflect those of Bluemango or the Dutch State Lottery. We would also like to thank Senior Editor Puneet Manchanda, the Associate Editor, two anonymous referees, Jean-Pierre Dubé, Ulrich Doraszelski, Eric French, Ulrich Kaiser, Carl Mela, Peter Newberry, Martin Peitz, Pasquale Schiraldi, Stephan Seiler, Andrey Simonov, Rik Pieters, Frank Verboven, Ken Wilbur, Joachim Winter, and participants of a seminar in Oxford, at eBay in San Jose, the 2015 Media Economics Workshop in Stellenbosch, the 2016 meeting of the IO group of the Verein für Socialpolitik, the 2016 Digital Marketing Conference at Stanford, the 2017 C.E.P.R. Conference on Applied Industrial Organization, the 2018 EEA/ESEM in Cologne, and the 2019 Choice Symposium in Chesapeake Bay for their for comments and suggestions.

# Advertising as a reminder: Evidence from the Dutch State Lottery*

Chen He and Tobias J. Klein†

March 2022

**Abstract**

Consumers who intend to buy a product may forget to do so because they suffer from limited attention. Therefore, they may value being reminded by an advertisement. This phenomenon could be important in many markets, but is usually difficult to document. We study it in the context of buying a product that has existed for almost 300 years: a ticket for the Dutch State Lottery. This context is particularly suitable for our analysis, because the product is simple, it is very well-known, and there are multiple fixed and known purchase cycles per year. Moreover, TV and radio advertisements are designed to explicitly remind consumers to buy a lottery ticket before the draw. This can conveniently be done online. We develop an approach to distinguish reminder effects of advertising from other effects, such as conveying information about the size of the jackpot. We use minute-level advertising and online sales data and find that the reminder effect of advertising is strong. Reaching one percent of the population leads to an increase in online sales of 1.7 percent in the first hour after the advertisement is aired. We also provide direct evidence that reminding consumers does not only affect the timing of purchases, but also leads to market expansion. Finally, we estimate a model of consumer behavior under limited attention to quantify the effect on total sales. We find that total sales would be 15.7 percent lower without the reminder effect of advertising and that shifting advertising to the week of the draw would lead to a 10.8 percent increase in sales.

**Key words:** Reminder advertising, limited attention, adoption model.

**JEL-classification:** M37, D12, D83.

*We are grateful to Blue Mango Interactive and the Dutch State Lottery for providing us with the data. We thank Menno Zevenbergen and Pieter van Geel for stimulating discussions and useful comments. Without them, this project would not have been possible. The views expressed in this paper, however, do not necessarily reflect those of Bluemango or the Dutch State Lottery. We would also like to thank Senior Editor Puneet Manchanda, the Associate Editor, two anonymous referees, Jean-Pierre Dubé, Ulrich Doraszelski, Eric French, Ulrich Kaiser, Carl Mela, Peter Newberry, Martin Peitz, Pasquale Schiraldi, Stephan Seiler, Andrey Simonov, Rik Pieters, Frank Verboven, Ken Wilbur, Joachim Winter, and participants of a seminar in Oxford, at eBay in San Jose, the 2015 Media Economics Workshop in Stellenbosch, the 2016 meeting of the IO group of the Verein für Socialpolitik, the 2016 Digital Marketing Conference at Stanford, the 2017 C.E.P.R. Conference on Applied Industrial Organization, the 2018 EEA/ESEM in Cologne, and the 2019 Choice Symposium in Chesapeake Bay for their for comments and suggestions.

†He: ShanghaiTech University; e-mail: hechen@shanghaitech.edu.cn. Klein: C.E.P.R., TILEC, CentER, Tilburg University; e-mail: T.J.Klein@uvt.nl.

# 1 Introduction

Models of consumer behavior usually assert that not making a purchase is a deliberate choice. However, in practice, not making a purchase could be driven by limited attention rather than preference. In this situation, consumers may value being reminded by an advertisement.

This is potentially important in many markets, ranging from markets for consumer packaged goods to markets for health insurance. Important questions in this context are how big the reminder effect of advertising is, at what times it is most effective to remind consumers, and whether reminder advertising mainly affects the timing of purchases or also the probability to buy at all. However, so far, little is known about reminder advertising. One of the reasons for this is that it is generally difficult to isolate its effects, as advertisements generally also have an effect on the inclination of consumers to buy a product, either because they convey information or because they have an effect on the valuation consumers have for the product.

In this paper, we develop an approach that allows us to estimate the effect of reminding consumers through an advertisement. Our approach leverages access to high frequency data at the minute level in a setting in which advertising reaches consumers either on TV or the radio and consumers have the possibility to buy the product online. The key idea behind our approach is that the effect of providing information or influencing the valuation consumers have for the product will last longer than just a few hours, while the reminder effect vanishes much more quickly.

We use the approach to study reminder advertising in the context of buying a product online that has existed for almost 300 years: a ticket for the Dutch State Lottery.[1] This context is particularly suitable for our analysis, because the product is simple, it is very well-known, there are known purchase cycles, and advertisements explicitly reminded consumers to buy a ticket. Our minute-level data allow us to credibly identify the short-run effects of TV and radio advertising on online sales. The exact timing of advertisements is beyond the control of the firm and therefore, the thought experiment we can undertake is to compare sales just before the advertisement was aired to sales right after this.

We find the short-run effects of advertising to be sizable. Reaching 1 percent of the population leads to an average increase of sales by about 1.7 percent within one hour. The effects last about 3 hours and are the bigger the less time there is until the draw. Advertisements also conveyed information about the size of the jackpot. This means that the short-run effect of advertising could be the combination of a reminder effect and the effect of that information. We estimate the total short-run effects within the first four hours to be 26 percent (0.87 percentage points) higher when the jackpot is high, as compared to draws with low jackpots. However, this difference is not statistically significantly different from zero. Our interpretation of this finding is that *next to the reminder effect*, there could be an effect of advertising on the inclination of

[1] About 90 percent of the sales still take place offline. In Section 3 we discuss implications of this for the interpretation of our results.

consumers to buy given that they are thinking about buying, which originates in the information contained in the advertisements. Unlike the reminder effect, it seems to be hard to precisely estimate this additional effect of the information on the jackpot size. Besides, we find that advertising has a short-run effect until the end of the period in which tickets can be bought. This means that reminding consumers does *not only* lead to purchase acceleration (individuals buying earlier rather than later), but *also* to market expansion (more people buying tickets online in total). The reason is that advertising effects shortly before the draw cannot be due to purchase acceleration, and hence reflect market expansion.[2]

Our reduced-form analysis is useful to estimate short-term effects of advertising, but is not suited to predict the effects of dynamic advertising strategies on total monthly sales. The reason is that it does not explicitly take into account that consumers who buy a ticket leave the market. Therefore, we also estimate a simple structural model of consumer behavior under limited attention. We use it to quantify the effect of counterfactual dynamic advertising strategies on total online sales for the entire month. Our counterfactual simulations suggest that total sales would be 15.7 percent lower without advertising. Shifting advertising to the week of the draw would lead to a 10.8 percent increase in sales. The reason is that in that week, consumers are more likely to buy once they are reminded.

We see our contribution as being threefold. First, we propose a new framework in which advertising can also act as a reminder. We use the framework to derive our reduced-form estimation equation and to show how reminder effects of advertising can be separated from other effects. Second, we estimate the reminder effect of advertising in a context that is particularly well-suited for this—the effect of TV and radio advertising on online sales of lottery tickets. Third, we propose a simple structural model that allows us to quantify the effects of reminder advertising on total monthly sales.

To the best of our knowledge, the idea that advertising may serve as a reminder is not prominently featured in the academic literature. One exception is Krugman (1972), who asked the question how often consumers should be reached by advertising and then argued that consumers need to first understand the nature of the stimulus, then evaluate the personal relevance, and finally are reminded to buy when they are in a position to do so.

Reminder advertising can be seen as a generalization of the concept of purchase facilitation. Originally, Rossiter and Percy (1987) described purchase facilitation as providing information to individuals who intended to buy a product, for instance about the closest retailer at which a product can be bought and how one can pay for it. Purchase facilitation resembles reminder advertising in the sense that it helps the consumer to turn the intent to buy a product into an actual purchase. However, reminder advertising overcomes a different challenge than purchase facilitation, as it reminds individuals of their purchase intent without providing new information.

The rest of this paper is structured as follows. Next, in Section 2, we relate our paper to the literature. Section 3 gives a brief overview over the market for lottery tickets in the Nether-

[2]We would like to thank Martin Peitz for this suggestion.

lands. Section 4 describes the data. In Section 5, we provide a precise definition of reminder advertising and develop our empirical approach to estimate the reminder effects of advertising. Section 6 presents our empirical results. Section 7 develops our structural model of lottery ticket demand with advertising effects and presents estimates and the results of counterfactual experiments. Section 8 concludes. The (intended) Online Appendix is attached at the very end. Appendix A contains additional tables and figures that we refer to in the main text. Appendix B contains additional results related to the reduced-form analysis, including a number of robustness checks. Appendix C contains details and many additional results related to the structural model.

## 2 Literature and contribution

The focus of our paper is on one particular role of advertising, namely to remind consumers to act on their preference. With this, first and foremost, we relate and contribute to the large literature on advertising effects.[3]

One strand of the advertising literature is concerned with characterizing the mechanism through which advertising affects consumers. Usually, a distinction is made between informative advertising, persuasive advertising, and advertising that acts as a complement to consumption. Contributions include Ackerberg (2001, 2003), Mehta et al. (2008), and Hartmann and Klapper (2017). However, this distinction abstracts from limited attention and therefore, advertising that acts as a reminder is not easy to fit into this taxonomy. DellaVigna and Gentzkow (2010) more recently distinguish between belief-based and preference-based models. Informative advertising changes beliefs, whereas the other two types affect behavior by affecting preferences. They remark that in belief-based models advertising never makes consumers worse off, *ceteris paribus*. At the same time, they point out that there is a "blurry" area in-between when advertising does neither convey information nor does it have an effect on preferences (p. 656). This is also true in our case, where consumers suffer from limited attention and advertising (also) makes them think about buying the product. A similar mechanism is at play in a model proposed by Shapiro (2006), where consumers forget about past consumption experiences and advertising helps them recall those. However, in our case, consumers do not forget about past consumption experiences, but forget to make an intended purchase. In both cases, advertising acts as a reminder, but the reminder effect we study operates *via* consumer attention, whereas in Shapiro (2006) consumers are reminded of past consumption experiences. Sahni (2015) studies yet another mechanism. He relates consumer learning about the existence of products to the spacing of advertising over time.

Another strand of the advertising literature studies the effects of TV advertising. Lodish et al. (1995) summarize the earlier literature on the effectiveness of TV advertising and docu-

[3]Summarizing this literature is beyond the scope of this paper. See Bagwell (2007) for an excellent survey on the economics of advertising.

ment a combination of no and positive effects. Hu et al. (2007) find that the effects have increased in later years. Dubé et al. (2005) estimate a goodwill stock model to show that dynamic advertising strategies can be optimal. More recent contributions include Stephens-Davidowitz et al. (2017), Shapiro (2018), and Shapiro et al. (2021). More and more papers use high frequency data and estimate the effect of TV advertising on behavior online, as we do. A first set of papers studies the effects on online search. See, e.g., Zigmond and Stipp (2010), Lewis and Reiley (2013), Joo et al. (2014), Joo et al. (2016), Chandrasekaran et al. (2018), and Du et al. (2019). Liaukonyte et al. (2015), and Lambrecht et al. (2020), among others, complement these papers with evidence from high-frequency advertising and sales (as opposed to search) data. Papers that use high frequency data generally find that advertising effects are strong. Liaukonyte et al. (2015) emphasize that this may be due to the fact that consumers can respond immediately to the advertisement, because they have a "second screen", for instance a smartphone, within reach and can use it to order right away. A related well-established finding is that promotions in stores, in particular feature and displays, have large effects (Blattberg et al., 1995). Also here, consumers can react directly and do not first have to travel to the store. However, Lambrecht et al. (2020) emphasize that the short-term effects do not translate into an increase in either browsing or sales over a period of multiple weeks, implying that the firm does not see an increase in revenue as a result of the TV advertising campaign.

More broadly, we think of reminder advertising as influencing consumer choice by inducing them to think about buying the product, or consider buying it. This establishes a link to the literature on product consideration. Roberts and Lattin (1991, 1997) summarize the early contributions to this literature. Bronnenberg and Vanhonacker (1996) propose a model of a two-stage choice process in which consumers first determine the choice set and then make a choice. Allenby and Ginter (1995) study the effects of display and feature advertising on consideration sets. Sovinsky Goeree (2008) and Draganska and Klapper (2011) estimate similar static models with a consideration stage. Terui et al. (2011) use scanner data and find that strong support for advertising effects on choice through an indirect route of consideration set formation that does not directly affect brand utility. Van Nierop et al. (2010), Manzini and Mariotti (2014), and Abaluck and Adams-Prassl (2021) discuss the more recent literature and demonstrate that consideration sets can be inferred from choice data.

Our paper is also related to the literature that studies the effects of inattention and information treatments.[4] A first set of papers in that literature studies low observed rates of switching between providers of a service (inertia). These include Hortaçsu et al. (2017), Ho et al. (2017), Heiss et al. (2021). A second set of papers studies the effects of reminders, typically by conducting field experiments (e.g. Calzolari and Nardotto, 2016).

Finally, our paper relates to the literature that is concerned with modeling the decision when to buy a product. Our model is static in the sense that consumers are not forward-looking, but nonetheless consumers decide at each point in time whether to buy a ticket or wait. Melnikov

[4] There is also a large, broader literature on inattention. See Gabaix (2019) for an excellent survey.

(2013) and De Groote and Verboven (2019) estimate more sophisticated, dynamic models.[5]

# 3 The Dutch State Lottery

The market for lottery tickets in the Netherlands is very concentrated, with three organizations conducting different types of lotteries. First, the Stichting Exploitatie Nederlandse Staatsloterij, from which we received the data, offers lottery tickets for the Dutch State Lottery (in Dutch: Staatsloterij) and the Millions Game (Miljoenenspel). The Dutch State Lottery has a history going back to the year 1726 and is run by the government. It is by far the biggest of its kind in the Netherlands. The second player is the De Lotto. It offers the Lotto Game (Lottospel), which is comparable but much smaller in size, next to other games such as Eurojackpot and Scratch Tickets (Krasloten) and sports betting. In 2016, these two organizations merged. The third player is Nationale Goede Doelen Loterijen offering a ZIP Code Lottery (Postcodeloterij), whose main purpose is to donate money to charity. For that reason, it is not directly comparable to the other two lotteries.[6]

The lottery run by the Dutch State Lottery is classical. There are 16 draws in a calendar year. 12 of them are regular draws and 4 of them are special draws. Regular draws take place on the 10th of every month. The dates of 4 additional special draws vary slightly from year to year. In 2014 (the year for which we have data), the 4 special draws were on April 26 (King's Day in the Netherlands), on June 24, October 1 and on December 31 (the New Year's Eve draw). All draws but the last in a year take place at 8pm (Central European Time). From 6pm onward, no more tickets can be bought for that draw.

A ticket has a combination of numbers and Arabic letters and a consumer can choose some of them. The size of the prize depends then on how many numbers and letters of a ticket match with the ones of the winning combination.[7] On top of that, there is a jackpot whose size varies over time. For all draws but the very last one in a year, consumers can choose between a full ticket that costs 15 euros and multiples of one fifth of a ticket. For the last draw, the price of a ticket is 15 euros and consumers can buy multiples of one half of a ticket. Winning amounts are then scaled accordingly.

The expected payoff for an individual depends on the number of people who hold a ticket on the day of the draw and on the size of the jackpot. While it is not communicated how many people have bought a ticket at any given point in time, the jackpot size for the next draw is

[5] In a previous version of this paper, we also estimated a dynamic model with an attention stage. See for instance Klein and He (2020). The results of counterfactual experiments were very similar.

[6] In 2014, the Dutch State Lottery had a turnover of 738 million euros with 579 million euros related to its lottery and De Lotto of 322 million euros with 144 million euros related to its lottery (https://over.nederlandseloterij.nl/over-ons/publicaties, accessed February 2022). The turnover of Nationale Goede Doelen Loterijen was 847 million euros in total with 624 million euros related to its charity ZIP code lottery (https://view.publitas.com/nationale-postcode-loterij-nv/npl-jaarverslag-2014/page/58-59, accessed February 2022).

[7] For an example see https://www.loten.nl/staatsloterij/ (accessed February 2022).

known right after the previous draw. Consumers can learn about it from billboards, posters at selling locations for tickets, or online. Some advertisements also contain information about the jackpot size (our empirical approach takes this into account).

Tickets can be purchased in two ways: they can either be purchased online via the official website of the Dutch State Lottery, or offline, for example in a supermarket or a gas station. To the best of our knowledge, about 90 percent of the sales were offline in the year for which we have data (the exact number is considered a trade secret), but nevertheless the online business was considered important. This means that we must make an important qualification. In this paper, we study the effects of advertising on online sales only. There could in addition be two effects of advertising on offline sales. The first effect is that some consumers bought offline in response to seeing the advertisement. The second effect is a cannibalization effect, namely that some of the consumers who were motivated to buy online would have bought offline anyway. Unfortunately, we cannot quantify these effects.[8] Nonetheless, as we have explained in the introduction, we believe that studying the effect of advertising on online sales allows us to document that advertising can also act as a reminder. This is the focus of this paper.

# 4 Data

## 4.1 Overview

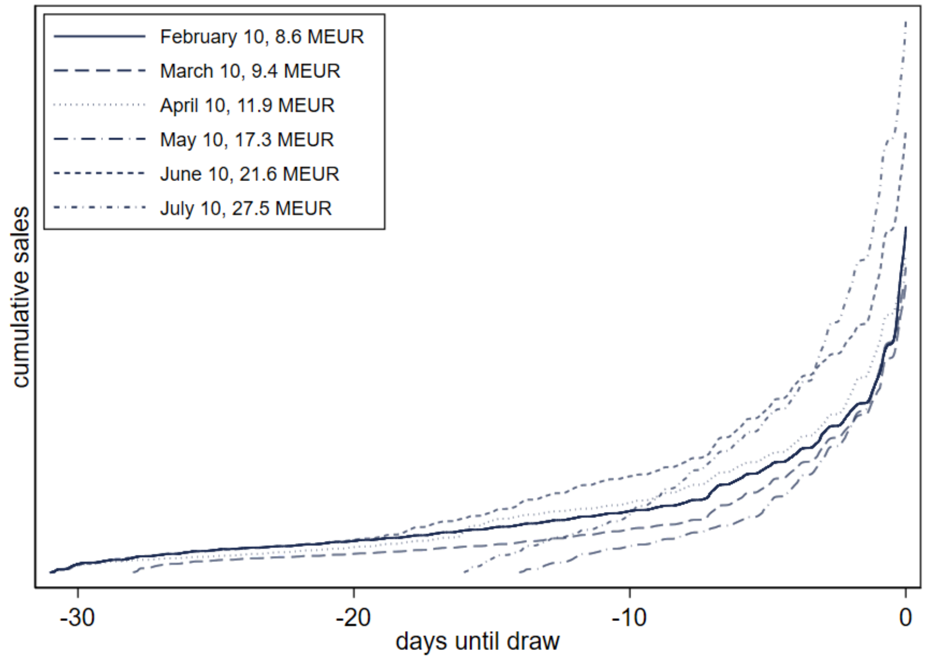
Our data are for 2014 and consist of 3 parts: online transactions, TV and radio advertising, and jackpot sizes. The transaction data are collected at the minute level. We observe the number of lottery tickets sold online.[9] The advertising data consists of minute-level measurements of gross rating points (GRP's), separately for TV and radio advertising. GRP's measure impressions as a percentage of the target population at a given point in time. For example, 5 GRP's in our data mean that in that minute 5 percent of the target population (in our case the general population) are exposed to an advertisement. This is a standard measure in the advertising industry.

Besides, we observe the jackpot size for the 12 regular draws in 2014. There is no jackpot size for the 4 special draws, as more involved rules apply to them. For example, on the drawing day, every 15 minutes consumers can win an additional 100,000 euros. In the empirical analysis, we will capture differences across draws in a flexible way.

[8]The effect on online sales is likely a lower bound on the effect on total sales. The first effect on offline sales is likely positive. The second effect only leads to online sales instead of offline sales.

[9]These data have been collected using Google Analytics. In particular, visits to the "exit page" confirming payment have been recorded. This means that we do not observe what type of ticket a consumer has bought. Advertising also affects offline sales, and therefore, ideally, we would also like to observe the number of lottery tickets sold offline. However, offline transactions are not observed in the data set. At the same time, it generally takes longer until an offline sale takes place after an individual listens to a radio advertisement or sees a TV advertisement. At the minimum, this will be the time it takes between listening to a radio commercial in the car and buying a ticket in a shop. Therefore, it will be much more challenging to measure advertising effects in offline data—a challenge we try to overcome with our high frequency online sales data. For the interpretation of our results below we focus on online sales.

Figure 1: Cumulative sales for selected draws




Notes: This figure shows cumulative sales for 6 selected regular draws. The respective jackpot sizes are given in the legend. See Figure A1 in the Online Appendix for the remaining draws.

We are not allowed to report levels of sales and advertising. Therefore, we will only present relative numbers and (semi-) elasticities in the tables and figures below and some vertical axis will have no units of measurements.

## 4.2 Descriptive evidence

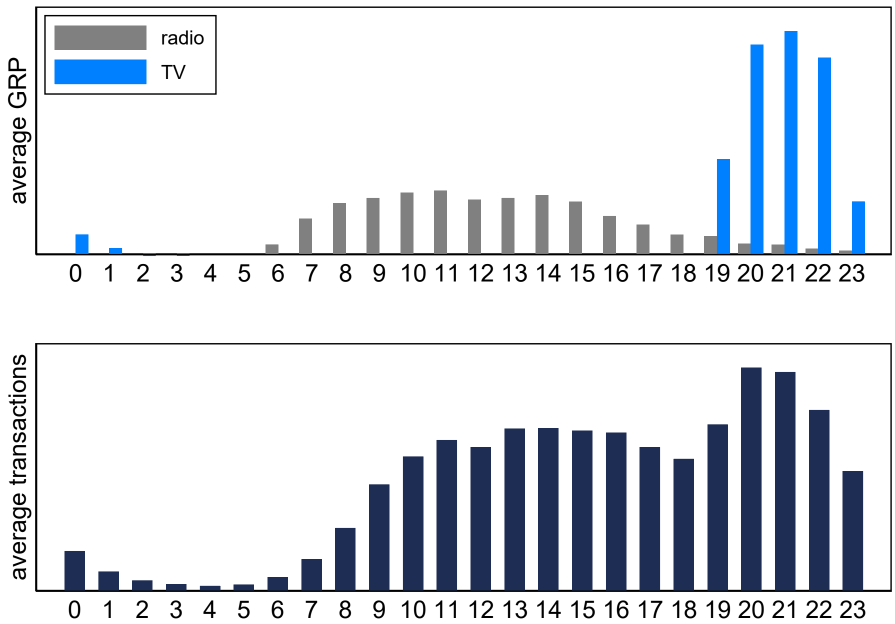
Figure 1 shows cumulative sales for 6 selected regular draws against the time until the draw, together with the respective jackpot size.[10] Some of the draws take place one full month after the previous draw, while others take place after less than a month. For example, the draw on July 10 follows on the one of June 24 and therefore the line for the draw on July 10 is only from June 24 (6:00 pm) to July 10 (5:59 pm).[11] The main take-away from this figure is that it strongly suggests that consumers value buying a ticket shortly before the draw. Interestingly, Inman and McAlister (1994) document a similar pattern for the effects of coupons that allow consumers to buy a product at a reduced price. Also here, the effects are bigger shortly before the expiration date.

The figure also shows that across draws there is a positive relationship between jackpot size and total sales (that is, cumulative sales on the day of the draw). The draw on July 10 has the

[10]Patterns for the other draws are similar. See Figure A1 in the Online Appendix for the remaining draws.

[11]We do not expect this to have big effects, however, because most tickets are sold in the week before the draw. But we do take this into account in our analysis.

Figure 2: Advertising and sales during the day




Notes: This figure shows average GRP's and sales for different times of the day. To produce this figure we first aggregate sales at the hourly level and then average over days and draws. We exclude the respective day of the draw because tickets can only be bought until 6pm on that day and advertising and sales are higher just before this deadline. See Figure A2 in the Online Appendix for the pattern on the day of the draw.

largest total sales of the 6 draws. It also has the largest jackpot size. The second largest total sales are for the draw on June 10, which also has the second largest jackpot size. However, in general, it is not true that larger jackpot size always implies larger total sales.

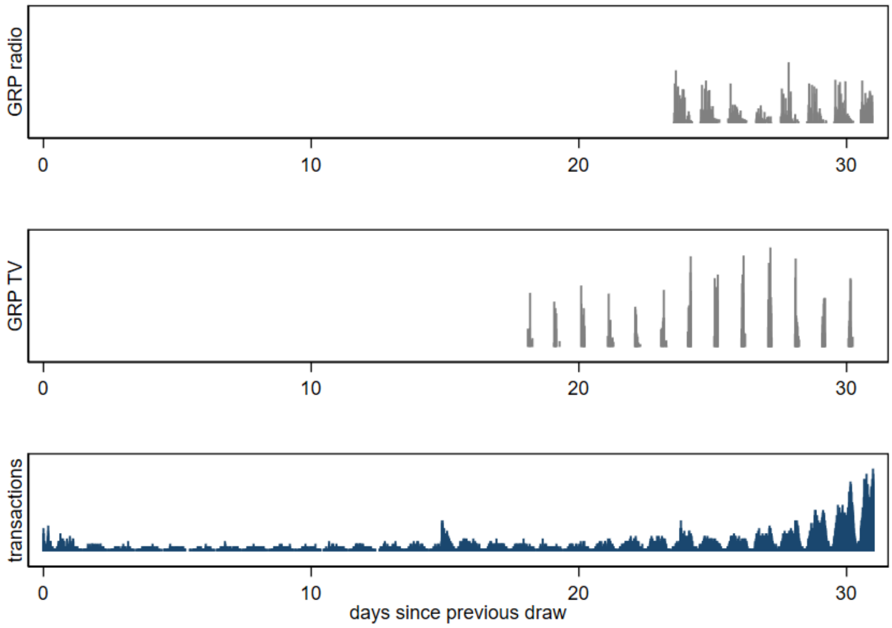
We further explore differences across draws by regressing the log of the total number of tickets sold online on the log of the jackpot size and the total number of days between the date of the previous and current draw.[12] Obviously, we only have 16 observations and jackpot size only varies among the 12 regular draws. Nevertheless, we find a statistically significant relationship between jackpot size and sales. We estimate the effect of a 1 percent increase in the jackpot size to be a 0.4 percent increase in total sales. We find no statistically significant effects of total sales in the previous draw on total sales in the current draw.

Figure 2 shows the pattern of sales and GRP's across different hours of a day.[13] We average over all days in 2014 except for the days of the draw. The reason for this is that the time until which tickets can be bought is 6pm and we observe that a large amount of sales occurs during the hours before 6pm. At the same time, we observe that sales are very low in the first several hours after 6pm on the day of the draw, as one would expect. So, by excluding those 16 drawing days, we can get a cleaner picture on how sales and GRP's are distributed over time during a

[12]See Table A1 in the Online Appendix for details.

[13]See Figure A2 in the Online Appendix for the pattern on the day of the draw.

Figure 3: GRP's at the minute-level for a regular draw




Notes: This figure shows radio and TV GRP's and sales at the minute level, for the regular draw on April 10, 2014. Tickets for the next draw can be bought from 6pm on the day of the previous draw, which is depicted as 0 days since the previous draw.

typical day.

We distinguish between radio and TV advertisements. TV advertisements are concentrated during evening and night hours, while radio advertisements are more likely to be aired in the morning and in the afternoon. This clear separation is due to the fact that in the Netherlands, TV advertisements related to gambling must not be aired during the day time, until 7pm.[14]

Figure 2 shows that GRP's are positively correlated with sales. During the hours in which sales are high, GRP's are also high. However, this does not necessarily mean that advertising has positive effects, because GRP's have not been assigned randomly. For instance, it could be that consumers have more time in the evening and are therefore more likely to buy a lottery ticket anyway.[15]

Figure 3 shows GRP's and sales at the minute level for one regular draw.[16] We see that the firm starts advertising on the 17th day after the last regular draw, while sales only increase in the last days before the draw.

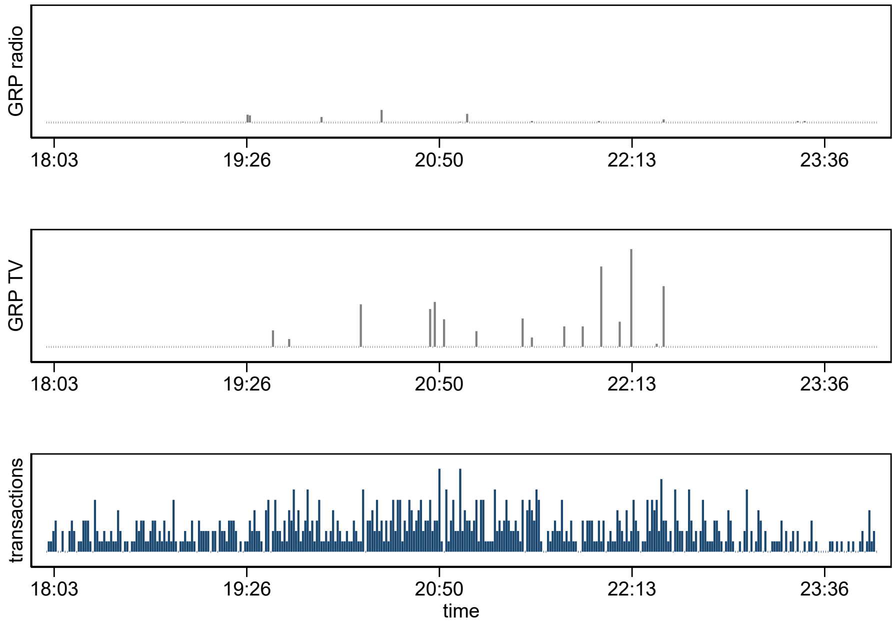
Finally, Figure 4 zooms in further and shows the pattern for one of the days in Figure 3.

[14]We have tried to exploit this regression discontinuity design to produce estimates of advertising effects. However, it turns out to be difficult to distinguish the discontinuity in the total number of GRP's from a flexible time trend. The reason is that the number of GRP's increases in a continuous manner between 7pm and 9pm and did not sharply jump to a high level right after 7pm.

[15]This is a well-known challenge for the analysis of advertising effects. Our identification strategy for overcoming this challenge is described in Section 5 below.

[16]Figure A3 in the Online Appendix shows GRP's and sales for the special draw on April 26. Patterns are similar.

Figure 4: GRP's at the minute-level for a short time window




Notes: This figure shows radio and TV GRP's and sales at the minute level for a short time window on April 3, 2014.

Related to our identification strategy described below, it is interesting to notice that the raw data presented in Figure 4 already show some evidence of short-run sales responses to advertising. For example, there are some spikes of GRP's just before 20:50, followed by spikes of sales several minutes later. In Section 6, we investigate this more systematically.

# 5 Advertising as a reminder

In this section, we lay out our framework. We use it to provide a precise definition of the reminder effect of advertising. Our framework also allows for other effects of advertising. We derive a reduced-form equation and show how one can use high frequency data to estimate the reminder effect of advertising.

## 5.1 Framework

There are $M$ consumers in the market. Time $t$ is discrete and measured in minutes. Consumers are well-informed about the existence of the product. In each period, consumers may think about buying a ticket. If they think about buying, they either buy or not.

First consider the baseline situation without advertising. Baseline quantities at time $t$ are indexed by $0t$. Denote the probability to think about buying by $P_{0t}(\text{think})$, the probability to buy given that a consumer is thinking about buying by $P_{0t}(\text{buy}|\text{think})$, and the probability to buy by $P_{0t}(\text{buy})$. From this it follows that the baseline probability to buy is given by $P_{0t}(\text{buy}) \equiv P_{0t}(\text{think}) \cdot P_{0t}(\text{buy}|\text{think})$ and that baseline sales are given by $sales_{0t} \equiv M \cdot P_{0t}(\text{think}) \cdot P_{0t}(\text{buy}|\text{think})$.[17]

Relative to this baseline, consider a counterfactual world in which an advertisement reached a consumer at $t_0$. For $t \geq t_0$, this advertisement has two effects, one on the probability to think about buying and one on the probability to buy given thinking.

The first effect of the advertisement is that it changes the probability to think about buying from $P_{0t}(\text{think})$ to $P_t(\text{think}) = P_{0t}(\text{think}) \cdot (1+\Delta_\tau)$, where $\tau \equiv t - t_0$ denotes time relative to $t_0$. We call $\Delta_\tau$ the reminder effect of advertising. Note that this reminder effect depends on the time since the advertisement reached the consumer. We think of it as short-lived, in the sense that we expect $\Delta_\tau$ to go to zero within a few hours. Formally, we assume that there is a $\check{\tau}$ so that $\Delta_\tau = 0$ for all $\tau \geq \check{\tau}$. The underlying way to think about consumers is that they are exposed to limits of information processing power and attention and may therefore forget to make an intended purchase. An advertisement reminds them of that, in the sense that it draws attention to this, but this effect on attention fades away quickly.

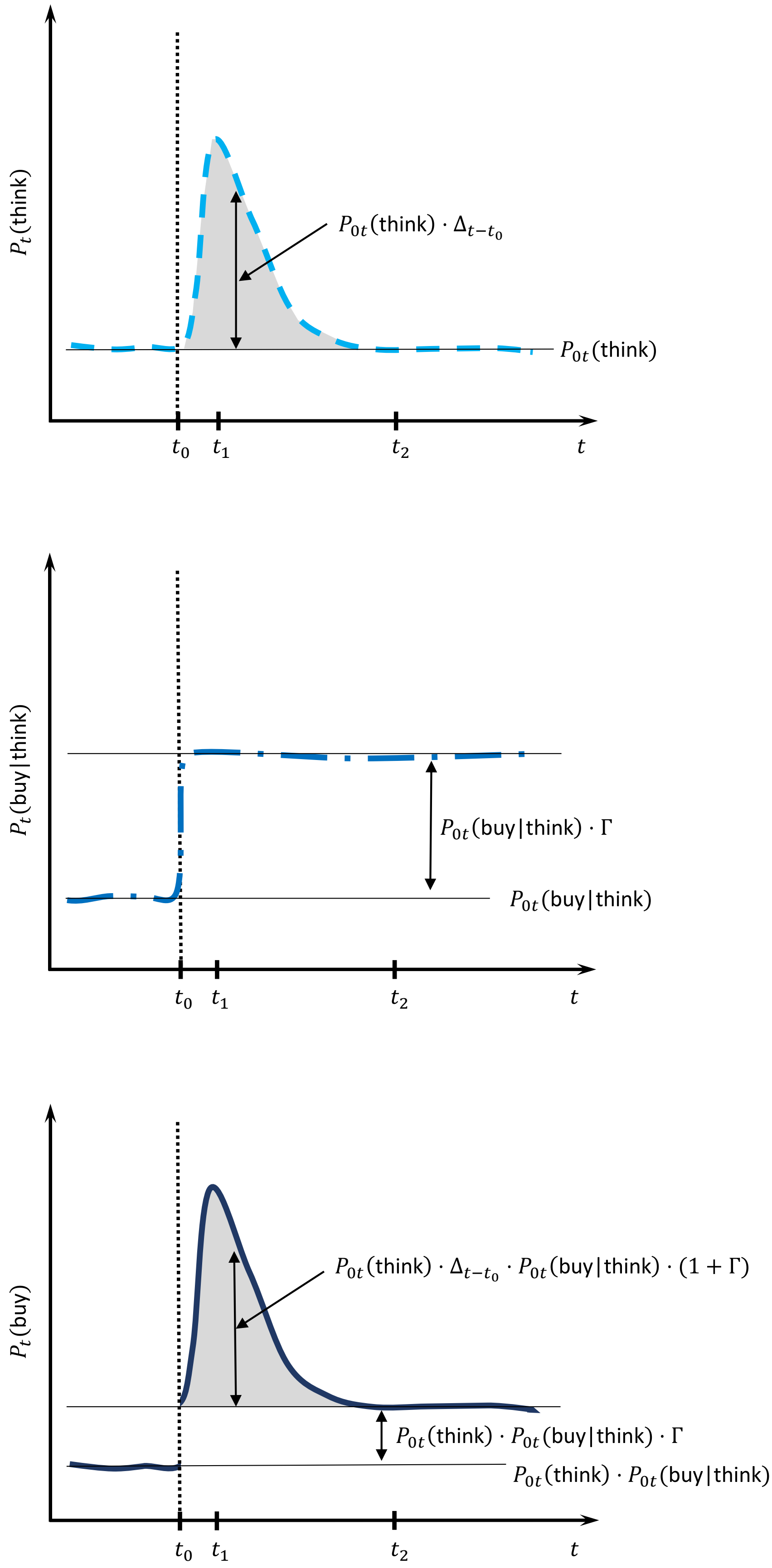
Figure 5 illustrates. The plot on top shows how the probability to think about buying, $P_t(\text{think})$, is influenced by the advertisement. Until $t_0$, the time at which the advertisement is aired, it is equal to the baseline probability $P_{0t}(\text{think})$. Then, it increases until $t_1$ and decreases again until $t_2$, when the probability to think about buying is back to the baseline level $P_{0t}(\text{think})$. The difference between $P_t(\text{think})$ and $P_{0t}(\text{think})$ is equal to $P_{0t}(\text{think}) \cdot \Delta_{t-t_0}$ and the figure shows that $\Delta_{t-t_0} = 0$ for $t < t_0$ and $t \geq t_2$.

The second effect of the advertisement is that it changes the probability to buy given that the consumer is thinking about buying to $P_{0t}(\text{buy}|\text{think}) \cdot (1+\Gamma)$. In our context, this is the effect of information about the jackpot size. More generally, this could be any effect advertising has on the attitude a consumer has towards buying the product given that she thinks about buying. This could be either because the advertisement changes preferences or acts as a complement to consumption, or because it contains information. In our context, advertisements often name the jackpot size. If it is higher than consumers expect, then $\Gamma$ could be positive. The defining characteristic of this effect is that it is long-lived. In line with this, note that $\Gamma$ is not indexed by $\tau$. In Figure 5, the plot in the middle shows a situation in which $\Gamma > 0$. Until $t_0$, $P_t(\text{buy|think})$ is equal to the baseline, $P_{0t}(\text{buy|think})$, and then changes to and remains at $P_{0t}(\text{buy|think}) \cdot (1+\Gamma)$.

The overall effect of advertising is a change in the probability of buying from $P_{0t}(\text{buy}) = P_{0t}(\text{think}) \cdot P_{0t}(\text{buy}|\text{think})$ to $P_t(\text{buy}) = P_{0t}(\text{think}) \cdot (1+\Delta_\tau) \cdot P_{0t}(\text{buy}|\text{think}) \cdot (1+\Gamma)$. The third plot of Figure 5 shows that before $t_0$, $P_t(\text{buy}) = P_{0t}(\text{buy})$. After $t_0$, $P_t(\text{buy})$ first increases until

[17] One can generalize this model by introducing consideration sets. A consideration set contains the products a consumer is aware of. Then, the probability of buying is the product of three probabilities: the probability to think about buying, the probability that the product is in the consideration set (which does not depend on whether the consumer thinks about buying), and the probability that the consumer buys given that the product is in the consideration set and that she thinks about buying. In our case, the consumer is well-informed about the existence of the product and therefore we have that the probability that the product is in the consideration set is equal to 1.

Figure 5: Advertising effects



$t_1$ to $P_{0t}(\text{buy}) \cdot (1+\Delta_\tau) \cdot (1+\Gamma)$ and then starts to decrease to the new level $P_t(\text{buy}) = P_{0t}(\text{buy}) \cdot (1+\Gamma)$, at which it remains from $t_2$ onward.[18]

## 5.2 Identification of $\Delta_\tau$ and $\Gamma$

Suppose an advertisement is aired in period $t-\tau$. The fraction of consumers who are reached by this advertisement is given by the number of GRP's divided by 100. This means that in $t$, for $M \cdot \frac{grp_{t-\tau}}{100}$ consumers the probability to buy is $P_{0t}(\text{think}) \cdot (1+\Delta_\tau) \cdot P_{0t}(\text{buy}|\text{think}) \cdot (1+\Gamma)$. For the remaining $M \cdot \left(1 - \frac{grp_{t-\tau}}{100}\right)$ consumers it is $P_{0t}(\text{think}) \cdot P_{0t}(\text{buy}|\text{think})$. Hence, sales in $t$ are

$$\begin{aligned} sales_t &= M \cdot \frac{grp_{t-\tau}}{100} \cdot P_{0t}(\text{think}) \cdot (1+\Delta_\tau) \cdot P_{0t}(\text{buy}|\text{think}) \cdot (1+\Gamma) \\ &\quad + M \cdot \left(1 - \frac{grp_{t-\tau}}{100}\right) \cdot P_{0t}(\text{think}) \cdot P_{0t}(\text{buy}|\text{think}) \\ &= M \cdot P_{0t}(\text{think}) \cdot P_{0t}(\text{buy}|\text{think}) \cdot \left[\left(1 - \frac{grp_{t-\tau}}{100}\right) + \frac{grp_{t-\tau}}{100} \cdot (1+\Delta_\tau) \cdot (1+\Gamma)\right] \cdot \end{aligned}$$

Taking logs and then performing a first order Taylor series expansion in $grp_{t-\tau}$ about $grp_{t-\tau} = 0$ gives

$$\log(sales_t) \approx \alpha_t + \beta_\tau grp_{t-\tau},$$

where $\alpha_t \equiv \log\left(M \cdot P_{0t}(\text{think}) \cdot P_{0t}(\text{buy}|\text{think})\right)$ and

$$\beta_\tau \equiv \frac{1}{100}[(1+\Delta_\tau) \cdot (1+\Gamma) - 1]. \tag{1}$$

So far, we have only considered the case in which there was one advertisement. We can make the above argument for any $\tau$. Then, a first order Taylor series expansion in $\bar{\tau} \geq \check{\tau}$ lags of GRP's about 0 gives[19]

$$\log(sales_t) \approx \alpha_t + \sum_{\tau=0}^{\bar{\tau}} \beta_\tau grp_{t-\tau}.$$

To take this to the data, we form time blocks and decompose $\alpha_t$ into a component that is specific to the time block $t$ lies in, $\alpha_{b(t)}$, and a second residual component $\varepsilon_t$ that also contains the approximation error. This gives

$$\log(sales_t) = \alpha_{b(t)} + \sum_{\tau=0}^{\bar{\tau}} \beta_\tau grp_{t-\tau} + \varepsilon_t. \tag{2}$$

[18]Part of the increase in the probability of buying between $t_0$ and $t_1$ could be driven by consumers going through the buying process on the website. Here, we attribute this to an increase in the probability to think about buying. The model could be generalized to allow for a delay between the decision to buy and the actual purchase. Then, we could model the probability to think about buying as jumping up at $t_0$ and then monotonically decreasing until $t_2$, without a peak at $t_1$. Importantly for our analysis, even without this generalization, we can still interpret the shaded area in the third plot as the reminder effect of advertising on sales, as it is *caused* by an increase in the probability to think about buying. The reason is that a delay between the decision to buy and the actual purchase does not change the size of that area. It would only shift the timing of additional purchases.

[19]Recall that $\Delta_\tau = 0$ for all $\tau \geq \check{\tau}$. Therefore, we need to consider at least $\check{\tau}$ lag terms.

For identification of $\beta_\tau$, $\tau = 0, \ldots, \bar{\tau}$, we make the following assumption.

**Assumption 1** (no endogeneity at minute level)**.** *$\varepsilon_t$ in period $t$ is independent of $grp_s$ in period $s$, for all combinations of $s$ and $t$.*

This assumption means that the number of GRP's in all periods $s$, including $t$, are not systematically related to $\varepsilon_t$. Importantly, it still allows for a relationship between $\alpha_{b(t)}$ and $grp_s$. For instance, it allows for more advertising on days closer to the day of the draw, more advertising in certain hours of the day, and more advertising for draws with a higher jackpot. This is important, because one would expect baseline sales and advertising levels to be positively correlated.

For two reasons, we expect Assumption 1 to hold in our context. First, advertising buying takes place several weeks in advance. The company chose not to determine exact times at which the advertisements were aired, because this would have been more expensive. It instead indicated in broad terms when and how much advertising should be aired. We do not know specifics about this, but we have been assured that the time window was at least several hours long.

Second, the company bought a certain quantity of consumers that were reached (measured in GRP's), and not a certain number of spots that were aired. For a given time in the future, it is uncertain how many viewers will be reached, as viewership demand depends on many factors other than the TV schedule, for instance the weather. This means that the target quantity bought by the firm is allocated to multiple spots, until the amount of advertising that was actually bought has been provided (see also Dubé et al., 2005). For each of these spots, even if the time was known in advance (which it isn't), the number of GRP would not be known at the time advertising is bought.

Under Assumption 1, we can estimate the coefficients on the lag terms, $\beta_\tau$, by regressing the log of sales on lags of GRP's, while controlling for time block fixed effects. To see why minute-level data are useful to estimate $\beta_\tau$, we can start from (2) and take the first difference in $t$. For any minute but the first and the last of a block this eliminates the time block fixed effect $\alpha_{b(t)}$. Consider a situation where there is only an advertisement in $t$. Then, $\log(sales_t) - \log(sales_{t-1})$ will be informative about $\beta_0$, $\log(sales_{t+1}) - \log(sales_t)$ will be informative about $\beta_1 - \beta_0$, and so on. Once we know $\beta_0$ and all $\beta_\tau - \beta_{\tau-1}$, we can recover all $\beta_\tau$. This means that we can also estimate $\alpha_{b(t)}$.[20]

To identify $\Delta_\tau$ and $\Gamma$, we re-state restrictions we impose on the advertising effects in the following assumption.

**Assumption 2** (advertising effects)**.** *(i) $\Delta_\tau = 0$ for $\tau \geq \check{\tau}$, (ii) $\Gamma$ is constant for $\tau \leq \bar{\tau}$, and (iii) $\bar{\tau} \geq \check{\tau}$.*

[20]This is obviously not a formal proof of identification. For a more formal argument, assume that Assumption 1 holds for $\log(sales_t) = \alpha_t + \sum_{\tau=0}^{\bar{\tau}} \beta_\tau grp_{t-\tau} + \alpha_{h(t)} + \varepsilon_t$. Then, the fixed effects estimator will consistently estimate the parameters $\beta_\ell$ under a standard rank condition.

Part (i) says that the reminder effect lasts for at most $\check{\tau}$ periods, part (ii) says that the effect of advertising on $P_t(\text{buy}|\text{think})$ is constant for at least $\bar{\tau}$ periods, and part (iii) says that we have included enough lag terms in (2). In Figure 5, $\check{\tau} = t_2 - t_0$, which means that the model would have to include at least $\bar{\tau} = \check{\tau}$ lag terms.

Under Assumption 2, it follows from (1) that $\beta_\tau = \frac{\Gamma}{100}$ for $\tau \geq \check{\tau}$ (recall that $\bar{\tau} \geq \check{\tau}$; this means that this argument applies for all $\beta_\tau$ with $\check{\tau} \leq \tau \leq \bar{\tau}$). Hence, $\Gamma$ is identified.[21] Once we know $\Gamma$, we can recover the parameters $\Delta_\tau$, as it follows from (1) that

$$\Delta_\tau = \frac{1 + 100 \cdot \beta_\tau}{1 + \Gamma} - 1 \tag{3}$$

for $\tau \leq \check{\tau}$.

# 6 Reduced-form analysis

## 6.1 Direct evidence for big advertisements

In our data, there are a number of relatively small advertisements. This means that there is often only a short amount of time between advertisements. For that reason, providing direct evidence on the effect of advertisements is challenging, as advertising effects may overlay each other. Our first approach to overcome this challenge is to *select* advertisements with at least 9 GRP's and to only keep the ones out of these advertisements for which we do not see another big advertisement in the hour before and after. This leaves us with 59 advertisements. Out of those, 51 were aired on TV. Figure A5 in Appendix A shows which advertisements were used.

We form time blocks that are 121 minutes long and cover the two hours around the time the advertisement was aired, respectively. We drop the rest of the data. We use these data to nonparametrically regress sales on the time to and since the advertisement was aired.[22] We rescale the estimates to obtain a figure that corresponds to $\beta_\tau$ in (2).[23]

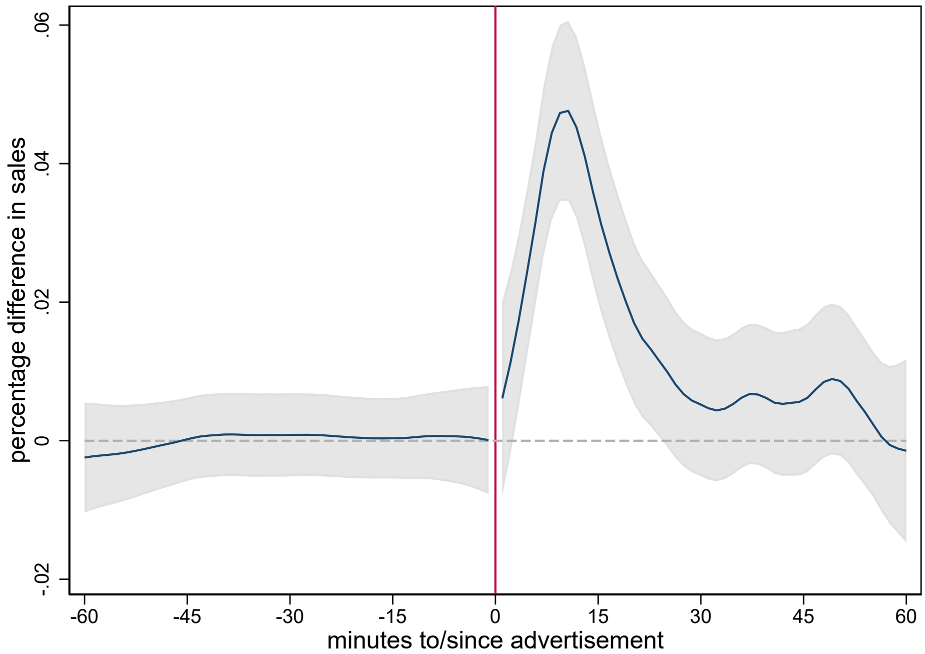
Figure 6 shows the resulting plot of the percentage difference in sales (relative to average sales in the hour before the advertisement was aired) against the time to and since the advertisement was aired. Notice that sales are flat in the 60 minutes before the advertisement was aired, in line with the idea these constitute a baseline that can be extrapolated. Against that baseline, we find that the effect of a big advertisement is an increase in sales that lasts for about 30 minutes. The effect is fairly immediate and dies out relatively quickly. It is a bit higher than 4 percent after a few minutes and overall leads to an increase of sales by 1.4 percent in the hour

[21] After $\bar{\tau}$ periods, this long-run effect is captured by the parameters $\alpha_{b(t)}$.

[22] We use separate nonparametric regressions for the time to and since the advertisement was aired, respectively. We used a local constant specification and the rule-of-thumb bandwidth.

[23] Denote estimates by $\widehat{sales}_\tau$, $\tau = -60, \ldots, 60$, average GRP by $\overline{grp}$, and average sales in the hour before the advertisement was aired by $\bar{y}$. Figure 6 shows $(\widehat{sales}_\tau - \bar{y})/\bar{y})/\overline{grp}$.

Figure 6: The effect of advertising on sales for big advertisements




Notes: This figure shows the effect of 1 GRP of advertising on sales. The shaded area depicts pointwise 95 percent confidence intervals. See text for additional details.

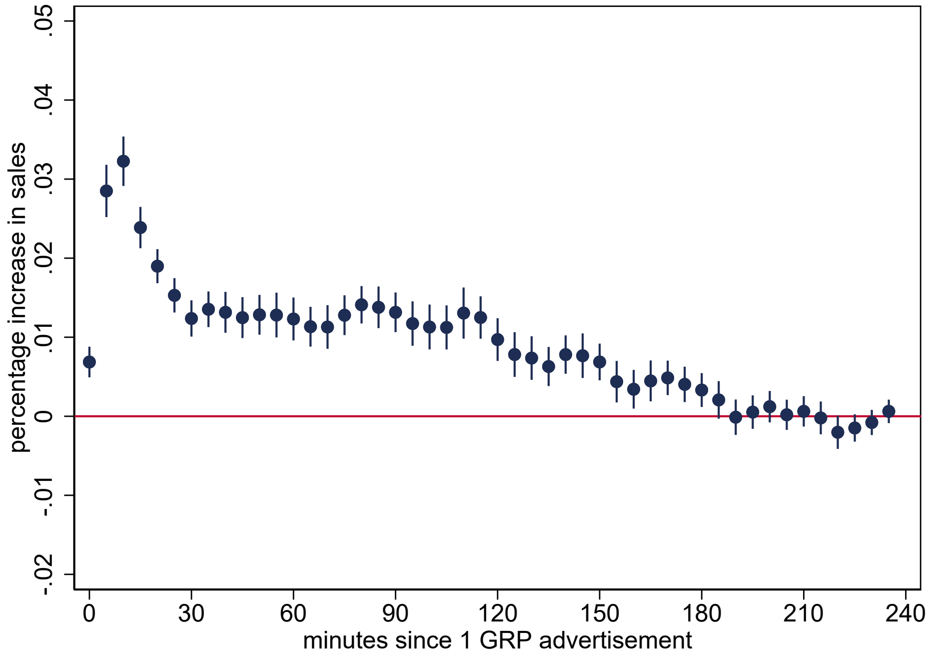
after it is aired.[24] Thereafter, sales return to the baseline before the advertisement, suggesting that, at least on average across all big advertisements (and therefore all draws), the information contained in advertisements does not have a longer-lasting effect, i.e. that on average $\Gamma$ is zero. It then follows from (3) that $\beta_\tau = \Delta_\tau$, which means that Figure 6 shows $\Delta_\tau$ for $\tau = -60, \ldots, 60$.

## 6.2 Evidence from a distributed lag model

To provide more systematic evidence without selecting advertisements, we next estimate a distributed lag model. For this, we start from (2) and assume that the length of a time block is an hour. Moreover, we impose that $\beta_\tau$ is piecewise-constant in $\tau$, in blocks of 5 minutes. The reason is that we specify $\bar{\tau} = 240$ in our baseline specification and would otherwise have to estimate 240 coefficients on 240 lag terms. This is in principle possible, but is too flexible in practice and produces noisy estimates. Using blocks of 5 minutes reduces this to 240/5=48 lag terms, which turns out to be manageable in our situation. To implement this, we generate regressors $sumgrp_{\ell,t} = \sum_{t'=t-5\cdot\ell+1}^{t-5\cdot\ell+5} grp_{t'}$ that respectively sum up lags of GRP's in groups of 5. The first lag term $sumgrp_{1,t}$ is the sum of the GRP's in $t$ and the previous 4 minutes and has coefficient $\bar{\beta}_1$. The second lag $sumgrp_{2,t}$ is the sum of the GRP's between 9 and 5 minutes in the past and has coefficient $\bar{\beta}_2$, and so on. Finally, we add 1 to sales before taking the log, as

[24] This is the average estimated value in the first hour.

Figure 7: The effect of advertising on sales




Notes: This figure shows estimates from a distributed lag model that was estimated at the minute level. There are 48 lags that are each for 5-minute intervals, so 4 hours in total. The dependent variable is the log of one plus sales. We control for hour of year fixed effects. The bars indicate 95 percent confidence intervals, which are based on standard errors that are clustered at the daily level. Table A2 in the Online Appendix shows the coefficients for the first 12 lags.

sales are sometimes 0 in our data.[25] With this, our estimation equation is

$$\log(1+sales_t)=\sum_{\ell=1}^{48}\bar{\beta}_\ell sumgrp_{\ell,t}+\alpha_{h(t)}+\varepsilon_t. \tag{4}$$

We then use the fixed effects estimator to estimate the 48 coefficients $\bar{\beta}_\ell$ on the 48 lag terms, controlling hour of year fixed effects $\alpha_{h(t)}$. As for inference, we cluster standard errors at the daily level. This means that we allow for a correlation of $\varepsilon_t$ within the day.

Figure 7 plots estimates of the coefficients on the lag terms. Table A2 in the Online Appendix reports the coefficient estimates for the first hour. We find that the effect of advertising increases until 10 to 14 minutes after the advertisement was aired and then decreases. The main effect is observed in the first hour, but there are effects thereafter. The maximal effect is an increase in sales of about 3.2 percent for each additional GRP of advertising, between 10 and 14 minutes after the advertisement was aired.

[25] Usually, $\bar{\beta}_\ell$ is interpreted as approximately a percentage change in sales, as it approximates the actual percentage change $\exp(\bar{\beta}_\ell)-1$ very well for small values of $\bar{\beta}_\ell$. This is still the case when we use the log of 1 plus sales as the dependent variable. To see this, assume that baseline sales are equal to 10 and the actual effect of 1 GRP of advertising is a 4 percent increase in sales to 10.4. Then, we have that $\log(1+10.4)-\log(1+10)=0.036$. This means that our coefficient estimate would be 0.036.

Table 1: Implied average effects of advertising

| | (1) baseline | (2) low jackpot | (3) high jackpot | (4) diff. (3)-(2) |
|---|---|---|---|---|
| first hour | 0.0171 | 0.0137 | 0.0178 | 0.0042 |
| | (0.0008) | (0.0012) | (0.0012) | (0.0017) |
| second hour | 0.0125 | 0.0122 | 0.0140 | 0.0018 |
| | (0.0011) | (0.0018) | (0.0016) | (0.0024) |
| third hour | 0.0062 | 0.0065 | 0.0096 | 0.0031 |
| | (0.0010) | (0.0018) | (0.0015) | (0.0023) |
| fourth hour | 0.0003 | 0.0012 | 0.0009 | -0.0004 |
| | (0.0007) | (0.0013) | (0.0012) | (0.0018) |
| total effect | 0.0361 | 0.0336 | 0.0423 | 0.0087 |
| | (0.0027) | (0.0048) | (0.0038) | (0.0061) |

Notes: This table shows implied percentage increases in sales for each of the first 4 hours after an advertisement was aired and the sum of these average effects. Average effects are calculated as the average of $\exp(\bar{\beta}_\ell)$ minus one. See also (4). The table also shows standard errors in parentheses. These were calculated using the delta method and account for correlations of error terms within days. In column (1) we report the results for the baseline model described in Section 6.2. In column (2) and (3) we report results by jackpot size, see Section 6.3. Column (4) reports results on the difference between (2) and (3). These were calculated using the delta method after estimating (2) and (3) jointly using a fully interacted model.

From these coefficient estimates, we can calculate implied effects on sales for each of the four hours after the advertisement was aired. Recall that we have estimated 12 coefficients $\bar{\beta}_\ell$ per hour. Denote those estimates by $\hat{\bar{\beta}}_\ell$. For the first hour, we calculate[26]

$$\left[\frac{1}{12}\sum_{\ell=1}^{12}\exp(\hat{\bar{\beta}}_\ell)\right] - 1.$$

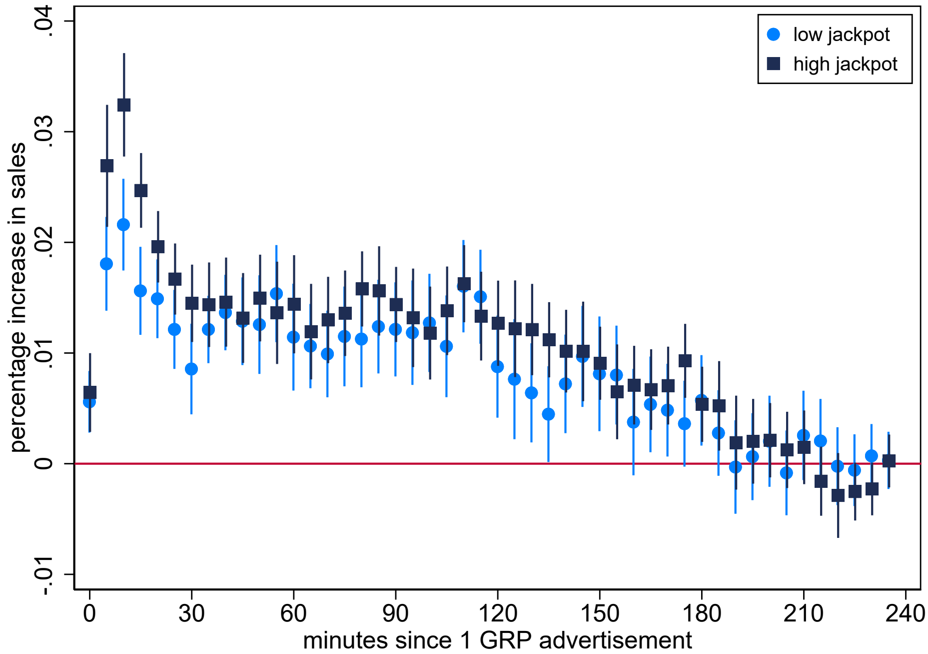
The average effects for the other three hours are calculated accordingly. Table 1 shows the results. The effect of advertising is a statistically significant increase in sales by about 1.7 percent of the baseline sales in the first hour. This compares to the 1.4 percent per GRP that we have measured for big advertisements. In the second hour it is 1.3 percent, in the third 0.6. In the fourth hour, it is not statistically significantly different from zero. The total effect is the sum of these four effects. We find that one GRP of advertising increases sales on average by 3.6 percent of the sales of one hour. The remaining columns of the table are discussed below.

## 6.3 Does the effect of advertising depend on the jackpot size?

So far, we have estimated reduced-form advertising effects, $\beta_\tau$. We have found them to be strong and to last for about three hours. Next, we investigate whether the effects depend on the

[26]This would be exactly the percentage change if our dependent variable was the natural logarithm of sales. However, we use the natural logarithm of one plus sales as the dependent variable. See also footnote 25.

Figure 8: Dependence of advertising effects on jackpot size



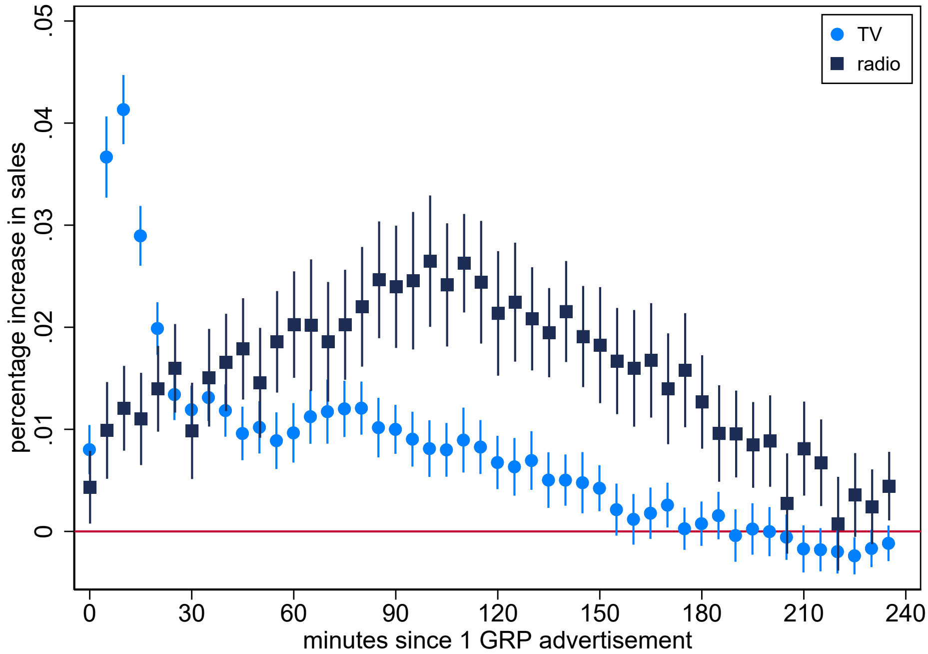

Notes: This figure shows estimates from a distributed lag model that was estimated at the minute level. The specification is the same as in Figure 7, except that we split the sample by jackpot size.

jackpot size. This could be the case, as advertisements also contained information about the jackpot size.

Recall that according to (1) in Section 5 we have that $\beta_\tau \equiv \frac{1}{100}[(1+\Delta_\tau)\cdot(1+\Gamma)-1]$. In words, the reduced-form advertising effect, $\beta_\tau$, is a function of the reminder effect, $\Delta_\tau$, which does depend on the time since the advertisement was aired, and the effect of the information contained in the advertisement, $\Gamma$, on $P_t(\text{buy}|\text{think})$. Information that lets consumers positively update their priors on the jackpot size may increase the probability to buy given that the consumer thinks about buying. Information that lets consumers negatively update their priors may lead to a decrease in that probability. Based on this, we now test the null hypothesis that $\Gamma$ does not depend on the jackpot size. We can do so by testing whether $\beta_\tau$ depends on the jackpot size.

For this, we split the sample by jackpot size. Figure 8 reports estimates for above and below median jackpots. The estimated effects are slightly higher for above-median jackpots, but the pointwise confidence intervals for both curves overlap most of the time. To test this formally, recall that we do not estimate 240 $\beta_\tau$'s for high and low jackpots, respectively, but always group 5 of those together, which means that we estimate 48 $\bar{\beta}_\ell$'s for high and low jackpots, respectively. Denote the 48 parameters for draws with a jackpot below the median by $\bar{\beta}_\ell^{\text{low}}$ and the 48 parameters for draws with a jackpot above the median by $\bar{\beta}_\ell^{\text{high}}$. We test the joint hypothesis that $\bar{\beta}_\ell^{\text{low}} = \bar{\beta}_\ell^{\text{high}}$ for all $\ell$. This hypothesis is rejected at conventional levels, with an $F$ statistic of 3.27 and a $p$-value of 0. This establishes that in a statistical sense, the two

Figure 9: TV vs. radio advertisements




Notes: This figure shows estimates from a distributed lag model that was estimated at the minute level. We estimate separate effects for TV and radio advertisements. This means that we include 96 instead of 48 lags. Otherwise, the specification is the same as in Figure 11.

estimated curves are different from one another.

To assess economic significance, we have calculated implied average percentage advertising effects for each of the four hours after the advertisement was aired, separately for above and below median jackpots. We have done so in the same way as in Section 6.2. Table 1 reports the results in column (2) and (3). Column (4) reports the difference between the two. The table also reports robust standard errors (clustered at the daily level). We can see that in the first hour the average effect is statistically significantly different between high and low jackpots. We estimate it to be 30 percent (0.42 percentage points) higher for high jackpots. The effects for the second, third and fourth hour are not statistically significantly different from one another, but point estimates are generally higher for high jackpots. Also the total effect is not statistically significantly different between high and low jackpots, but is estimated to be 26 percent (0.87 percentage points) higher. Overall, this suggests that *next to the reminder effect*, there is an effect of advertising on the inclination of consumers to buy given that they are thinking about buying, which originates in the information contained in the advertisements. However, it seems to be hard to estimate it precisely.

## 6.4 Effect by type of advertisement

The main question we study in this paper is whether advertising can also act as a reminder. So far, we have documented strong short-run effects of advertising both for a selected set of big advertisements and for all advertisements in our data. However, comparing Figure 6 for big advertisements to Figure 7 shows that the shape of the effect differs. If advertising acts mainly as a reminder, then a natural explanation for this finding is that big advertisements were predominantly aired on TV and that the effects for TV advertisements materialize quicker. The reason for this is that consumers are more likely to be at home when they watch TV and can therefore react faster once they are reminded.

To investigate this further, we have estimated a general version of the model in Section 6.2. Recall that (4) specifies a distributed lag model with explanatory variables $sumgrp_{\ell,t}$, which are 48 5-minute blocks of sums of GRP's in the past 4 hours. We generalize this model by calculating these explanatory variables separately for TV and radio advertisements. Denote them by $sumgrptv_{\ell,t}$ and $sumgrpradio_{\ell,t}$, respectively, with coefficients $\bar{\beta}_\ell^{\text{TV}}$ and $\bar{\beta}_\ell^{\text{radio}}$. The model we estimate is

$$\log(1+sales_t) = \sum_{\ell=1}^{48} \left( \bar{\beta}_\ell^{\text{TV}} sumgrptv_{\ell,t} + \bar{\beta}_\ell^{\text{radio}} sumgrpradio_{\ell,t} \right) + \alpha_{h(t)} + \varepsilon_t. \qquad (5)$$

Figure 9 shows the results. As expected, the effect of TV advertisements is more immediate and dies out faster.[27]

## 6.5 The dependence of advertising effects on the time until the draw

Reminder advertising can help consumers remember to buy a product. However, as pointed out and documented by Lambrecht et al. (2020), this does not necessarily mean that total sales increase as well. If consumers would have bought the product anyway at a later point in time, then advertising will only lead to purchase acceleration, but not to market expansion.

Our empirical setup offers the opportunity to study whether advertising leads *also* to market expansion by making use of the fixed ending time up to which lottery tickets for a particular draw can be bought. The idea is that if advertising has an effect until shortly before the draw, then this must be due to market expansion, as there is no later point in time at which a ticket can be bought for that draw.

To study whether this is the case, we estimate a version of our model that allows for a dependence of advertising effects on the time until the draw. The reason why we use a more general specification is that baseline sales are naturally higher shortly before the draw (see Figure 1) and our estimation equation (4) implies that short-term advertising effects are proportional to baseline sales. In our framework, proportionality arises naturally because advertising that acts

[27] In Figure A6 in the Online Appendix, we show additional results for combinations of jackpot size and type of advertisement. Table A3 reports implied total effects of advertising.

as a reminder affects the probability to think about buying, which multiplies the probability to buy given that the consumer thinks about buying. However, we do not want to impose that here.

The more general specification includes additional interaction terms between the days until the draw and the GRP lags, so that the model is flexible enough to also capture a situation where advertising has no effect shortly before the draw. Our estimation equation is

$$\log(1+sales_t) = \sum_{\ell=1}^{48} (\bar{\beta}_\ell + \gamma_{H(\ell),DTD(t)}) \cdot sumgrp_{\ell,t} + \alpha_{d(t)} + \alpha_{DTD(t)} + \alpha_{HOD(t)} + \varepsilon_{it}. \quad (6)$$

Here, $H(\ell)$ denotes the hour since the advertisement was aired, $DTD(t)$ denotes the days to the draw, $d(t)$ denotes the draw, and $HOD(t)$ the hour of day. The model differs from the baseline model, (4), in two ways. First, we use draw fixed effects, $\alpha_{d(t)}$, days-to-draw fixed effects, $\alpha_{DTD(t)}$, and hour of day fixed effects, $\alpha_{HOD(t)}$. This replaces the more flexible hour of year fixed effects. We use this slightly more restrictive specification, because it makes it easier to predict baseline sales in order to produce a figure where we plot absolute advertising effects against the days until the draw.[28]

The second and main way in which the model differs is that advertising effects are allowed to depend on the number of days until the draw. As in (4), we have 48 5-minute lag terms with coefficients $\bar{\beta}_\ell$. Now we have in addition 4 times 20 interaction terms between the lag length at the hourly level, $H(\ell)$, and the last 20 days until the draw. This is best illustrated with an example. Suppose we are interested in the effect of an advertisement on sales 90 to 94 minutes later, when there are 6 days until the draw. $\ell$ indexes time blocks with length 5 minutes, relative to the time of the advertisement (the first block takes minute 0 to 4, the second block 5 to 9, and so on). This means that 90 to 94 minutes later corresponds to $\ell = 19$. This is in the second hour after the advertisement was aired, so $H(\ell) = 2$. We are interested in the effect when there are 6 days until the draw, so $DTD(t) = 6$. Therefore, in this example, the effect of interest is given by $\bar{\beta}_{19} + \gamma_{2,6}$.

Figure 10 plots predicted *absolute* advertising effects for 1 GRP of advertising in the first hour after an advertisement was aired relative to the *absolute* effect on the day of the draw. These are given by $\exp\left(\alpha_{DTD(t)} + \gamma_{1,DTD(t)}\right)$.[29] The figure also shows respective 95% confidence intervals.[30] Towards the time of the draw, the effects increase, which suggests that advertising does *not only* lead to purchase acceleration, but *also* to market expansion.

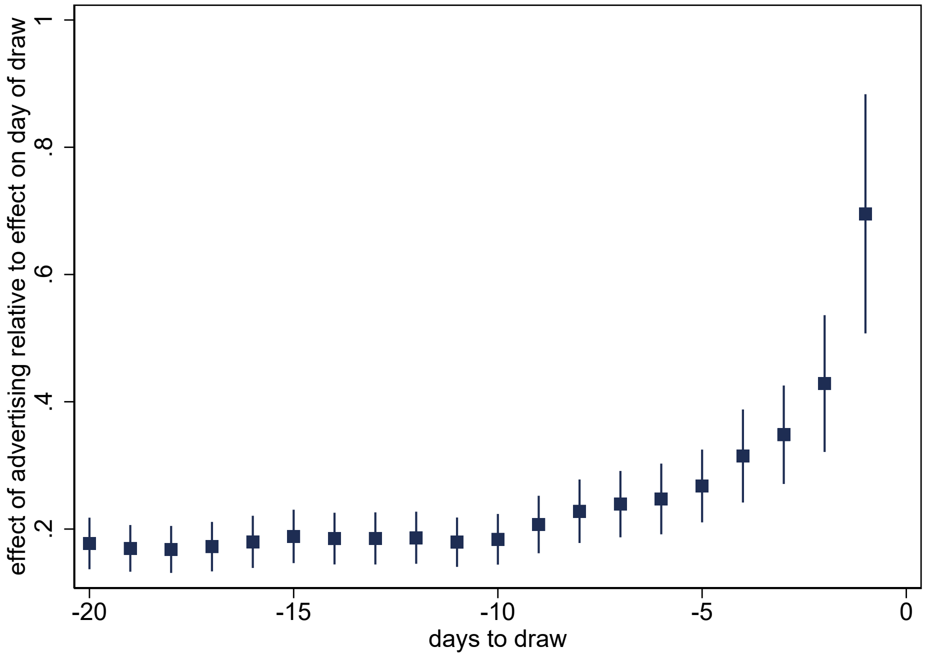
One can also see this analysis as a robustness check related to the proportionality built into

[28]This is supported by our first robustness check in Section 6.6 below.

[29]We normalize $\gamma_{H(\ell),DTD(t)} = 0$ for the day of the draw, $\alpha_{d(t)} = 0$ for the first draw, $\alpha_{DTD(t)} = 0$ for the day of the draw, and $\alpha_{HOD(t)} = 0$ for the first hour of the day. It follows from (6) that $1 + sales_t = \exp\left(\sum_{\ell=1}^{48} (\bar{\beta}_\ell + \gamma_{H(\ell),DTD(t)}) \cdot sumgrp_{\ell,t} + \alpha_{d(t)} + \alpha_{DTD(t)} + \alpha_{HOD(t)} + \varepsilon_{it}\right)$. The normalizations imply that on the day of the draw, keeping everything else equal, this is $1+sales_t = \exp\left(\sum_{\ell=1}^{48} \bar{\beta}_\ell \cdot sumgrp_{\ell,t} + \alpha_{d(t)} + \alpha_{HOD(t)} + \varepsilon_{it}\right)$. Taking the ratio of the two and evaluating it for one GRP in the hour before $t$ gives $\exp\left(\alpha_{DTD(t)} + \gamma_{1,DTD(t)}\right)$.

[30]We have also tried to “zoom in” and show that the effect is there in the very last hours before the draw, but we only have data for 16 draws, with a limited number of advertisements in the last hours before the draw.

Figure 10: Effect of timing




Notes: This figure shows the absolute effect of advertising on the 20 days before the draw, relative to the absolute effect on the day of the draw. See text for details.

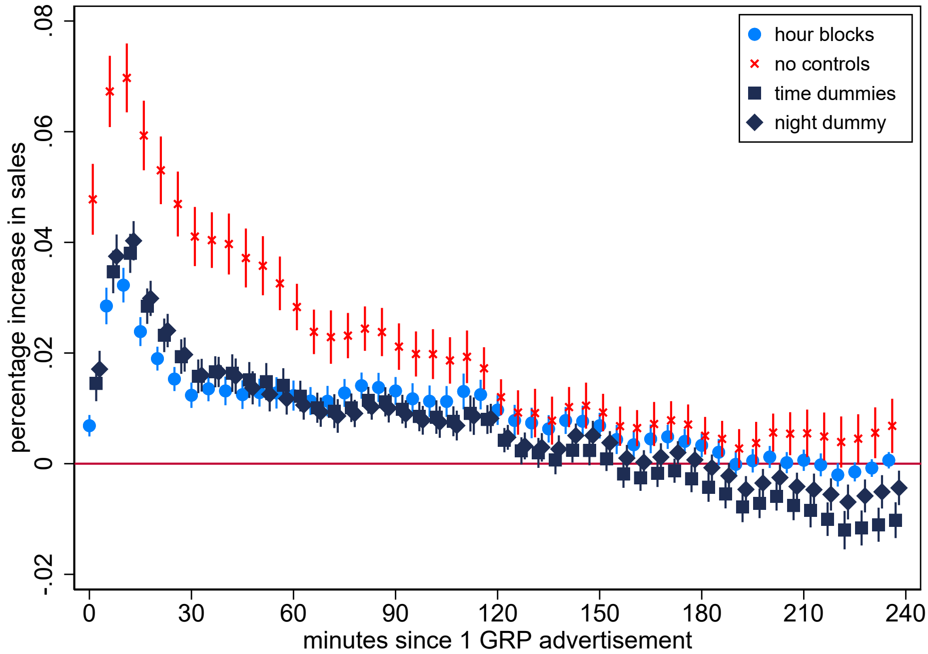
our framework and our baseline specification (4). Figure A11 in the Online Appendix shows that the *percentage* effect of advertising, $\gamma_{1,DTD(t)}$, does *not* depend statistically significantly on the time until the draw. This means that advertising effects are approximately proportional to baseline sales.

## 6.6 Robustness

In this section, we report results from two robustness checks and provide a brief overview over additional robustness checks that we report on in the Online Appendix.

Sales and advertising levels are higher on certain days and at certain times of the day. Therefore, it is important to control for time effects. Figure 11 shows how the estimated advertising effect depends on the way we control for time effects. Corresponding coefficient estimates for the first 12 lags are reported in Table A2 in the Online Appendix. We expect that advertising effects are overestimated if we do not control for time effects at all. The figure confirms that. At the same time, we see that advertising effects are relatively similar when we control either for hour of year fixed effects, as we do in our baseline specification in Figure 7, time dummies, as we do in Figure 10, and even when we only control for days to draw and a night dummy that takes on the value one in the hours between midnight and 7am. This last specification is inspired by Figure 2 showing that there is a big difference in both sales and advertising levels between day and night. The results suggest that controlling for the time of the day in terms of

Figure 11: Estimates for varying sets of controls



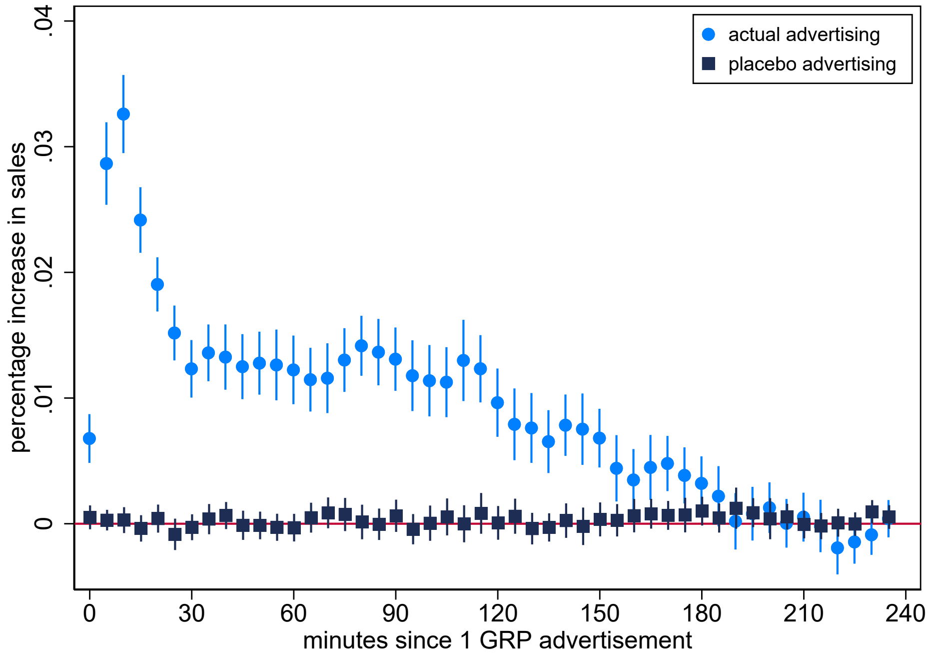

Notes: This figure shows estimates from a distributed lag model that was estimated at the minute level. There are 48 lags that are each for 5 minute intervals, so 4 hours in total. The four curves are for four different sets of controls: hour of year fixed effects, no controls, time dummies, and a night dummy. The dependent variable is the log of one plus sales. The bars indicate 95 percent confidence intervals, which are based on standard errors that are clustered at the daily level. Table A2 in the Online Appendix shows the coefficients for the first 12 lags.

day and night is sufficient to alleviate endogeneity concerns. We will make use of this when we estimate the structural model.

Our second robustness check is to report results for a set of placebo advertisements. We generate the placebo advertisements by first generating an indicator for there being an advertisement in a given minute. We take the likelihood that this is the case from the data, pooling over time. Then, we draw the number of GRP from a log normal distribution with the same mean and variance as the distribution of GRP's conditional on there being an advertisement in the data. This means that we will have twice as many advertisements in the new data than before. Starting from this we estimate a model that simultaneously estimates the effect of the actual advertisements and the placebo advertisements. Figure 12 shows the results. For this, we use the same specification as in Figure 7, but add 48 additional lag terms for the placebo advertisements. The effect of the placebo advertisements is estimated to be very close to zero, with very tight confidence intervals, whereas the estimated effect of the actual advertisements is very close to the one we have estimated above.

Taken together, these two robustness checks suggest that advertising effects are well-identified. In the Online Appendix, we present additional results and carry out a number of additional robustness checks. We show that aggregating data to the hourly level leads to similar estimates of

Figure 12: Effects of actual and placebo advertisements




Notes: This figure shows estimates from a distributed lag model that was estimated at the minute level. We jointly estimate the effects of the actual advertisements that were aired and a set of placebo advertisements that were placed at random times. See text for details.

advertising effects (Table A4); that estimating the model with shorter time blocks leads to similar results (Figure A7); that the estimated effect per GRP does not strongly depend on whether many or not so many consumers have been reached (Figure A8); that the effects for big TV advertisements are similar to the effects of all TV advertisements (Figure A9); that coefficients on lead terms are close to zero (Figure A10), which can be seen as an additional placebo check; and that percentage advertising effects do not on the time until the draw (Figure A11).

# 7 A model of lottery ticket demand

Our reduced-form analysis is well-suited to estimate the short-run effects of advertising. However, we cannot use it to conduct counterfactual experiments in which we change the advertising schedule for a longer time period. The reason is that our reduced-form analysis does not capture that consumers who are motivated to buy a ticket will be less inclined to buy another one in the future, even if they are reminded again.

In this section we estimate a very simple structural model to gain more insight into the medium-run effects of advertising. The model is static and the length of one period is one hour. Consumers stay in the market until they have either bought a ticket or the draw takes place. Therefore, the mechanism through which medium-run advertising effects arise is that consumers buy and leave the market instead of staying in the market and possibly buying a

Table 2: Effect of various advertising strategies

| strategy | sales |
|---|---|
| data (reference point) | 100% |
| no advertising at all | 84.3% |
| spreading advertisements equally in the last 4 days before the draw | 115.5% |
| shift advertising from third week to fourth week | 110.8% |
| shift advertising from fourth week to third week | 87.7% |

Notes: This table shows the effect of using alternative dynamic advertising strategies for the February draw. See text for a description of these strategies. Simulations are based on the parameter estimates reported in Table A5.

ticket.

The probability to think about buying a ticket is given by a simple logit specification,

$$P_{it}(\text{think}) = \frac{1}{1 + \exp\left(-(\gamma_0 + \gamma_1 g_{it}^a)\right)}.$$

Here, advertising enters through an advertising goodwill stock $g_{it}^a$ whose evolution we simulate for a hypothetical set of consumers. It depreciates from period to period with rate $\lambda$. Our data are informative about how many consumers are reached by an advertisement. This information is used to randomly select the appropriate number of simulated consumers whose goodwill stock is in addition increased by one unit, respectively.

The probability to buy given that the consumer thinks about buying is specified as

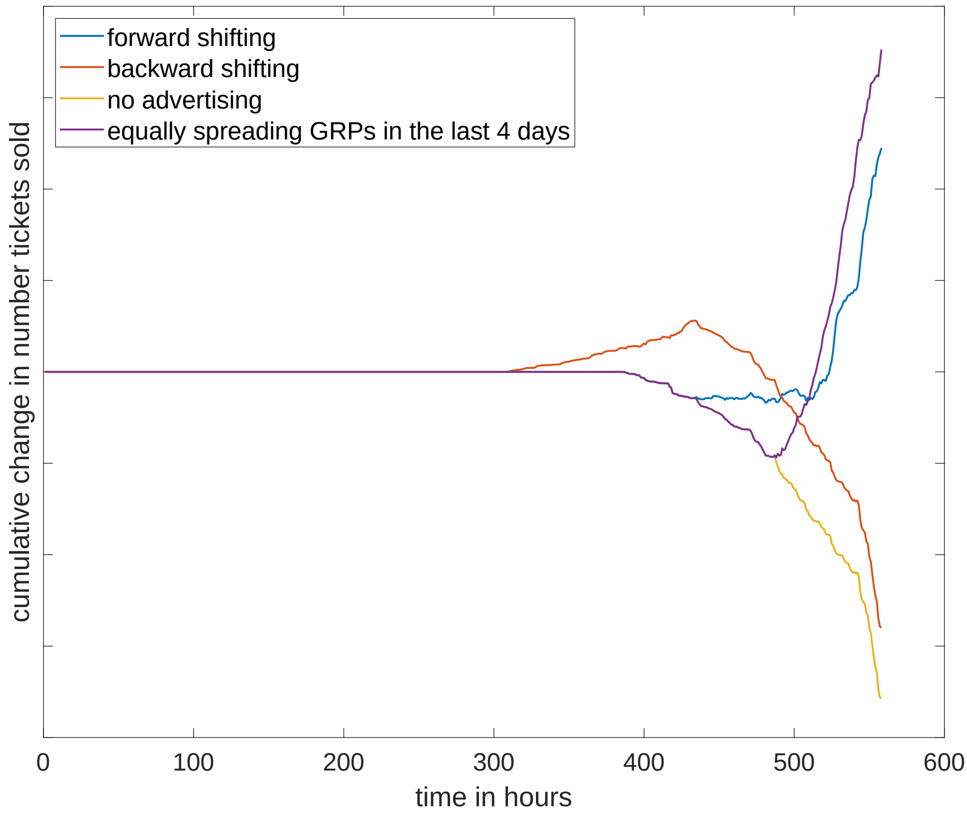
$$P_{it}(\text{buy}|\text{think}) = \frac{\exp\left(\frac{-p+\delta^{T-t}\psi}{\sigma}\right)}{\exp\left(\frac{-p+\delta^{T-t}\psi}{\sigma}\right)+1}.$$

Here, we normalize the price coefficient to one.

We estimate the model using the method of simulated moments.[31] We now return to the question what the effect of alternative dynamic advertising schedules would have been, including the case of no advertising at all. We do not have access to data on the profitability of an additional sold ticket, and also not on the cost of one GRP. It is, however, not unreasonable in our context to assume as an approximation that the price of one GRP does not vary over time, especially because the firm decided to not buy advertising at pre-specified times and regular draws are on the 10th day of the month, independent of the weekday. In that sense it is mean-

[31] We provide additional technical details in Online Appendix C.1, C.2, and C.3. Online Appendix C.4 discusses identification of the structural parameters. Technical details related to estimation are discussed in Online Appendix C.5. Online Appendix C.6 reports parameter estimates and shows that the model fits the data well. Online Appendix C.7 shows how we can use the model to calculate the elasticity of sales with respect to advertising. We find that it is 0.130, which is higher than in some other contexts. We discuss potential reasons for this. In Online Appendix C.8 we assess the robustness to making alternative assumptions on the market size and viewership behavior.

Figure 13: Effect of different advertising strategies




Notes: This figure shows the difference between the cumulative number of tickets sold at each point in time for a counterfactual advertising strategy and the cumulative number of tickets sold given the actual advertising schedule for the February draw. Based on parameter estimates reported in Table A5.

ingful to ask the question whether it is possible to sell more tickets when one allocates the same number of GRP's in a different way.

We consider 4 alternative strategies and compare the total number of tickets sold to the simulated one for the original GRP schedule in the data. The first alternative strategy is to remove all advertising. Comparing sales under this strategy to baseline sales will allow us to estimate what the overall effect of advertising was. In the second alternative strategy, we allocate all advertising to the last four days before the draw and distribute it equally over all hours on those four days. This allows us to estimate the effect of shifting advertising to times at which consumers are more likely to buy given that the consumer thinks about buying. This counterfactual strategy is however completely artificial and does not take scheduling constraints into account. To gain some insight into the effect of more realistic changes, the last two counterfactual strategies take, respectively, the schedule as it is in the third week and move it to the fourth week (i.e. adding it to the one in the fourth week), and *vice versa*.

Table 2 shows the results. Not advertising at all leads to 84.3 percent of the original sales. Generally, allocating advertising to later points in time increases sales. The most drastic measure we considered was to move it to the last four days before the draw. This leads to an increase in sales of 15.5 percent. The more realistic experiment of moving advertising from the third to the fourth week still leads to an increase in sales of 10.8 percent. The converse, namely to move advertising from the fourth week to the third, lead to a decrease in sales by 12.3 percent.

Figure 13 shows the underlying dynamics. We plot the *difference* between the cumulative sales for a given strategy and the baseline strategy. As an illustration, consider the strategy of shifting all the GRP's from the fourth week to the third week. As expected, sales in the third week increase faster than sales for the baseline strategy (hence the difference in cumulative sales is positive). Thereafter, they fall behind (hence the difference is negative). Overall, fewer tickets are sold, which is reflected in the lower end point.

# 8 Concluding remarks

In 2018, global advertising spending amounted to 552 billion US dollars. 208 billion of this were spent in the US, which is about 1 percent of GDP.[32] There is still no consensus among academic scholars on how one should think about this in general. Possible reasons include that the effect of advertising is highly context-specific and depends on the type of advertisement.

Usually, a distinction is made between informative advertising, persuasive advertising, and advertising that acts as a complement to consumption (Bagwell, 2007). In this paper, we provide empirical support for the view that advertising can also act as a reminder when consumers suffer from limited attention. For this, we focus on one particular context—online sales of lottery tickets in the Netherlands—that is particularly well-suited for this purpose. We propose a new framework in which advertising can act as a reminder and develop an empirical approach that allows us to isolate reminder effects from other effects of advertising, in particular those influencing consumer choice because advertisements contain information about the jackpot size. We then use high frequency data to credibly identify and estimate the short-run reminder effects of advertising. We also propose a simple structural model with an attention stage and estimate its parameters to gain insights into medium-run effects.

We show that reaching one percent of the population leads to an increase in online sales of 1.7 percent in the first hour after the advertisement is aired. This can be attributed to advertising acting as a reminder in the short run. We show that reminder advertising can lead to market expansion and that shifting reminder advertising to times at which consumers want to be reminded can have substantial effects on sales. Our counterfactual predictions suggest that shifting advertising to the week of the draw can lead to an increase in total online sales of 10.8 percent.

The context of our study is the market for lottery tickets that are bought online. This context is particularly helpful for obtaining model-free evidence on the effects of advertising and to

[32]Taken from the winter 2018 update to the MAGNA advertising forecast, available at https://magnaglobal.com/magna-advertising-forecasts-winter-2018-update/ (accessed May 2019). More than half of global spending, 301 billion US dollars, was on non-digital advertising that includes TV, radio and print advertising.

attribute the effects to advertising acting as a reminder. Focusing on this context however means that our results do not directly generalize. This opens up opportunities for future research. Our conjecture is that next to other effects, such as informing consumers, advertising generally draws attention to thinking about buying a product, which fades away quickly, and that the short-run effect of advertising is therefore closely related to how easy it is for consumers to buy the product once they think about doing so. In our context, it is very easy to buy a lottery ticket online after a consumer sees an advertisement on TV. Related to that, thinking about feature and displays in stores as reminding consumers could at least partly explain why they have large effects. Also here, it is easy for consumers to buy the product once they think about doing so. It would be interesting to study more broadly in which settings and for which products reminder advertising is particularly successful.

Another follow-up to our study could be a comprehensive content analysis of TV and radio advertisements that measures how much of the impressions are related to advertising conveying information, changing consumer preferences, acting as a complement to consumption, or reminding them to turn an intended purchase into an actual one. We also hope that more light can be shed on the question to what extent slogans such as “Drink Coca Cola” and “Just do it.” are actually reminders rather than persuasion or a complement to consumption.

Reminder advertising stimulates consumers to act on their preferences. By providing evidence for this pathway we hope to stimulate more work on the topic. This seems important to us, as many fundamental results in economics and marketing hinge on the presumption that consumers actually act on their preference.

# References


Abaluck, J. and A. Adams-Prassl (2021). What do consumers consider before they choose? identification from asymmetric demand responses. *Quarterly Journal of Economics 136*(3), 1611–1663.

Abbring, J. H. and Ø. Daljord (2020). Identifying the discount factor in dynamic discrete choice models. *Quantitative Economics 11*(2), 471–501.

Ackerberg, D. A. (2001). Empirically distinguishing informative and prestige effects of advertising. *RAND Journal of Economics*, 316–333.

Ackerberg, D. A. (2003). Advertising, learning, and consumer choice in experience good markets: an empirical examination*. *International Economic Review 44*(3), 1007–1040.

Allenby, G. M. and J. L. Ginter (1995). The effects of in-store displays and feature advertising on consideration sets. *International Journal of Research in Marketing 12*(1), 67–80.

Bagwell, K. (2007). The economic analysis of advertising. In M. Armstrong and R. Porter (Eds.), *Handbook of Industrial Organization*, Volume 3, pp. 1701–1844. Elsevier.

Blattberg, R. C., R. Briesch, and E. J. Fox (1995). How promotions work. *Marketing Science 14*(3), G122–G132.

Bronnenberg, B. J. and W. R. Vanhonacker (1996). Limited choice sets, local price response, and implied measures of price competition. *Journal of Marketing Research 33*(2), 163–173.

Calzolari, G. and M. Nardotto (2016). Effective reminders. *Management Science 63*(9), 2915–2932.

Chandrasekaran, D., R. Srinivasan, and D. Sihi (2018). Effects of offline ad content on online brand search: Insights from super bowl advertising. *Journal of the Academy of Marketing Science 46*(3), 403–430.

De Groote, O. and F. Verboven (2019). Subsidies and time discounting in new technology adoption: Evidence from solar photovoltaic systems. *American Economic Review 109*(6), 2137–72.

DellaVigna, S. and M. Gentzkow (2010). Persuasion: empirical evidence. *Annual Review of Economics 2*(1), 643–669.

Draganska, M. and D. Klapper (2011). Choice set heterogeneity and the role of advertising: An analysis with micro and macro data. *Journal of Marketing Research 48*(4), 653–669.

Du, R. Y., L. Xu, and K. C. Wilbur (2019). Immediate responses of online brand search and price search to tv ads. *Journal of Marketing 83*(4), 81–100.

Dubé, J.-P., G. J. Hitsch, and P. Manchanda (2005). An empirical model of advertising dynamics. *Quantitative Mareting and Economics 3*(2), 107–144.

Gabaix, X. (2019). Behavioral inattention. In *Handbook of Behavioral Economics: Applications and Foundations 1*, Volume 2, pp. 261–343. Elsevier.

Hartmann, W. R. and D. Klapper (2017). Super bowl ads. *Marketing Science 37*(1), 78–96.

Heiss, F., D. McFadden, J. Winter, A. Wuppermann, and B. Zhou (2021). Inattention and switching costs as sources of inertia in medicare part d. *American Economic Review 111*(9), 2737–81.

Ho, K., J. Hogan, and F. Scott Morton (2017). The impact of consumer inattention on insurer pricing in the medicare part d program. *RAND Journal of Economics 48*(4), 877–905.

Hortaçsu, A., S. A. Madanizadeh, and S. L. Puller (2017). Power to choose? an analysis of choice frictions in the residential electricity market. *American Economic Journal: Economic Policy 9*(4), 192–226.

Hu, Y., L. M. Lodish, and A. M. Krieger (2007). An analysis of real world TV advertising tests: A 15-year update. *Journal of Advertising Research 47*(3), 341.

Inman, J. J. and L. McAlister (1994). Do coupon expiration dates affect consumer behavior? *Journal of Marketing Research 31*(3), 423–428.

Joo, M., K. C. Wilbur, B. Cowgill, and Y. Zhu (2014). Television advertising and online search. *Management Science 60*(1), 56–73.

Joo, M., K. C. Wilbur, and Y. Zhu (2016). Effects of TV advertising on keyword search. *International Journal of Research in Marketing 33*(3), 508–523.

Klein, T. J. and C. He (2020). Advertising as a reminder: Evidence from the Dutch State Lottery. C.E.P.R. Discussion Paper 12948, C.E.P.R., London, U.K.

Krugman, H. E. (1972). Why three exposures may be enough. *Journal of Adverting Research 12*(6), 11–14.

Lambrecht, A., C. E. Tucker, and X. Zhang (2020). TV advertising and online sales: The role of inter-temporal substitution. Working Paper, London Business School, London, U.K.

Lewis, R. A. and D. H. Reiley (2013). Down-to-the-minute effects of super bowl advertising on online search behavior. In *Proceedings of the fourteenth ACM conference on Electronic commerce*, pp. 639–656. ACM.

Liaukonyte, J., T. Teixeira, and K. C. Wilbur (2015). Television advertising and online shopping. *Marketing Science 34*(3), 311–330.

Lodish, L. M., M. Abraham, S. Kalmenson, J. Livelsberger, B. Lubetkin, B. Richardson, and M. E. Stevens (1995). How TV advertising works: A meta-analysis of 389 real world split cable tv advertising experiments. *Journal of Marketing Research*, 125–139.

Loewenstein, G., D. Read, and R. F. Baumeister (2003). *Time and decision: Economic and psychological perspectives of intertemporal choice*. Russell Sage Foundation.

Magnac, T. and D. Thesmar (2002). Identifying dynamic discrete decision processes. *Econometrica 70*(2), 801–816.

Manzini, P. and M. Mariotti (2014). Stochastic choice and consideration sets. *Econometrica 82*(3), 1153–1176.

McFadden, D. (1989). A method of simulated moments for estimation of discrete response models without numerical integration. *Econometrica 57*(5), 995–1026.

Mehta, N., X. Chen, and O. Narasimhan (2008). Informing, transforming, and persuading: Disentangling the multiple effects of advertising on brand choice decisions. *Marketing Science 27*(3), 334–355.

Melnikov, O. (2013). Demand for differentiated durable products: The case of the us computer printer market. *Economic Inquiry 51*(2), 1277–1298.

Newey, W. and D. McFadden (1994). Large sample estimation and hypothesis testing. In R. F. Engle and D. McFadden (Eds.), *Handbook of Econometrics*, Volume 4, pp. 2112–2245. Amsterdam: North Holland.

Roberts, J. H. and J. M. Lattin (1991). Development and testing of a model of consideration set composition. *Journal of Marketing Research*, 429–440.

Roberts, J. H. and J. M. Lattin (1997). Consideration: Review of research and prospects for future insights. *Journal of Marketing Research*, 406–410.

Rossiter, J. R. and L. Percy (1987). *Advertising and promotion management.* McGraw-Hill.

Sahni, N. S. (2015). Effect of temporal spacing between advertising exposures: Evidence from online field experiments. *Quantitative Marketing and Economics 13*(3), 203–247.

Shapiro, B. T. (2018). Positive spillovers and free riding in advertising of prescription pharmaceuticals: The case of antidepressants. *Journal of Political Economy 126*(1), 381–437.

Shapiro, B. T., G. J. Hitsch, and A. E. Tuchman (2021). TV advertising effectiveness and profitability: Generalizable results from 288 brands. *Econometrica 89*(4), 1855–1879.

Shapiro, J. M. (2006). A "memory-jamming" theory of advertising. Working Paper, University of Chicago, Chicago, IL.

Sovinsky Goeree, M. (2008). Limited information and advertising in the US personal computer industry. *Econometrica 76*(5), 1017–1074.

Stephens-Davidowitz, S., H. Varian, and M. D. Smith (2017). Super returns to super bowl ads? *Quantitative Marketing and Economics 15*(1), 1–28.

Terui, N., M. Ban, and G. M. Allenby (2011). The effect of media advertising on brand consideration and choice. *Marketing Science 30*(1), 74–91.

Van Nierop, E., B. Bronnenberg, R. Paap, M. Wedel, and P. H. Franses (2010). Retrieving unobserved consideration sets from household panel data. *Journal of Marketing Research 47*(1), 63–74.

Zigmond, D. and H. Stipp (2010). Assessing a new advertising effect: Measurement of the impact of television commercials on internet search queries. *Journal of Advertising Research 50*(2), 162–168.

# Online Appendix

# A Additional descriptive evidence

Table A1: Differences across draws

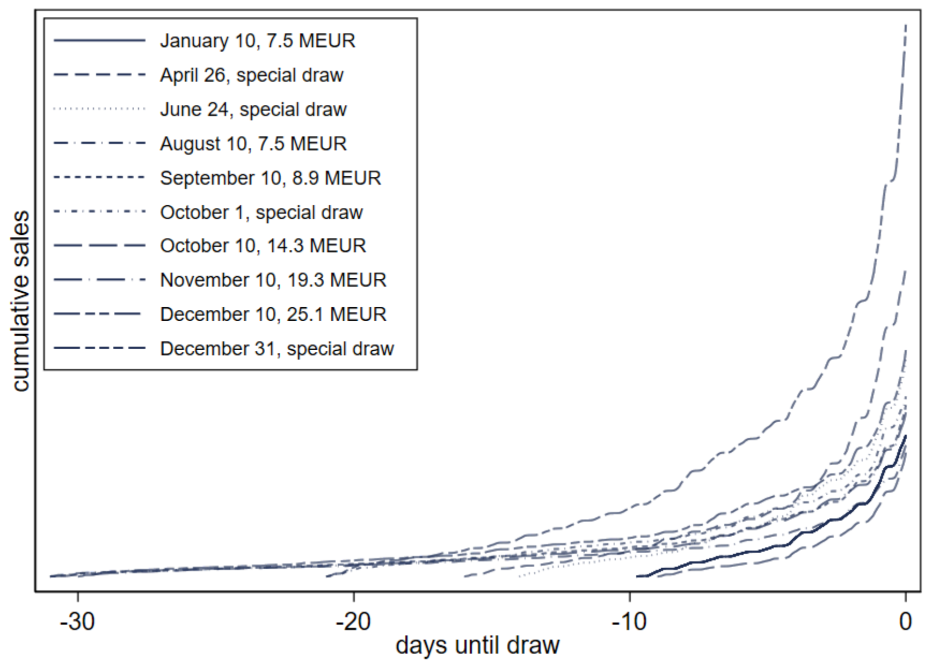
| | (1)<br>all draws | (2)<br>regular draws | (3)<br>special draws | (4)<br>all draws |
|---|---|---|---|---|
| log jackpot size | 0.367* | 0.366*** | | 0.507* |
| | (0.178) | (0.107) | | (0.248) |
| special draw | 1.511*** | | | 2.121*** |
| | (0.493) | | | (0.629) |
| log number of days | 0.173 | 0.147 | 0.722 | 0.430 |
| | (0.171) | (0.104) | (1.638) | (0.547) |
| log jackpot size previous draw | | | | -0.240 |
| | | | | (0.294) |
| special draw in previous draw | | | | -0.0622 |
| | | | | (0.971) |
| log number GRP previous draw | | | | 0.962 |
| | | | | (0.567) |
| Observations | 16 | 12 | 4 | 15 |
| $R^2$ | 0.560 | 0.602 | 0.089 | 0.730 |

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

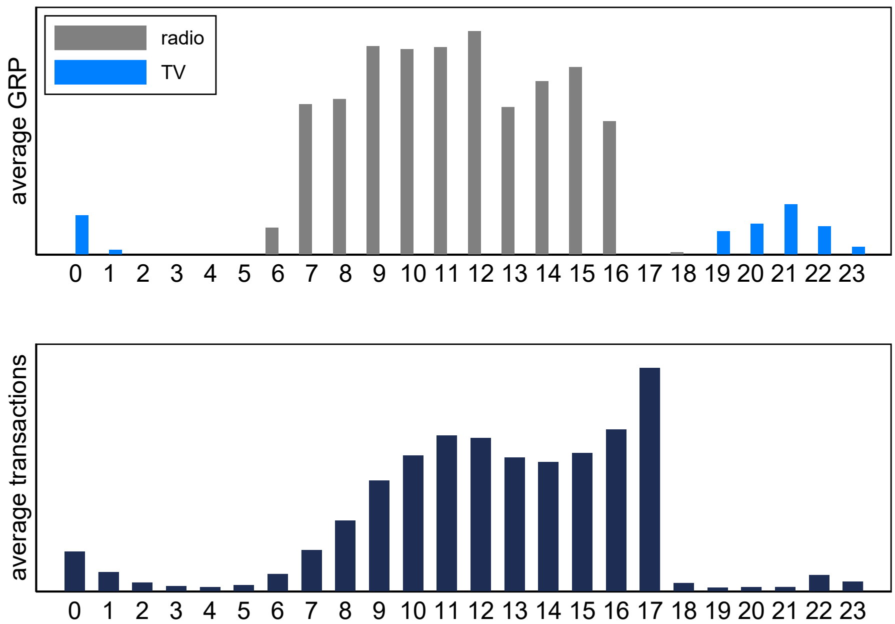
Notes: This table shows the results of a regression of the log of total sales on the total number of days on which tickets could be bought and the jackpot size if the draw was regular. In column (1) and (4) we pool across regular and special draws and set the log of the jackpot size to zero for the latter. One observation is one draw. There are only 15 observations for the last specification because we lack data on the previous draw for the first one that is in our data.

Figure A1: Cumulative sales for remaining draws



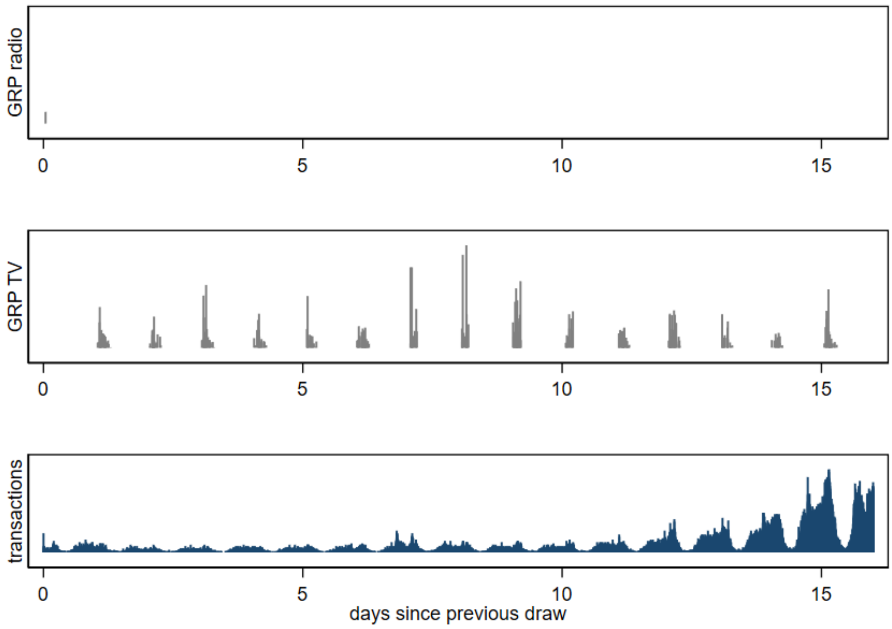

Notes: Figure 1 shows cumulative sales for 6 selected regular draws. This figure shows them for the remaining draws.

Figure A2: Advertising and sales during the day of the draw



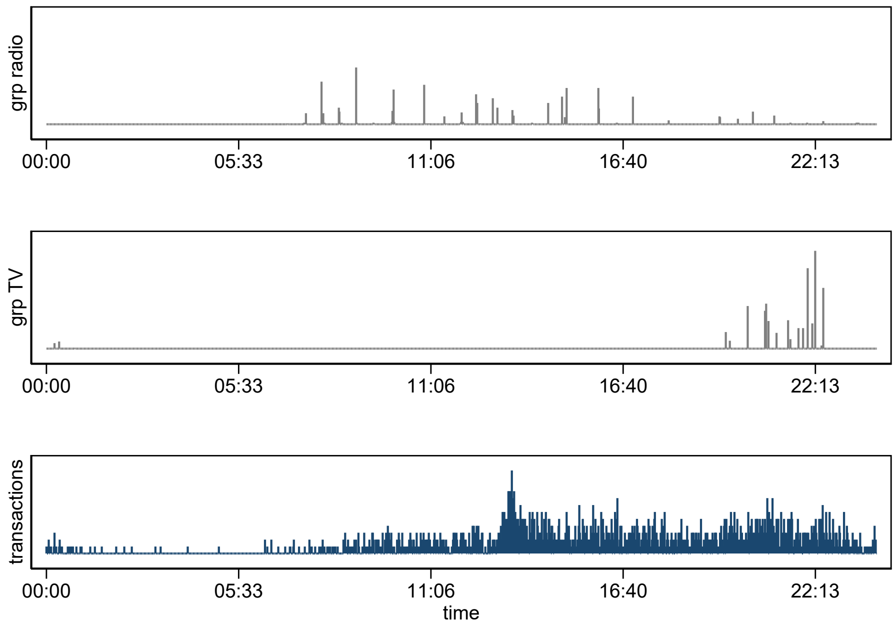

Notes: This figure shows average GRP's and sales for different times of the day. To produce this figure we first aggregate sales at the hourly level and then average over draws. On the day of the draw tickets for this draw can only be bought until 6pm. See Figure 2 for the pattern on the remaining days.

Figure A3: GRP's at the minute-level for a special draw



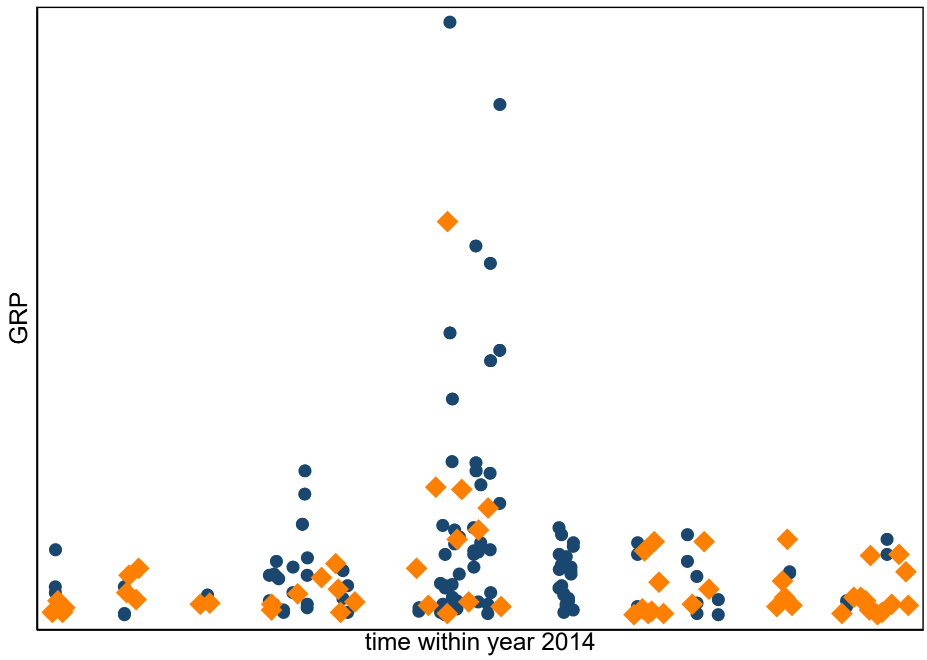

Notes: Figure 3 shows GRP's and sales for the draw on April 10, 2014. This figure shows GRP's and sales at the minute level for the special draw on April 26, 2014 (King's Day). The last regular draw took place on April 10, 2014. Tickets for the next draw can be bought from 6pm on the day of the previous draw, which is depicted as 0 days since the previous draw.

Figure A4: GRP's at the minute-level for an entire day




Notes: This figure shows GRP's and sales at the minute level for the entire day on April 3, 2014.

Figure A5: Advertisements that were used to construct Figure 6




Notes: This figure shows which advertisements were used in the sample for Figure 6. It shows a dot for each advertisement with at least 9 GRP, with the number of GRP's plotted against time. The diamonds are the advertisements that were used.

# B Additional results related to the reduced-form analysis

## B.1 Coefficient estimates for first hour

Table A2 shows coefficient estimates for the four models in Figure 11 in the main text. We report the first 12 estimates. 36 additional ones have been included, but are omitted here. Note that the number of observations for the model with hour of year fixed effects is one less. The reason for this is that there was one singleton group that was dropped.

## B.2 Additional results on heterogeneity of advertising effects

Figure A6 shows advertising effect separately for TV and radio advertising and by jackpot size. Table A3 shows implied effect sizes.

Table A2: The effect of advertising on sales

| | (1)<br>hour blocks | (2)<br>no controls | (3)<br>time dummies | (4)<br>night dummy |
|---|---|---|---|---|
| GRP between 0 and 4 minutes ago (lag) | 0.00686*** | 0.0478*** | 0.0145*** | 0.0171*** |
| | (0.000985) | (0.00325) | (0.00162) | (0.00168) |
| 5 and 9 minutes | 0.0285*** | 0.0673*** | 0.0347*** | 0.0375*** |
| | (0.00168) | (0.00327) | (0.00196) | (0.00200) |
| 10 and 14 minutes | 0.0323*** | 0.0697*** | 0.0380*** | 0.0403*** |
| | (0.00159) | (0.00316) | (0.00180) | (0.00181) |
| 15 and 19 minutes | 0.0239*** | 0.0593*** | 0.0285*** | 0.0299*** |
| | (0.00133) | (0.00320) | (0.00163) | (0.00162) |
| 20 and 24 minutes | 0.0190*** | 0.0530*** | 0.0232*** | 0.0241*** |
| | (0.00110) | (0.00311) | (0.00154) | (0.00153) |
| 25 and 29 minutes | 0.0153*** | 0.0469*** | 0.0194*** | 0.0197*** |
| | (0.00111) | (0.00298) | (0.00159) | (0.00157) |
| 30 and 34 minutes | 0.0124*** | 0.0410*** | 0.0158*** | 0.0160*** |
| | (0.00116) | (0.00272) | (0.00155) | (0.00150) |
| 35 and 39 minutes | 0.0135*** | 0.0404*** | 0.0166*** | 0.0166*** |
| | (0.00115) | (0.00256) | (0.00143) | (0.00139) |
| 40 and 44 minutes | 0.0132*** | 0.0397*** | 0.0164*** | 0.0158*** |
| | (0.00132) | (0.00280) | (0.00173) | (0.00163) |
| 45 and 49 minutes | 0.0125*** | 0.0372*** | 0.0151*** | 0.0137*** |
| | (0.00132) | (0.00271) | (0.00165) | (0.00152) |
| 50 and 54 minutes | 0.0128*** | 0.0358*** | 0.0148*** | 0.0125*** |
| | (0.00128) | (0.00271) | (0.00170) | (0.00157) |
| 55 and 59 minutes | 0.0128*** | 0.0326*** | 0.0142*** | 0.0117*** |
| | (0.00144) | (0.00247) | (0.00156) | (0.00146) |
| 36 additional GRP lags | Yes | Yes | Yes | Yes |
| hour of year dummies | Yes | No | No | No |
| draw dummies | No | No | Yes | Yes |
| days to draw dummies | No | No | Yes | Yes |
| hour of day dummies | No | No | Yes | No |
| night dummy | No | No | No | Yes |
| Observations | 525000 | 525001 | 525001 | 525001 |
| $R^2$ | 0.772 | 0.195 | 0.626 | 0.598 |

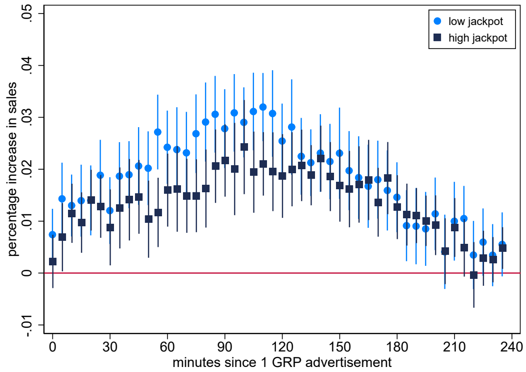
Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

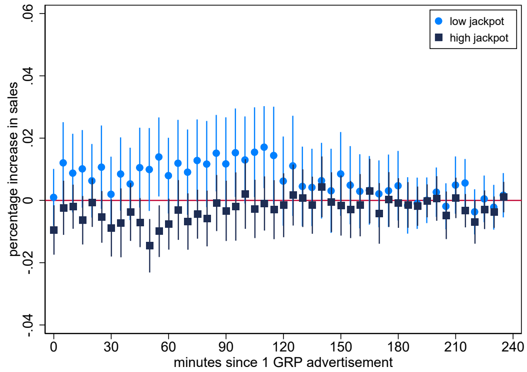
Notes: This table shows the results of regressions of the log of one plus sales on GRP's of advertising and lags thereof. Regressions were carried out at the minute level and standard errors are clustered at the daily level. The specifications differ by the set of control variables. Specification (1) is most flexible and uses hour of year dummies. Specification (2) uses no controls. Specification (3) uses draw, days to draw, and hour of day dummies. Specification (4) is like specification (3), with the only difference that a night dummy is used instead of the hour of day dummies. The night dummy takes on value 1 in the time between midnight and 7am.

Figure A6: Heterogeneity of advertising effects

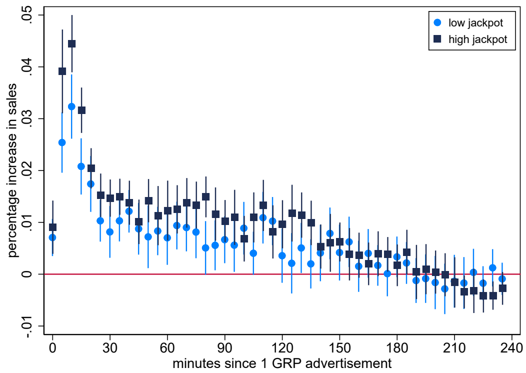
(a) radio AM




(b) radio PM




(c) TV




Notes: This figure shows estimates from a distributed lag model that was estimated at the minute level. We estimate the effects by type of advertising and time of the day. The first two sub-figures are for radio advertising in the morning (8am until Noon) and afternoon/evening (Noon until 11pm). The third sub-figure for tv advertising was estimated using data from 8am until 11pm. Estimates were obtained jointly using a fully interacted model. We restricted the effect of radio advertising to last for a maximum of 4 hours and the effect of TV advertising to last for a maximum of 2 hours.

Table A3: Total effects of advertising by type of advertising and jackpot size

| | (1) all draws | (2) low jackpot | (3) high jackpot | (4) diff. (3)-(2) |
|---|---|---|---|---|
| radio | 0.0619 | 0.0748 | 0.0551 | -0.0198 |
| | (0.0065) | (0.0093) | (0.0078) | (0.0123) |
| TV | 0.0312 | 0.0248 | 0.0373 | 0.0125 |
| | (0.0026) | (0.0048) | (0.0048) | (0.0068) |

Notes: This table shows implied total effects of advertising. They were calculated in the same way as the effects in the last row of Table 1. See notes to that table for details. Estimates were obtained jointly using a fully interacted model. We restricted the effect of radio advertising to last for a maximum of 4 hours and the effect of TV advertising to last for a maximum of 2 hours.

## B.3 Robustness

We also estimated the model at the hourly level. Table A4 shows the results. We report results for both a specification with hour of day dummies and a night dummy. The latter corresponds to our specification for the structural model that we estimate at the hourly level and where we treat all hours in the night as one. For both specifications at the hourly level we find that advertising has a similar effect in the hour in which it is aired as it has in the following hour: on average, one GRP of advertising leads to about a 1.2 percent increase in the amount of tickets sold. The effect is about one third of this two and three hours after the advertisement was aired, respectively. Comparing Table A2 to Table A4 shows that if anything aggregating the data to the hourly level will lead to slightly smaller estimated effects and in that sense estimates from hourly data will be conservative. This is important, because it would not be feasible to estimate a structural model at the minute level.

Our next robustness check concerns the length of the time blocks we control for. Figure A7 shows that results are very similar when we control for half hour or two hour blocks instead of the one hour blocks (hour of year fixed effects) in our baseline specification. This is not surprising, as Figure 11 already suggests that it makes not much of a difference whether we control for the days until the draw and a night dummy, or instead for hour of year fixed effects.

Our main results use a specification with the log of sales plus one as the dependent variable and lagged GRP's and controls as regressors. This means that we implicitly assume that the percentage change in sales is (roughly) proportional to the number of GRP's. To assess whether this is a good approximation we re-estimated the model allowing for departures from this. The most straightforward way to do this is to estimate the model with additional lags. In our baseline specification that underlies Figure 7, we use 12 lags per hour, for 4 hours, so 48 lags in total. The model we estimate here contains 96 lags, 48 related to advertisements with less than 14 GRP and 48 related to advertisements with at least 14 GRP. We chose to make the split at 14 because the average GRP in a 5 minute block that contains one of the big advertisements that we used to create Figure 6 was about 14. Figure A8 shows that the estimated percentage effect of an additional GRP in advertising does not depend strongly on the actual number of GRP. Most of the time, confidence intervals overlap. This suggests that our baseline specification is a good approximation.

Figure A9 addresses the related question whether the estimates for the effects of advertising from the more selected sample of big advertisements presented in Figure 6 is representative. To investigate this, we have estimated the distributed lag model for the sample of big advertisements. 51 out of the 59 big advertisements are TV advertisements. Therefore, in Figure A9 we also show estimates for all TV advertisements in our baseline sample. Estimates for the smaller set of big advertisements are less precise. The patterns are very similar, especially in the first 30 minutes after the advertisement was aired.

In our next specification check, we include in addition 48 lead terms. If our specification is

Table A4: Evidence from a distributed lag model at the hourly level

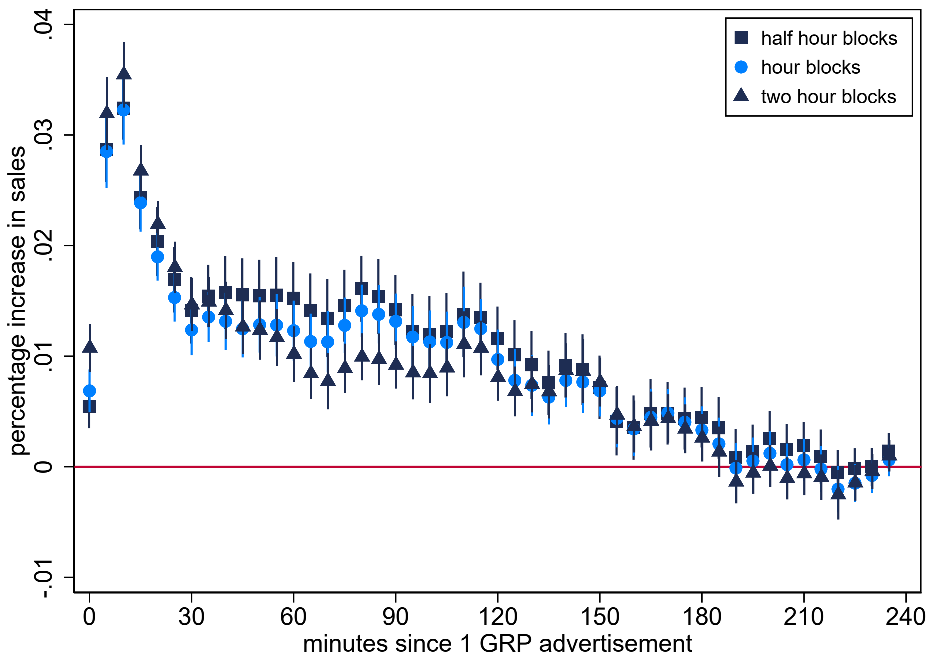
| | (1) baseline | (2) night dummy |
|---|---|---|
| GRP current hour | 0.0135*** | 0.0195*** |
| | (0.00150) | (0.00186) |
| GRP 1 hour lagged | 0.0135*** | 0.00889*** |
| | (0.00132) | (0.00156) |
| GRP 2 hours lagged | 0.00315* | 0.0124*** |
| | (0.00128) | (0.00160) |
| GRP 3 hours lagged | 0.00148 | 0.0208*** |
| | (0.00130) | (0.00161) |
| draw dummies | Yes | Yes |
| days to draw dummies | Yes | Yes |
| hour of day dummies | Yes | No |
| night dummy | No | Yes |
| Observations | 8751 | 8751 |
| $R^2$ | 0.915 | 0.846 |

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

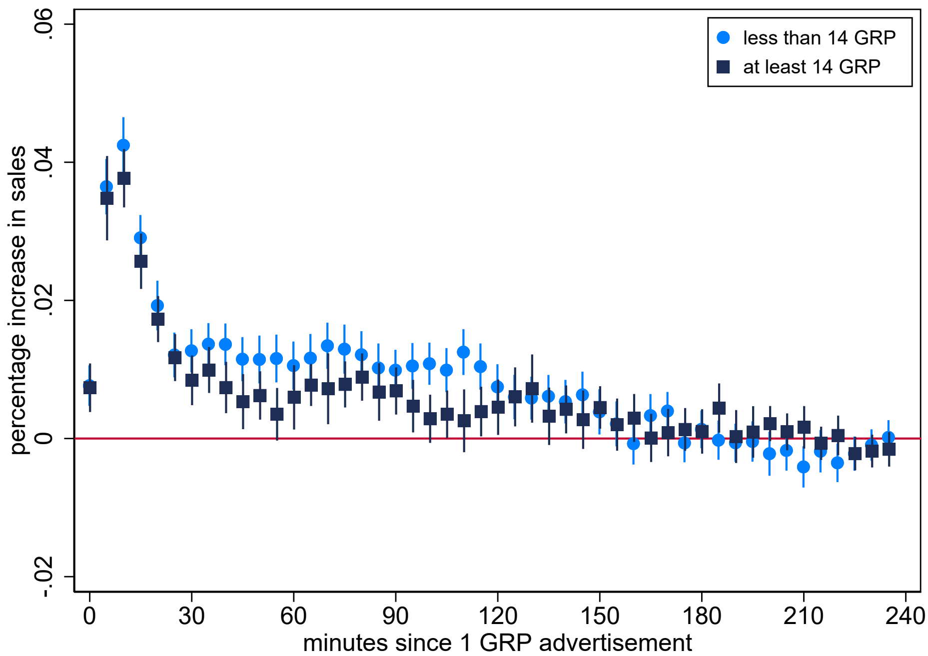
Notes: This table shows the results of regressions of the log of one plus sales on GRP's of advertising and lags thereof. Regressions were carried out at the hourly level and standard errors are clustered at the daily level.

Figure A7: Changing the length of the time blocks



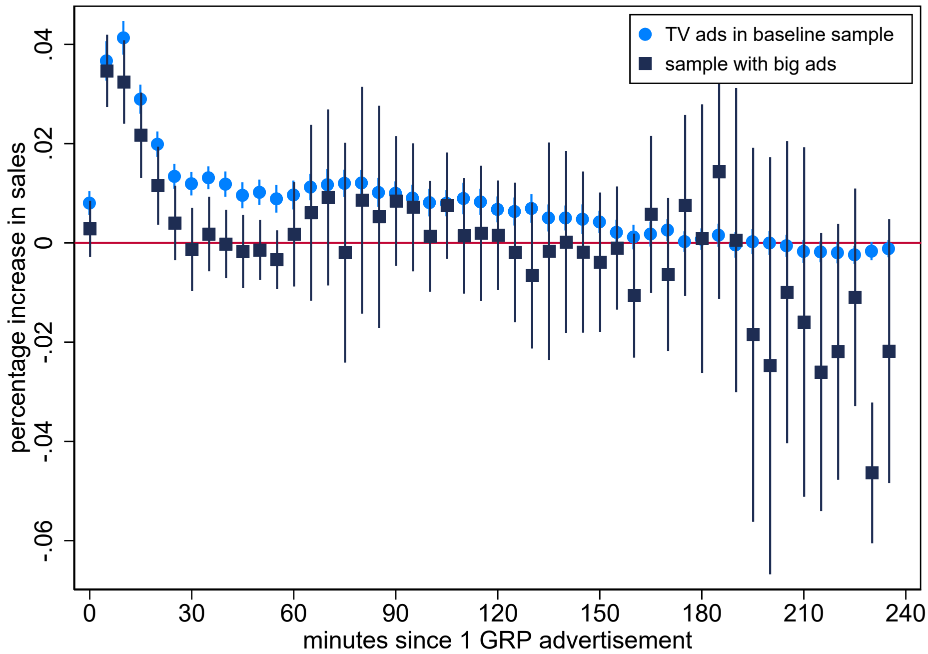

Notes: This figure shows estimates from a distributed lag model that was estimated at the minute level. The baseline specification is the same as in Figure 11 with hour blocks (hour of year fixed effects). The other two specifications use half hour and two hour blocks instead.

Figure A8: Dependence of effect on number GRP




Notes: This figure shows estimates from a distributed lag model that was estimated at the minute level. We estimate separate effects for big and small advertisements. Otherwise, the specification is the same as in Figure 11.

Figure A9: Representativeness of the advertising effects for big advertisements

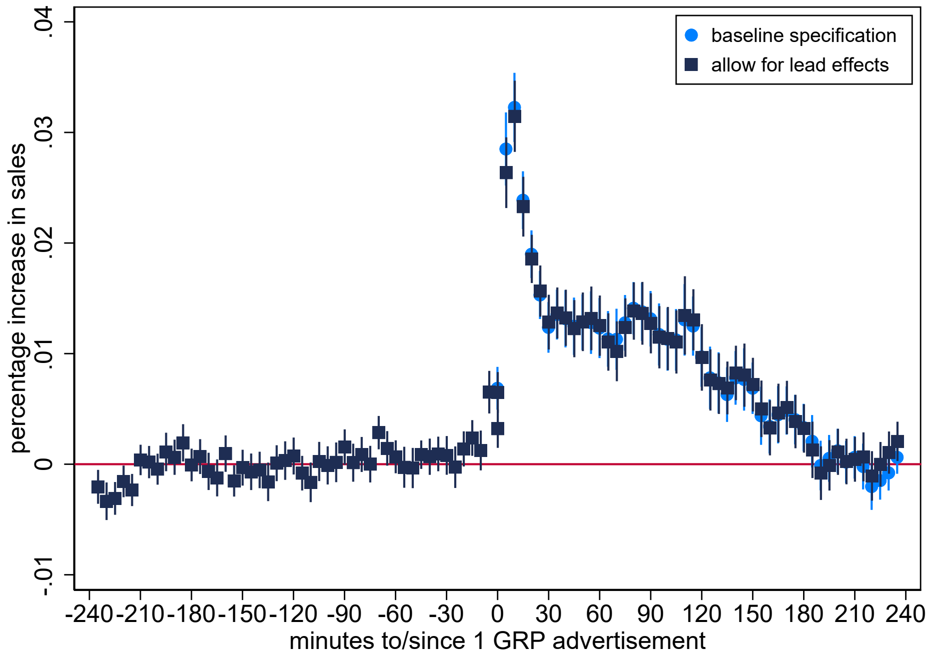



Notes: This figure shows estimates from a distributed lag model that was estimated at the minute level. We report estimates for TV advertisements in the sample with big advertisements that was used for Figure 6 and for the TV advertisements in the baseline sample. The specification is the same as in Figure 9.

appropriate, then the coefficients on these lead terms should be close to zero and the estimated effect of the lag terms should be similar to the one we obtain using our baseline specification. Figure A10 shows the result. The average implied effect of advertising before it has taken place, calculated from the lead terms using the approach described in footnote 26, is 0.00021, with a standard error of 0.00014, i.e. not statistically significant. This can be seen as another placebo test, complementing the one in Figure 12.

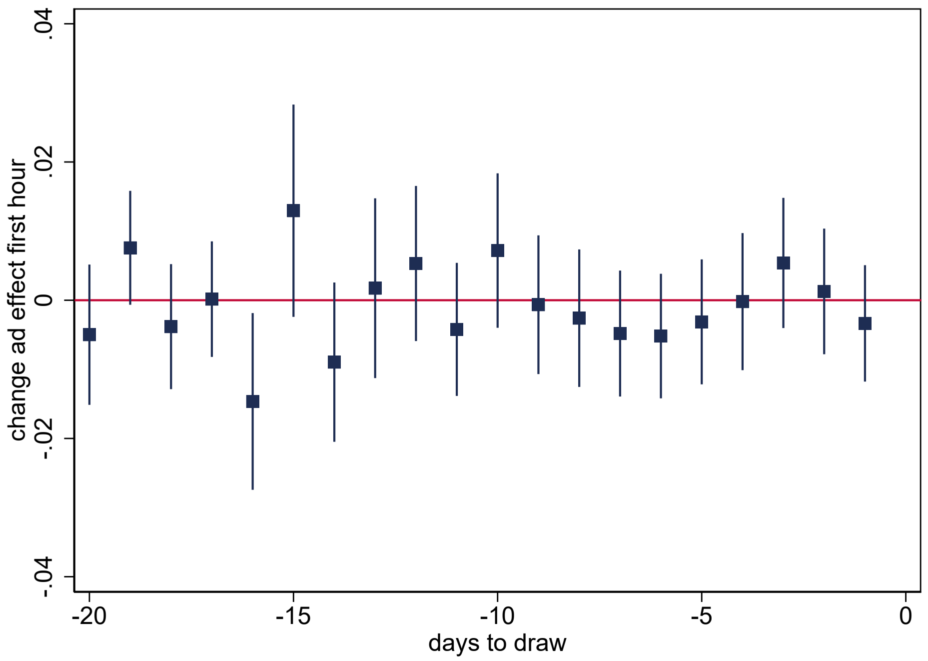
Figure 10 shows that advertising effects are stronger the less time there is until the draw. The figure is based on estimating the distributed lag model (6), where hour blocks of lag terms are interacted with days until the draw dummies. Figure A11 shows how the effect in the first hour depends on the days until the draw. Only one day has a confidence interval that does not include 0 (but is very close to it). The effect in the first hour is estimated to be slightly lower 14 days before the draw, but overall there does not seem to be a clear pattern. Therefore, the figure suggests that the percentage effect of advertising does not depend on the time until the draw.

Figure A10: Specification with leads




Notes: This figure shows estimates from a distributed lag model that was estimated at the minute level. We report our baseline estimates in Figure 7 and estimates when we include 48 additional lead terms. Otherwise, the specification is the same as in Figure 11.

Figure A11: Dependence of advertising effects on time until draw




Notes: This figure shows estimates from a distributed lag model that was estimated at the minute level. We report estimates of the interaction terms between the indicator for the first 12 lag terms for the last 20 days until the draw, $\gamma_{H(\tau),DTD(t)}$ in (6) for $H(\tau) = 1$,with $DTD(t)$ on the horizontal axis. Otherwise, the specification is the same as in Figure 11.

# C Structural model

This appendix provides additional details on our structural model and on how we estimate it. We also present additional results, including results of a number of robustness checks.

## C.1 General structure

Our structural model can be seen as an empirical version of the framework in Section 5. We briefly describe the main parts in Section 7. Here we provide details.

For each draw, there are $N$ individuals who can each buy at most one ticket. Choice is independent across draws. For each draw, time $t = 1, 2, \ldots, T$ is discrete and measured at the hourly level. $T$ is the hour of the draw and the last moment at which consumers can buy a ticket. Here, we describe the model for one draw, but in fact $T$ differs across draws.

In every hour, an individual thinks about buying a ticket with a certain probability. Advertising has a positive effect on this probability. If an individual thinks about buying a ticket, then she actively makes a decision whether to buy it. If she buys a ticket, then she receives a one-off flow of utility and cannot make any decisions anymore.[A1] Otherwise, she continues in the next period and has the option of buying a ticket there.

## C.2 Thinking about buying

In our model, advertising affects the likelihood that a consumer thinks about buying a ticket through an advertising goodwill stock. This goodwill stock increases if the individual is exposed to an advertisement. In many models of consumer choice, the advertising goodwill stock affects the utility from buying the product (e.g. Dubé et al., 2005). In our case the goodwill stock affects instead the probability to think about buying a ticket in the short-run.

Denote the goodwill stock of individual $i$ at the beginning of period $t$ by $g_{it}$. We will refer to the goodwill stock after the time at which the individual can be reached by an advertisement as the augmented goodwill stock. It is denoted by $g^a_{it}$. The augmented goodwill stock affects consumer choice by affecting the likelihood to think about buying and depreciates exponentially over time. Let $\lambda$ denote the depreciation rate and assume that the initial goodwill stock is 0. The law of motion for the (augmented) goodwill stock is

$$g^a_{it} = \begin{cases} g_{it} & \text{if } i \text{ did not see an advertisement in } t \\ g_{it} + 1 & \text{if } i \text{ saw an advertisement in } t \end{cases}$$

with initial condition

$$g_{i0} = 0$$

[A1] This is isomorphic to a model in which she pays a price today and expects to receive a flow utility in the future, provided that she cannot make any decisions in the meantime.

and

$$g_{it+1} = (1-\lambda)\cdot g_{it}^{a}.$$

The augmented advertising goodwill stock affects the probability to think about buying a ticket. We specify this probability as

$$P_{it}(\text{think about buying}) = \frac{1}{1+\exp\left(-(\gamma_0+\gamma_1 g_{it}^{a})\right)}.$$

The parameter $\gamma_0$ captures the baseline probability to think about buying a ticket in the absence of advertising. $\gamma_1$ captures the effect of advertising that runs through the augmented goodwill stock.

## C.3 Purchase decision

Once she thinks about it, a consumer decides whether or not to buy a lottery ticket. Buying a ticket yields flow utility

$$u_{it} = -p+\delta^{T-t}\psi+\sigma\varepsilon_{i1t},$$

where $p$ is the price of the ticket, $\delta$ is the hourly discount factor, $\psi$ is the draw-specific value of holding a ticket at the time of the draw (it will be estimated as a draw fixed effect), and $\varepsilon_{it}$ is a type 1 extreme value distributed taste shock (recentered, so that it is mean zero). The coefficient on the price is normalized to be minus 1, which means that flow utility is measured in terms of money. Specifying flow utility to depend on $-p+\delta^{T-t}\psi$ means that a consumer has a taste for buying the ticket as late as possible because she has to pay for it immediately but only receives a discounted benefit from this. This feature of our model is meant to capture the empirical pattern in Figure 1 that most sales occur in the last days before the draw.

If a consumer chooses not to buy, then she gets the flow utility $\sigma\varepsilon_{i0t}$, where again $\varepsilon_{i0t}$ is a type 1 extreme value distributed taste shock.

It follows that the probability of buying given that a consumer thinks about buying is given by the usual binary logit formula,

$$P_{it}(\text{buy}|\text{think about buying}) = \frac{1}{1+\exp\left(-\frac{-p+\delta^{T-t}\psi}{\sigma}\right)}.$$

The unconditional probability of buying is

$$P_{it}(\text{buy}) = P_{it}(\text{think about buying})\cdot P_{it}(\text{buy}|\text{think about buying}).$$

## C.4 Identification

We estimate the model using aggregate sales and GRP measurements at the hourly level. For this, we treat the price of a ticket, $p$, and the market size, $N$, as known. Here, we informally discuss which variation in the data identifies our parameters, respectively. We proceed in three steps.

### Parameters capturing the evolution of baseline sales

For the moment assume that the baseline probability to pay attention is known and abstract from advertising. As we observe sales and the market size, we observe at any point in time the remaining number of consumers who have not yet bought a ticket, which we can then multiply with the baseline probability to pay attention to calculate the number of consumers who think about buying at any given point in time. Our model for buying given that a consumer thinks about buying is a logit model. Therefore, the identification argument for the parameters that capture the evolution of sales, $\delta$, the 16 $\psi$'s, and $\sigma$, is the same as in the logit model. The question is whether there is only one combination of these parameters that can generate the patterns in sales over time, which at any point in time are given by the logit probabilities for these parameters multiplied by the known number of consumers who think about buying.

To assess this, we have used the model to generate predictions of sales trajectories. Here we describe the results of this. When $t$ approaches $T$, sales depend on $\psi$ and $\sigma$,but less and less on $\delta$. In the limit, when $t = T$, sales do not depend on $\delta$. Conversely, long before $T$, i.e. when $t$ is small, sales can only be small (which is what we see in the data), if $\sigma$ is sufficiently small. The reason for this is that for large values of $\sigma$, consumer choice resembles the outcome of flipping a coin, meaning that the probability of buying is approaching 0.5 and does not depend on $t$ anymore. Taken together, this suggests that sales levels long (small $t$) and short ($t$ close to $T$) before the draw identify all 16 $\psi$'s for all draws and $\sigma$.

$\delta$ mainly changes the way sales evolve in-between. The lower we set $\delta$ the steeper the curve becomes and the more convex it is. The underlying reason is that buying a ticket leads to an immediate cost and a delayed benefit. The longer the delay, the more the benefit is discounted. This can be seen as an exclusion restriction.[A2]

### Baseline level to think about buying

The baseline level of paying attention is directly linked to $\gamma_0$. There are recent advances in the literature related to identifying this baseline level, in particular Abaluck and Adams-Prassl (2021) and Heiss et al. (2021). However, the necessary variation is not in our aggregate data. Therefore, we set $\gamma_0 = -4.5$ so that the baseline probability to think about buying is 1 percent. In Section C.8.2, we re-estimate the model with a different value for $\gamma_0$. Parameter estimates

[A2] Magnac and Thesmar (2002) and Abbring and Daljord (2020) formally show that exclusion restrictions are useful for identification of the discount factor in dynamic models.

change in intuitive ways. For instance, if the baseline probability to pay attention is higher, then we can match observed sales when the value to holding a ticket on the day of the draw is smaller, which is what we find. Overall, results are robust. In particular, predicted adverting effects and counterfactual predictions remain similar.

**Parameters capturing advertising effects**

We work with aggregate data, but nevertheless wish to allow for some heterogeneity across consumers that arises over time because advertising reaches different consumers at different points in time. We do so using simulation (details are provided in Section C.5.2 below). The number of GRP's in our data are informative about how many consumers are reached at a given point in time. We use this information to simulate a number of goodwill stocks for a population of simulated consumers. For each simulated consumers we then calculate a simulated goodwill stock and the corresponding probability to think about buying. Aggregating across simulated consumers gives the predicted number of consumers who think about buying, which is then multiplied by the (now treated as known, as a result of our previous discussion) probability to buy given that a consumer thinks about buying at that point in time. The advertising effects can be matched to the patterns in the data, which are identified by the arguments given in Section 6. The decay of those advertising effects identifies the depreciation rate $\lambda$. The dependence of the magnitude of the effect on the number of GRP pins down $\gamma_1$.

## C.5 Econometric implementation

In this subsection, we provide details on the estimation procedure.

### C.5.1 Empirical setup

The data contain minute-level information on ticket sales and advertising activities for 16 draws. Estimating our structural model at the minute-level and with many covariates, mirroring the flexible reduced-form specifications we have used in Section 6, is however infeasible. Therefore, based on the results reported in Figure 11 and Table A4 we take the following approach: we aggregate the data to the hourly level, treat the time between midnight and 7am as one hour, and specify utility to only depend on the number of hours until the draw and a draw-specific parameter. This means that every day in the model has 18 hours and that we estimate the model at the hourly level.

The first period in our aggregated data is 00:00-06:59 on Jan 1 and the last period is 17:00-17:59 on Dec 31, as the draws always take place at 6pm and the last one is on December 31. Thus, the total number of periods is $\tau = 6,564$ ($\tau$ is not to be confused with $T$, which we have defined in the context of our model). We divide them up into sub-periods, one for each draw. We account for the fact that the draws differ with respect to the total number of hours in which a ticket can be bought (the draw-specific $T$ in the model) and the draw-specific value to holding a ticket ($\psi$ in the model).[A3]

In our aggregate data, we only observe that a consumer has bought a ticket, but not which ticket. Therefore, we assume that everybody buys the same ticket and that the price of that ticket is 3 euros.[A4]

### C.5.2 Simulated consumers

There is an inner and an outer loop. In the inner loop, we simulate choices of a population of simulated consumers for given values of the parameters and compute the value of a method of simulated moment (MSM) objective function. In the outer loop we estimate the parameters. The moments we use are related to sales at a given point in time given the advertising activity before that, and the evolution of cumulative sales.

We assume the market size for Dutch online lottery tickets market is 250,000 and we simulate choices of 1,000 consumers.[A5] Thus, each simulated consumer represents 250 real consumers. In the background, there is again a trade-off between computational burden and how realistic the model is. We found that simulating 1,000 consumers works well for capturing heterogeneity in the simulated advertising goodwill stock (see Section C.4) while the computational burden is still not too high. To implement this, we take aggregate sales and divide them by 250. The thought experiment that underlies our approach is that we match simulated sales to the expectation thereof, across 250,000 actual consumers, which is given by our data.

We pay particular attention to the fact that different consumers have different advertising stocks at a given point in time, as it is random whether or not they are exposed to advertisements in the periods before that. Tentatively, there will be dynamic selection in the short-run, because those consumers with higher advertising goodwill stocks will be more likely to buy, so that those with lower advertising goodwill stocks remain. Our strategy allows and controls for that. For example, think of 250,000 individuals who may in principle buy a ticket (the market size we assume). Suppose that there are 3 GRP's of advertising in a given hour and that there have not been any advertisements before that. Then, in expectation, 7,500 individuals will be reached. Now suppose that there are 4 GRP's of advertising in the hour after this. This reaches in total 10,000 individuals. Some of those individuals were among the 7,500 who have already seen an

[A3] This means that $T$ and $\psi$ need to indexed by the draw, because they differ across draws. For the ease of the exposition, in Section 7, we have described the model only for one draw. Within each draw, $t$ runs from 1 to the draw-specific $T$.

[A4] 3 euros is the price for the smallest ticket one can buy. See Section 3 for details. See also footnote 9. Assuming a different price will only re-scale the parameters, but will not change the results of counterfactual experiments. To see this, suppose we double the price and double at the same time $\Psi$ and $\sigma$. Then, importantly, $P_{it}(buy|consider)$ and $P_{it}(buy)$ will stay exactly the same. This shows that both models are observationally equivalent. Consequently, simulated sales under counterfactual advertising strategies will be the same. This means that our main assumption here is that everybody buys the same ticket.

[A5] This market size is considerably more than the maximum number of tickets that was sold in each month in our data. We experimented with different market sizes and found that results of the counterfactual simulations are not very sensitive to it. In Appendix C.8 we also present results when we assume that the market size is higher.

advertisement before and some of those will not. We assume that it is independent over time who is reached and therefore 300 individuals will see both advertisements.

For each simulated consumer we compute the implied choice probabilities by the model. These probabilities will differ across consumers, because simulated advertising goodwill stocks will differ. Denote simulated augmented advertising goodwill stocks by $\tilde{g}^a_{it}$ and simulated probabilities of buying by $\tilde{P}_{it}(\tilde{g}^a_{it})$.

We could in principle generate simulated choices by combining these probabilities with random draws $u_{it}$ from the standard uniform distribution for each consumer at each point in time. If $\tilde{P}_{it}(\tilde{g}^a_{it}) \geq u_{it}$, then the simulated choice $\hat{d}_{it}$ is one and otherwise zero. A consumer can buy at most one ticket and therefore we set $\hat{d}_{it}$ to zero after a consumer has bought for the first time. Aggregating gives simulated aggregate demand $\tilde{q}_t$, which we can then match to (rescaled, as described above) actual aggregated demand.

One challenge that arises when we pursue this approach is that the simulated choice probabilities or, in our case, simulated demand, is not a smooth function in the parameters. This is due to the fact that in discrete choice models individuals are assumed to either choose to buy or not at a given point in time. Consequently, for each simulated consumer, small changes in parameters will either have no effect on her decision (which stays at 0 or 1), or change it discretely. Such non-smoothness can lead to problems with the usual methods for finding an optimum of the objective function because of flat spots.

We overcome this problem by using a smoothed accept-reject simulator to make the demand function fully smooth in the parameters. We use this very conservatively, however, and only to avoid that the estimator gets stuck on a flat spot.

Following McFadden (1989), the simulator that we choose has the logit form. Instead of generating choices for individual $i$ in $t$ that are either 0 or 1, we generate smoothed choices

$$\hat{d}^{smooth}_{it} = \frac{1}{1+\exp\left(\frac{u_{it}-\tilde{P}_{it}(\tilde{g}^a_{it})}{s}\right)},$$

where $s$ is the smoothing parameter. The higher $s$ the more smoothing there is. When $s$ becomes very small, then there is no smoothing anymore. In our case, it is sufficient to use very little smoothing. We specify $s = 0.00015$.[A6] In the model, consumers can buy at most once. Here, we account for this by first calculating the cumulative smoothed sales up to the previous period, $1-\sum_{s=1}^{t-1}\hat{d}^{smooth}_{is}$. Then, we replace $\hat{d}^{smooth}_{it}$ by those cumulative sales if it exceeds it.

### C.5.3 Method of simulated moments

The set of estimated structural parameters that do not change across draws is $\{\lambda, \gamma_1, \delta, \sigma\}$. In addition, we estimate 16 values $\psi_1, \ldots, \psi_{16}$ of holding a ticket at the time of the respective draw.

[A6]We have experimented with different values of $s$ and the result is not sensitive to the choice of $s$ for values of $s$ around 0.00015.

Thus the full set of structural parameters to be estimated is $\theta \equiv \{\lambda, \gamma_1, \delta, \sigma, \psi_1, \ldots, \psi_{16}\}$.

Recall that we only have access to aggregate data. Let $\hat{u}_t(\theta) \equiv q_t - \tilde{q}_t(\theta)$ be the difference between actual aggregate demand $q_t$ in the data, divided by 250, and the model prediction $\tilde{q}_t(\theta)$. Starting from this we specify a set of moments $\mathbb{E}[m(z_t, \hat{u}_t(\theta))] = 0$, where $z_t$ is a vector of exogenous variables constructed from the data so that the left hand side is a column vector and the right hand side is a vector of zeros and the expectation is taken over hours.

Let $\bar{m}(\tilde{\theta})$ be the average of $m(z_t, \hat{u}_t(\theta))$, over time in hours across all draws (thus over $\tau$ time periods), evaluated at any candidate parameter vector $\tilde{\theta}$. The MSM estimator is

$$\hat{\theta} = \arg\min_{\tilde{\theta}} \bar{m}(\tilde{\theta})' W \bar{m}(\tilde{\theta}),$$

where $W$ is a positive definite weighting matrix. Section C.5.4 below provides details on the specification of the moments and the weighting matrix.

In Section C.4, we have provided an informal discussion of identification. The (technical) condition for identification is that the moment conditions hold at the true parameters $\theta$ (see for instance Newey and McFadden, 1994). Then, $\hat{\theta}$ is consistent. An estimator of the variance-covariance matrix is given by (Newey and McFadden, 1994)

$$\widehat{var(\hat{\theta})} = \frac{1}{\tau}(A'WA)^{-1} B (A'WA)^{-1},$$

where

$$A = \frac{\partial \bar{m}(\hat{\theta})}{\partial \hat{\theta}'}$$

and

$$B = A'W(m(\hat{\theta}) - \bar{m}(\hat{\theta}))(m(\hat{\theta}) - \bar{m}(\hat{\theta}))'WA.$$

### C.5.4 Moments and weighting matrix

$z_t$ in the moments $\mathbb{E}[m(z_t, \hat{u}_t(\theta))] = 0$ contains 3 sets of exogenous variables: a full set of dummy variables for the number of days until the draw, the number of GRP's in $t$, $t-1$, $t-2$, and $t-3$, and variables calculating cumulative sales up to point $t$. This means that we attempt to pick the parameters so that the model captures well the evolution of sales over time and the reaction to advertisements.

To implement this we stack all $\hat{u}_t(\theta)$ into a vector $\hat{u}(\theta)$ of dimension $\tau \times 1$ and define a $\tau \times M$ matrix of exogenous variables $Z = (\mathbf{1}, Z_1, Z_2, Z_3)$, where $\mathbf{1}$ is a vector of ones, $Z_1$ contains times until draw dummies in the columns, $Z_2$ contains GRP's and lags thereof in the columns, and $Z_3$ is a matrix with indicators such that it takes cumulative sales at the daily level, separately for each draw. $Z_3$ is block-diagonal with sub-matrices $Z_{3,r}$ on the diagonal ($r$ indexing draws). Each column of these sub-matrices is for one day and contain a set of ones on top and zeros in the bottom, such that the cumulative prediction error is calculated on a daily level when we multiply $Z_3'$ with $\hat{u}_t(\theta)$.

After eliminating linearly dependent columns, $Z$ has $M = 376$ columns, meaning that we have 376 exogenous variables.[A7] Using this, we calculate

$$\bar{m}(\tilde{\theta}) = \frac{1}{\tau} Z' \hat{u}(\tilde{\theta}).$$

Regarding the weighting matrix, we follow the usual two-step GMM procedure. We first choose the weighting matrix $W$ to be

$$W = ((\frac{1}{\tau}\sum_t m(z_t, \hat{u}_t(\theta_s))m(z_t, \hat{u}_t(\theta_s))/\tau)^{-1}$$

to get consistent estimates $\hat{\theta}_s$ where $\theta_s$ is a vector of starting values that we chose. Then we calculate the optimal weight matrix at $\hat{\theta}_s$: $W^o(\hat{\theta}_s) = (\frac{1}{\tau}\sum_t m(z_t, \hat{u}_t(\hat{\theta}_s))m(z_t, \hat{u}_t(\hat{\theta}_s))')^{-1}$ and then get the efficient estimates.

## C.6 Parameter estimates and fit

In this subsection, we present our estimation results and assess the fit of the model. Table A5 shows the estimated parameters. The effect of advertising on sales depreciates quickly, at an hourly rate of about 41.3 percent. Recall that we set $\gamma_0$, the intercept of the probability to think about buying to -4.5. This means that about $1/(1+exp(-(-4.5))) \approx 1$ percent of the consumers will think about buying a ticket in the absence of advertising. A one unit increase in the goodwill stock from zero to one, driven by seeing an advertisement, will increase the probability of to think about buying to $1/(1+exp(-(\gamma_0+\hat{\gamma}_1))) \approx 0.0358$. One hour later, the goodwill stock is $1-0.413 = 0.587$ and the probability to think about buying is $1/(1+exp(-(\gamma_0+\hat{\gamma}_1 \cdot 0.587))) \approx 0.0221$ if no advertisement reaches the consumer. Yet another hour later it is 0.3446 and the probability to think about buying is 0.0166. If the consumer is instead reached by another advertisement in the hour after she was first reached, then the augmented goodwill stock becomes 1.587 in the second period and the probability to think about buying is 0.0701. And when she is reached again one hour later, it is 1.9316 and the probability to think about buying is 0.1026.

The hourly discount factor is estimated to be 0.994. One month before a draw, the value consumers attach to a ticket is only $0.994^{18 \cdot 31} = 3.48\%$ of the value on the day of the draw. For that reason, consumers will value buying tickets late and being reminded at later points in time. This suggests that in the specific context we study here and for the specific product, consumers discount future utility at a very high discount rate.[A8] One of the findings in the literature on

[A7] $Z_1$ originally has 30 columns. $Z_2$ contains GRP's and 3 lags thereof, so it has 4 columns. $Z_3$ has 365 columns. Most columns in $Z_1$ are linear combinations of columns in $Z_3$. After dropping those, $Z_1$ has 7 columns left. Thus, we have in total $1+7+4+364 = 376$ columns.

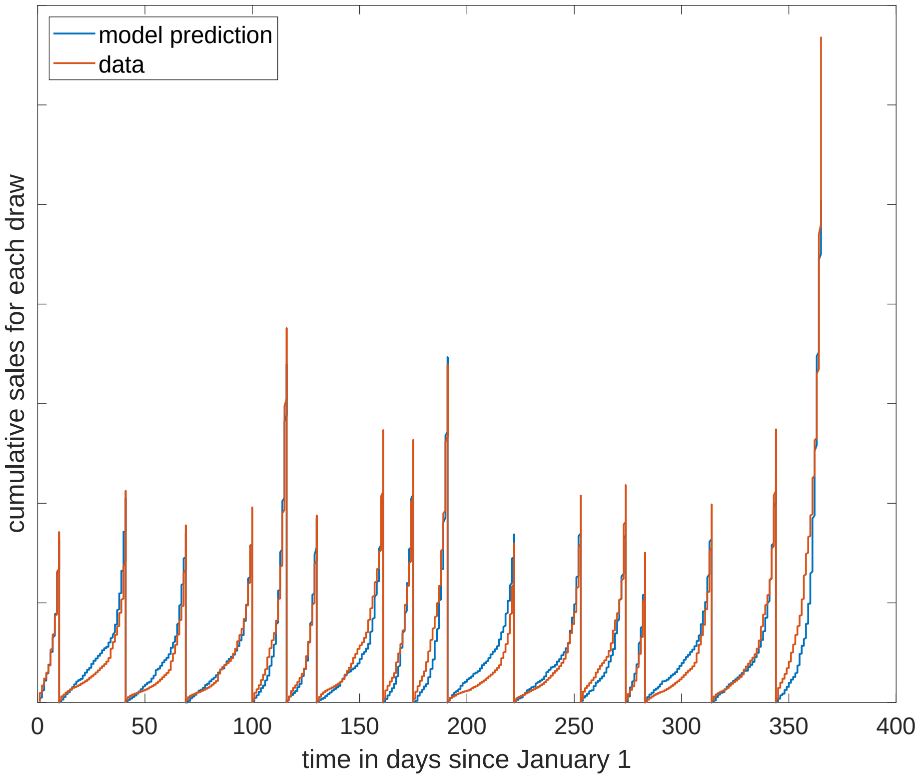
[A8] The implied daily discount factor is $0.994^{18} = 0.897$ and the implied weekly discount factor is $0.994^{18 \cdot 7} =$

Table A5: Parameter estimates

| parameter | estimate | ste. |
|---|---|---|
| depreciation rate goodwill stock ($\lambda$) | 0.413 | 0.053 |
| effect of goodwill stock on probability to think about buying ($\gamma_1$) | 1.207 | 0.083 |
| hourly discount factor ($\delta$) | 0.994 | 0.000 |
| multiplying factor taste shock ($\sigma$) | 0.480 | 0.020 |
| value to having a ticket on the day of the draw ($\psi$) | | |
| 10 January, 2014 | 2.355 | 0.072 |
| 10 February, 2014 | 2.579 | 0.064 |
| 10 March, 2014 | 2.259 | 0.084 |
| 10 April, 2014 | 2.231 | 0.067 |
| 26 April, 2014 (King's Day) | 3.236 | 0.079 |
| 10 May, 2014 | 2.518 | 0.058 |
| 10 June, 2014 | 2.702 | 0.070 |
| 24 June, 2014 (Orange draw) | 2.852 | 0.076 |
| 10 July, 2014 | 2.995 | 0.082 |
| 10 August, 2014 | 2.108 | 0.089 |
| 10 September, 2014 | 2.557 | 0.065 |
| 1 October, 2014 (special 1 October draw) | 2.635 | 0.065 |
| 10 October, 2014 | 2.409 | 0.066 |
| 10 November, 2014 | 2.253 | 0.074 |
| 10 December, 2014 | 2.656 | 0.077 |
| 31 December, 2014 (New year's eve draw) | 3.968 | 0.132 |

Notes: Structural estimates. Obtained using the method of simulated moments. See Section C.5 for details on the estimation procedure. The probability to think about buying is specified as $P_{it}(\text{think about buying}) = 1/(1+\exp(-(\gamma_0+\gamma_1 g^a_{it}))$. We set $\gamma_0 = -4.5$.

Figure A12: Model fit




Notes: This figure shows actual cumulative sales and cumulative sales predicted by our structural model using the estimated parameters reported in Table A5.

time discounting is that discount rates are domain dependent, meaning that discount rates are highly specific to the context and the type of decision individuals make. See Loewenstein et al. (2003) for a review. Arguably, our estimate of the discount factor is very low because we study the decision to buy a relatively cheap product. In contrast, De Groote and Verboven (2019) estimate a model of solar panel adoption, where households make a big investment decision. Their estimate of the monthly discount factor is 0.9884. In line with the idea that this is a much more important decision with higher stakes, their estimate is much higher than ours.

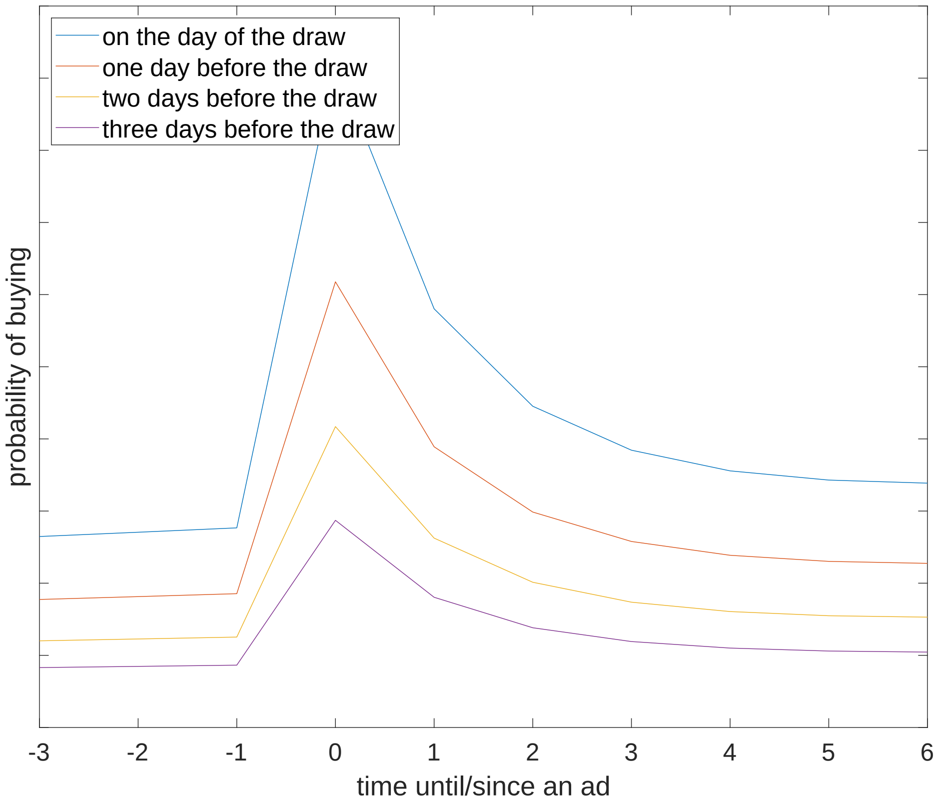
The estimated standard deviation of the taste shock is $0.480 \cdot \sqrt{\pi^2/6} \approx 0.616$ euros ($\sqrt{\pi^2/6}$ is the standard deviation of a type 1 extreme value random variable). Finally, the 16 estimates of the draw fixed effects are in line with expectations and positively related to the size of the jackpot, mirroring the pattern in Figure 1.

Figure A12 shows the model fit. Arguably, with only a few parameters, the model fits the overall patterns in the data relatively well.

In the following, we present the results of our structural analysis. We first show our estimates of implied short-term advertising effects and their dependence on the days until the draw. We also compare the estimates of short-term advertising effects from the structural model to those

0.468. The corresponding discount rates are given by $1/0.994 - 1 = 0.6\%$ at the hourly level, $1/0.994^{18} - 1 = 11.4\%$ at the daily level, and $1/0.994^{18 \cdot 7} - 1 = 113.5\%$ at the weekly level.

Figure A13: Dependence of predicted effect of advertising on timing




Notes: This figure shows how the predicted absolute increase in sales that is due to seeing an advertisement depends on the time until the draw. Obtained from our structural model using the estimated parameters reported in Table A5.

from the model-free evidence. Then we use the model estimates to evaluate counterfactual dynamic advertising strategies.

A key quantity the model predicts is the immediate effect of advertising on sales and how this effect depends on the time until the draw. Figure A13 shows, for our structural estimates, how the probability of buying a ticket, if the individual has not done so yet, changes when she is exposed to an advertisement. There are four lines for four different time periods, on the day of the the draw, and 1 day, 2 days and 3 days before that. As one can see, the closer the time of the advertisement is to the deadline, the more effective is the advertisement—in line with the model-free evidence that we have presented in Section 6. The figure shows that our model can generate this effect.

Next, we compare the effect of advertising on sales implied by the structural parameter estimates with that implied by the model-free evidence in Section 6. Table A6 shows the results. The effect of advertising in the first hour is 1.92% for the structural model and 1.7% for the reduced-form analysis. The next row shows the numbers for the total effect. They are also similar. The next four rows of the table compare the implied evolution of sales. For example, the model predicts that 84.92% of the consumers who are in the market to buy a ticket have not

Table A6: Comparison of results from structural model with reduced-form results

| | structural model | reduced-form |
|---|---|---|
| average effect of advertising in the 1st hour | 0.0192 | 0.0171 |
| effect of advertising in the first 4 hours | 0.0344 | 0.0361 |
| % consumers who have not bought yet on the day of draw | 0.8492 | 0.8480 |
| % consumers who have not bought yet 1 day before draw | 0.8767 | 0.8830 |
| % consumers who have not bought yet 2 days before draw | 0.8977 | 0.9030 |
| % consumers who have not bought yet 3 days before draw | 0.9116 | 0.9190 |

Notes: This table compare the effect of advertising and the percentage consumers who have not bought yet that are implied by the structural estimates to the same quantities implied by the reduced-form evidence. We have obtained the numbers for the structural model by performing counterfactual simulations. Footnote 26 describes how average advertising effects have been calculated for our reduced-form evidence. The percentage consumers who have not bought yet was estimated by regressing that fraction on days-until-draw indicators, pooling across draws.

done so on the day of the draw (averaging across hours on that day). This compares to 84.80% in the data. Overall, the results from the structural model and the model-free evidence are very similar. This is remarkable, because we estimate advertising effects in our reduced form analysis at the minute level using a very flexible specification with hour of year fixed effects, while we conduct our structural analysis at the hourly level and cannot control for time effects at such a level. Also, the bottom part of the table suggests that our parsimonious specification of the probability of buying given that the consumer thinks about buying in the structural model is able to capture the patterns in the data well.

## C.7 The elasticity of sales with respect to advertising

To get an idea about the size of the implied advertising effects, we have increased all GRP's by 10% and have simulated the effect of this on sales. From this we calculate an elasticity of sales with respect to advertising of 0.130.

To put this into perspective, the average elasticity Hu et al. (2007) report for their meta study of offline effects of classical TV advertising is 0.113. Shapiro et al. (2021) estimate elasticities for 288 consumer packaged goods (CPG) brands and find an average elasticity of 0.014, with about 65 to 75 percent of the elasticity estimates not being statistically different from zero.

Our estimate of 0.130 is similar to the average Hu et al. (2007) report, but higher than the average elasticity in Shapiro et al. (2021). This could be the case for various reasons. First, the product is different. Shapiro et al. (2021) look at classical CPG's and not lottery tickets. Second, in general advertising serves many purposes. In our case advertising is designed to remind consumers to buy a ticket. It could be that reminder advertising is particularly effective. Third, in our case consumers have the possibility to immediately buy the product online, whereas they might forget doing so when they need to first visit the store at some future point in time, as it is the case for the classical CPG's. Related to this, we use our minute-level advertising and online sales data to estimate the short-term responsiveness of sales, as opposed as the effect on sales at a lower frequency, such as days or weeks. Finally, the measure of advertising in our case is GRP's instead of advertising expenditures.

## C.8 Robustness

In this appendix, we assess how robust our parameter estimates are to assuming a different market size (Appendix C.8.1), to assuming a different baseline probability of paying attention (Appendix C.8.2), and to allowing for serial correlation in viewership behavior (Appendix C.8.3). We also assess how these alternative specifications affect our results of counterfactual experiments (Appendix C.8.4).

### C.8.1 Assumption on market size

We first assess the robustness to making alternative assumptions about the market size. For this, we re-estimate the model assuming a market size of 300,000. We continue to impose that $\gamma_0 = -4.5$ and, focusing on baseline sales for the moment, expect that the higher market size will lead to estimates that imply a lower probability to buy given that a consumer thinks about buying. As discussed in Section C.4, the evolution of sales over time is closely related to the discount factor, which we do not expect to change much for that reason. But when the variance of the taste shocks, that is directly related to $\sigma$, decreases, then the model will produce lower sales given that a consumer thinks about buying (as the average value to holding a ticket is about 2.64 euros and the price is 3 euros, so in the model consumers always buy due to taste shocks; in fact more so the less time there is until the draw).

parameter estimates. We find that indeed, $\sigma$ is estimated to be lower. This produces choice probabilities given that a consumer thinks about buying that are roughly 73% as big as for our baseline model with a market size that is 20% bigger than the baseline case one day before the draw. The estimate of the depreciation rate of the advertising goodwill stock increases slightly and the discount factor is almost unchanged. $\gamma_1$ is slightly different. In section C.8.4, we show that the implied results from different counterfactual advertising strategies are very similar across different specifications.

### C.8.2 Assumption on probability to think about buying

Next, we assess the robustness to making alternative assumptions about the baseline probability to think about buying in the absence of advertising. For this, we re-estimate the model imposing that $\gamma_0 = -2.944$ so that the probability to think about buying without advertising increases from 1% (baseline specification) to 5%. Focusing on baseline sales for the moment, we expect that the higher baseline probability to think about buying will lead to estimates that imply a lower probability to buy given that a consumer thinks about buying.

Table A7: Robustness checks

| parameter | (1) baseline specification | ste. | (2) 120 percent market size | ste. | (3) increased $\gamma_0$ | ste. | (4) generalized model | ste. |
|---|---|---|---|---|---|---|---|---|
| depreciation rate goodwill stock ($\lambda$) | 0.413 | 0.053 | 0.592 | 0.068 | 0.664 | 0.123 | 0.710 | 0.067 |
| effect of goodwill stock on probability to think about buying ($\gamma_1$) | 1.207 | 0.083 | 1.365 | 0.099 | 1.171 | 0.099 | 1.464 | 0.099 |
| hourly discount factor ($\delta$) | 0.994 | 0.000 | 0.996 | 0.001 | 0.995 | 0.000 | 0.994 | 0.000 |
| multiplying factor taste shock ($\sigma$) | 0.480 | 0.020 | 0.370 | 0.036 | 0.369 | 0.015 | 0.468 | 0.021 |
| value to having a ticket on the day of the draw($\psi$) | | | | | | | | |
| 10 January, 2014 | 2.355 | 0.072 | 2.435 | 0.066 | 1.954 | 0.056 | 2.382 | 0.074 |
| 10 February, 2014 | 2.579 | 0.064 | 2.520 | 0.063 | 2.059 | 0.056 | 2.553 | 0.067 |
| 10 March, 2014 | 2.259 | 0.084 | 2.388 | 0.081 | 1.941 | 0.060 | 2.391 | 0.081 |
| 10 April, 2014 | 2.231 | 0.067 | 2.350 | 0.074 | 1.898 | 0.064 | 2.322 | 0.060 |
| 26 April, 2014 (King's Day) | 3.236 | 0.079 | 3.051 | 0.059 | 2.488 | 0.038 | 3.226 | 0.073 |
| 10 May, 2014 | 2.518 | 0.058 | 2.530 | 0.057 | 2.072 | 0.049 | 2.533 | 0.065 |
| 10 June, 2014 | 2.702 | 0.070 | 2.655 | 0.057 | 2.124 | 0.058 | 2.734 | 0.061 |
| 24 June, 2014 (Orange draw) | 2.852 | 0.076 | 2.759 | 0.048 | 2.303 | 0.034 | 2.820 | 0.066 |
| 10 July, 2014 | 2.995 | 0.082 | 2.873 | 0.054 | 2.378 | 0.049 | 3.033 | 0.075 |
| 10 August, 2014 | 2.108 | 0.089 | 2.246 | 0.091 | 1.802 | 0.075 | 2.164 | 0.086 |
| 10 September, 2014 | 2.557 | 0.065 | 2.582 | 0.051 | 2.099 | 0.050 | 2.542 | 0.060 |
| 1 October, 2014 (special 1 October draw) | 2.635 | 0.065 | 2.637 | 0.062 | 2.121 | 0.052 | 2.682 | 0.066 |
| 10 October, 2014 | 2.409 | 0.066 | 2.447 | 0.070 | 1.993 | 0.057 | 2.451 | 0.064 |
| 10 November, 2014 | 2.253 | 0.074 | 2.369 | 0.077 | 1.954 | 0.065 | 2.413 | 0.075 |
| 10 December, 2014 | 2.656 | 0.077 | 2.622 | 0.061 | 2.221 | 0.048 | 2.768 | 0.081 |
| 31 December, 2014 (New year's eve draw) | 3.968 | 0.132 | 3.531 | 0.129 | 2.804 | 0.043 | 3.915 | 0.143 |

Notes: Structural estimates. See Appendix C.5 for details on the estimation procedure. Estimates for the baseline specification in Table A5 are repeated in column (1). The second set of parameter estimates was obtained under the assumption that the market size is 300,000 instead of 250,000. The fourth set of estimates is for the generalized model with serially correlated viewership behavior described in Appendix C.8.3. The probability to think about buying is specified as $P_{it}(\text{think about buying}) = 1/(1+\exp(-(\gamma_0+\gamma_1 g_{it}^a)))$. For specifications (1), (2), and (4), we set $\gamma_0 = -4.5$. The fourth set of estimates was obtained by setting $\gamma_0 = -2.944$ so that the baseline probability of paying attention is 5%.

Table A7 shows the resulting parameter estimates. Similarly to the specification of 120% market share, the model produces a lower estimate of $\sigma$, which gives rise to lower sales given that a consumer thinks about buying and thereby offsets the higher probability to think about buying. The estimated depreciation rate of the goodwill stock is higher. The estimate of the discount factor and of $\gamma_1$ are almost unchanged.

#### C.8.3 A model with serially correlated viewership

So far, we have assumed that the probability that a consumer $i$ is reached in $t$ by an advertisement is given by the number of GRP's. Implicitly, this assumes that reaching a consumer in $t$ is independent of reaching the same consumer in another period $t'$, for instance $t-1$. This can only be the case if viewership behavior is not serially correlated.

While this is likely violated at the minute level, it may be a reasonable approximation at the hourly level at which we estimate our model. We have no data to directly quantify how likely it is that the same consumer is reached when there are advertisements in two consecutive hours. Therefore, we assess whether this assumption substantially affects our estimates and the main conclusions we draw from them by extending our model to allow for serial correlation in viewership behavior and by re-estimating it for a given degree of serial correlation.

In this extended version of the model, there are two states for each consumer: watching TV or listening to the radio and not watching TV or listening to the radio. When estimating the model we proceed in two steps. We first simulate, for each individual, whether they are watching TV or are listening to the radio in a given time period. Then we impose that advertising can only reach those consumers who are actually watching TV or are listening to the radio.

Formally, let state $k=1$ be the state of not watching or listening and state $k=2$ the one of watching TV or is listening to the radio. Specify a 2-by-2 Markov transition matrix

$$\Pi = \begin{bmatrix} 0.8 & 0.2 \\ 0.4 & 0.6 \end{bmatrix}.$$

This means that if an individual is watching TV or is listening to the radio at time $t$, then there will be 40% chance that she will stop watching and 60% chance that she will continue watching in period $t+1$. From this we compute the implied 2-by-1 vector $P^{\infty}$ of stationary probabilities. Then, we use $P^{\infty}$ to simulate individual viewership demand in the first period and $\Pi$ to simulate paths in subsequent periods.

Note that here, we treat the transition probabilities as known. We could estimate them if we had data at the consumer level.

Table A7 shows the estimation result for this more general model. Now, the same pattern in the data is rationalized by a model in which viewership demand is serially correlated. The results show that this can be achieved by a higher depreciation rate of the goodwill stock. The remaining parameters are almost not affected.

#### C.8.4 Counterfactual experiments with different specifications

In this sub-section, we assess how these alternative specifications will affect our results of counterfactual experiments. For this, we re-produce the result of counterfactual experiments in Section 7 using three different specifications. The results thus correspond to the ones in the first column of Table 2.

Table A8 shows the result. The result for the baseline specification is repeated in the first column. Column (2) displays the result using the set of estimates that we obtain when we specify the market size to be 120 percent of what we used before. Results for a baseline probability to think about buying of 5% are reported in column (3). Finally, column (4) shows the results using the estimates of the generalized model. Overall, we find similar magnitudes across different specifications.

Table A8: Robustness checks: counterfactual experiments with alternative specifications

| strategy | (1) baseline specification | (2) 120 percent market size | (3) increased probability of paying attention | (4) generalized model |
|---|---|---|---|---|
| data (reference point) | 100% | 100% | 100% | 100% |
| no advertising at all | 84.3% | 88.5% | 93.5% | 86.5% |
| spreading advertisements equally in the last 4 days before the draw | 115.5% | 112.3% | 104.2% | 112.3% |
| shift advertising from third week to fourth week | 110.8% | 109.2% | 103.0% | 108.8% |
| shift advertising from fourth week to third week | 87.7% | 88.5% | 94.4% | 89.3% |

Notes: This table shows the results of the counterfactual experiments in Section 7 using three alternative specifications.