Forthcoming
Management Science

Behavioral-Based Advertising*

Qiaowei Shen
(Peking University)

J. Miguel Villas-Boas
(University of California, Berkeley)

November, 2016

* E-mail addresses: qshen@gsm.pku.edu.cn, and villas@haas.berkeley.edu.

Behavioral-Based Advertising

Abstract

This paper considers the effect of firms sending advertising messages to consumers based on their past purchase behavior. If past purchase behavior on a product category is positively correlated with a consumer having high preferences in another category, a firm may want to advertise more intensively to those consumers that purchased the former category, if possible. The paper finds that this can lead to lower prices in the initial category if the annoyance of receiving advertising is large, and to higher prices if the annoyance of receiving advertising is not too large, as consumers expect a possible additional surplus from the category affected by behavioral-based advertising. If receiving advertising does not yield too much annoyance to consumers, firms end up better off due to both higher prices, and the increased demand of better matching of advertising. Behavioral-based advertising may also lead firms to sell less in the initial category than without behavioral-based advertising, as a way to be able to better target the most valuable consumers. If the consumer annoyance of receiving advertising is large, firms may end up serving only a few consumers initially, as attracting more consumers requires prices that are too low, and the initial consumers are attracted because of the possibility of lower prices in the following period. The paper also investigates the effects of joint behavioral-based pricing and advertising, and of different firms benefiting from the purchase information.

1. Introduction

One major result of firms being able to keep track of individual consumer purchases is that firms can now, in some cases, advertise differently to consumers with different purchase histories. This practice of behavioral-based advertising is currently widespread with the development of the Internet, and one can argue that it is one of the most (if not the most) visible practices by firms as a result of customer recognition. For example, a consumer who purchases a book from Amazon.com gets personalized recommendations of books with a similar taste. A customer who has shopped at a fashion web site receives customized emails about the latest style and products of her preference based on her past purchase record. Hotel chains inform their customers in the database about the new openings and services. From advertising on book titles, to automobile models, to travel packages, consumers receive targeted advertising that is the result of their past purchases. These effects can also be seen in the context of targeted sales efforts to consumers who have already bought an initial product from a firm. For example, insurance agents commonly suggest additional insurance programs to their existing clients based on their individual needs and current insurance policies. An important role of financial advisors is to inform their customers of new financial products that are most appropriate for their needs and secure a sale.

This paper explores the effects of behavioral-based advertising when preferences for one product may be correlated with preferences for another product. For example, preferences for a particular historical novel may be correlated with preferences for other historical novels, preferences for a model of a luxury automobile manufacturer may be correlated with other models of that manufacturer a few years later, or preferences for one financial product can be correlated with preferences for other financial products or services.

This ability to do behavioral-based advertising may lead consumers in the initial periods to realize that their purchases may lead to possibilities to earn surplus in future periods. Targeted advertising to the consumers who bought the initial products makes these consumers more likely to be aware of other products in future periods that they may potentially value. This then is a force for consumers to be willing to pay higher prices in the initial periods, and to be less price sensitive (if advertising annoyance is not too large). Behavioral-based advertising can then be helpful for firms both because of better matching of advertising in the later periods and because of higher prices in the initial periods. Strategic firms may also want to sell less in the initial periods to better identify the consumers that are more likely to value most the product, and charge them more in the future periods. In order to fully understand these effects, one also has to consider the dis-utility created by receiving advertising. These dis-utility effects of receiving advertising can lead to lower prices, which can ultimately benefit consumers if the drop in prices is sufficiently large.

If the dis-utility of receiving advertising is not too large, the firm sells to a sufficiently large set of consumers in the initial periods, such that in future periods concentrates only on pricing for those consumers. On the other hand, if the dis-utility of receiving advertising is sufficiently large, the firm would have to lower the price too much in the initial periods to attract consumers. The firm then prefers to supply only a limited number of consumers in the initial periods, who realize that they will be offered a lower price in later periods, and that will compensate them for the dis-utility of receiving extra advertising.

The existing literature on behavior-based market offerings has concentrated mostly on price discrimination effects (e.g., Villas-Boas 1999, 2004, Fudenberg and Tirole 2000).[1] There is also a literature on the personalization of product characteristics (e.g., Fudenberg and Tirole 1998, Zhang 2011). Neither of these literatures has investigated the possibility of doing consumer advertising contingent on the purchase histories of consumers. Behavioral-based advertising has specific effects that are not present in these literatures on behavioral-based pricing or personalization of product characteristics. First, behavioral-based advertising can be beneficial to firms because of better matching between who receives advertising, and who is likely to buy the product, which does not have an equivalent effect in those literatures. In fact, in a monopoly setting, both the firm and consumers can benefit from behavioral-based advertising, while this is not the case for behavioral-based pricing. Second, with behavioral-based advertising demand may be reduced in the initial periods because of strategic behavior of the firms but not consumers, while in behavioral-based pricing demand may be reduced in the initial periods because of strategic behavior by either consumers or firms. Third, a potential benefit for firms of behavioral-based pricing is price discrimination in the later periods, which does not have an equivalent with just behavioral-based advertising. In Section 5 below we directly further compare behavioral-based advertising with behavioral-based pricing, and investigate what happens when there is both behavioral-based advertising and behavioral-based pricing.

There is a related literature on static targeting of advertising based on consumer preferences (e.g., Stegeman 1991, Roy 2000, Iyer et al. 2005, Bergemann and Bonatti 2011). In relation to that literature this paper can be seen as considering how the firms learn about consumer preferences by tracking consumer purchase histories, and investigating what happens when the firms decide initially on price to determine the number of consumers that the firms are going to identify as having a preference for their product. Behavioral segmentation is becoming increasingly important with the increased ability of firms to collect information on the behavior of consumers, and it is important to investigate the dynamic effects of the strategic behavior of firms and consumers

[1] See also Chen (1997), Taylor (2003), Acquisti and Varian (2005), Pazgal and Soberman (2008), Chen and Zhang (2009), Chen and Pearcy (2010), Esteves (2010), Shin and Sudhir (2010) among others.

about how the behavioral segmentation ends up occurring. One important dimension of potential customers' behavior that reveal some information on preferences is the purchase history, which is the focus of this paper.[2] Esteves (2009) considers behavior-based price discrimination when initially the set of consumers who can buy the product is determined by advertising, but in that paper a consumer receiving advertising is not contingent on the purchase history of that customer.

There is also an empirical literature investigating the effectiveness of advertising based on the prior behavior of the customers. Malthouse and Elsner (2006) show that identifying the segments of consumers by their past purchases allows for more effective online advertising. Goldfarb and Tucker (2011) show that privacy regulations that limit the ability of firms to tailor advertising to a consumer's behavior may reduce online advertising effectiveness. This paper formalizes a mechanism for this effect by allowing consumers past purchases to be indicative of the consumers' preferences for the good, and therefore, advertising that informs the consumers with past purchases of the related product is more likely to result in a consumer purchase. This then illustrates that privacy regulations may have negative welfare effects by reducing the benefit of advertising helping the consumers find the products that best fit them (that is, advertising turns into a sale, advertising is effective). Furthermore, this paper explores the early periods strategic effects of how to price given this later behavioral-based advertising.[3]

The remainder of the paper is organized as follows. The next section introduces the model, and Section 3 considers what happens after consumers have already bought the initial product, and studies the firm advertising targeting decisions. Section 4 considers the consumer and firm decisions in the initial period, aware that decisions in the initial period will affect the behavioral-based advertising in the later period. Section 5 discusses the possibility of both behavioral-based pricing and behavioral-based advertising. Section 6 concludes. The proofs are presented in the Appendix.

2. The Model

Consider a two-period market, where two related but different products are sold in sequence by a monopolist. Production costs are set at zero. Consumers have a valuation for each of the two products, a valuation θ_1 for the product in the first period, and a valuation θ_2 for a product in the

[2] As noted below, another important form of customer behavior not studied here that may be revealing of customer preferences is the consumer search behavior (see Armstrong and Zhou 2015 for some research on this dimension), or social network information (e.g., Iijima and Kamada 2016). Past purchase behavior may reveal some important dimension of consumer preferences, and this dimension is explored here.

[3] Other examples of empirical work on the effectiveness of forms of behavioral-based advertising include Dias et al. 2008, Hauser et al. 2009, Lambrecht and Tucker 2013, among others.

second period. The consumers and firm discount their second period payoffs with a discount factor δ_C and δ_F, respectively, with $\delta_C, \delta_F < 1$. This allows us to consider different degrees of looking forward by consumers and firms. The case of myopic consumers has $\delta_C = 0$, and is a particular case considered below. As consumers become more aware that the advertising messages that they receive may depend on their purchase behavior, it becomes more important to consider $\delta_C > 0$.[4]

The two-period model allows us to identify in a parsimonious way the market effects of the firms managing the size of the set of consumers who reveal to like the product in the first period, of the consumers behaving strategically with respect to their purchase behavior given the future advertising, and of the firms choosing their different advertising levels to consumers who purchased and did not purchase their previous product.[5]

We consider θ_1 positively correlated with θ_2 such that if a consumer places a greater value on the first-period product (s)he is also more likely to place a greater value on the second-period product. To model the correlation of preferences across periods we consider that the second period valuation is the same as in the first period with probability p. With probability $1-p$ the second-period valuation is independent of the first-period valuation, and with the same marginal distribution as the first-period valuation.[6] Denoting the cumulative marginal probability distribution of the first period valuation as $F(\theta_1)$, we have that the cumulative joint probability distribution of the first and second-period valuations is determined by

$$F_{\text{joint}}(\theta_1, \theta_2) = p \min[F(\theta_1), F(\theta_2)] + (1-p)F(\theta_1)F(\theta_2).$$

We assume that the support of θ_1 and θ_2 is $[0, \overline{\theta}]$.

The parameter p can be seen as an index of the correlation between the first and second-period valuations. We use the general notation $F(\theta_i)$ in most of the analysis. In order to obtain sharper results we sometimes restrict attention to the case in which $F(\theta_i)$ is the uniform distribution on $[0, 1]$.

Consumers are assumed to be completely aware of the first-period product, hence their decision to purchase in the first period is only a function of the first-period price, P_1.[7] In the second period

[4] An example of this effect at work is the behavior from a colleague to choose to buy a book from a retailer where the price was higher in order to benefit from advertising in the future from that retailer about books that that consumer may like.

[5] In future work, it would be interesting to investigate what happens in an infinite horizon game, where these effects are occurring simultaneously in each period.

[6] This particular dependence between the first and second period valuations is called a copula Family 11 (Joe, 1997, p. 148). See also Chen and Pearcy (2010) for a similar formulation.

[7] The assumption of consumers being fully aware of the first-period product is made to simplify the model, as

each consumer only becomes aware of the product if the firm advertises sufficiently to him. The cost to the firm of advertising to a consumer such that that consumer is likely to become aware with probability β is denoted as $K(\beta)$ with $K(0) = 0$, and $K'(\beta), K''(\beta) > 0$. This captures the idea that a consumer may need to receive several advertisements before becoming aware of a product. This modeling of the cost of advertising also restricts attention to the case in which the advertising costs are proportional to the number of consumers being advertised to, which could be seen as a first approach to considering behavioral-based advertising. It would also be interesting to consider the situation where the firm may not be able to do direct advertising to separate consumers. Examples of advertising costs that can be seen as proportional to the number of consumers advertised to are direct mailings, telephone calls, or purchase of ad space in web sites where the ad shown to each consumer depends on that consumer characteristics (for example, past purchases).[8]

The role of advertising creating awareness of products has been widely considered in the literature (e.g., Nelson 1974, Mahajan and Muller 1986). "Awareness" can be here broadly defined as a firm giving sufficient information about the existence of a product and/or its characteristics such that it makes the product appealing for consumers to consider purchasing it. As in any case of gaining information about a product, consumers can also be the ones actively searching for information on products that they may potentially be interested in. This possibility of consumer search is assumed away here for the paper to focus on the essential objective of investigating behavioral-based advertising.[9] One important dimension of the behavioral-based advertising considered here is that firms are able to advertise more intensively to consumers that are likely to value more the product. Alternatively, consumers could also sign up for some form of mailing lists to get information on products of potential interest. For example, some sellers could ask buyers if they would like to receive further information on related products in the future. However, there are important segments of markets where these activities are not possible or are greatly reduced (for example,

this captures the essence of the effects being considered, of the the initial consumption revealing something about preferences, and how that affects the behavioral-based advertising in the later periods. To better capture the real-world, one would have to consider also the advertising decisions in the initial periods.

[8] One could also think that a firm could make a fixed investment in an electronic recommendations system, and that once that investment is made the costs of making a recommendation to a consumer would be free. Even in that case, one could think that making a recommendation to a particular consumer would not be completely free as that consumer may start disregarding recommendations if too many are made.

[9] If consumer search is allowed for in the model, the conclusions presented would likely follow through for some parameter values, with some potential softening of the effects, but with the same main messages as capturing what is present in the real-world in terms of behavioral-based advertising. Another possibility is consumers who bought the product in the first period being more likely to become aware independently about the second period product. In terms of the model, this would just mean that the results should be interpreted in terms of final awareness levels in the second period, rather than in terms of advertising intensities, but the same messages will carry through. From anecdotal evidence, it seems that firms advertise more intensively to their previous customers, so this effect of automatic greater awareness to the other products of the same firm may not be too strong.

in some cases there are some "transaction costs," different than purchasing a product, that limit the extent of these activities), and this paper is focussing on those situations.[10] Furthermore, the information revealed about preferences by a purchase can be different than the information revealed by "signing up to a mailing list", and this paper focusses on the effects of information revealed about preferences by a purchase.[11]

In order to get sharper results we sometimes restrict attention to the case in which $K(\beta) = \frac{\alpha}{2}\beta^2$, where α is a parameter indicating the cost of advertising. We allow the possibility of the firm advertising differently to the consumers who bought or did not buy the product in the first period. We denote by β the probability of becoming aware of the consumers that did not buy the first period product, and by $\beta + \phi$ the probability of becoming aware of the consumers that bought the product in the first period. The term β can be seen as the base advertising received by all consumers, and the term ϕ is the extra advertising received by the consumers who purchased the first period product. In the main case considered we do not allow the firm to price differently across consumers that bought and did not buy in the first period. This could be potentially justified by arbitrage arguments. The case of firms pricing differently across consumers with different purchase histories is considered in Section 5.

We also consider the possibility that a consumer who receives advertising that creates a probability of that consumer becoming aware of the product with probability β getting a dis-utility of $G(\beta)/\beta$, with $G'(\beta) \geq 0$. In expected value, a consumer targeted with β advertising then gets an expected dis-utility of $G(\beta)$ as the consumer becomes aware with probability β. This accounts for the possibility that advertising can be a nuisance for consumers, and that more advertising to some targeted consumers can generate greater dis-utility for those consumers. This could also be seen as some possible costs of the loss of privacy. One extreme case to consider is one in which consumers do not have this dis-utility of receiving advertising, $G(\beta) = 0$ for all β. At some points we consider $G(\beta) = \gamma\beta$, where γ is a parameter indicating the importance of this dis-utility of advertising.

The utility of a consumer aware of the product, purchasing the product in the second period, and having received advertising β is then $U(\theta_2, \beta; \text{Aware and Buy}) = \theta_2 - P_2 - G(\beta)/\beta$, where P_2

[10] For example, many small sellers use online platforms to get access to consumers instead of owning their individual websites. They seldom offer the possibility of sign-up to mailing or email lists. Only past customers can get alerts of the latest offering from the seller. Examples include eBay in US and Taobao.com (owned by Alibaba group) in China. In addition, our survey results among online shoppers seem to suggest that subscribing to product-related email lists may not be a common practice for consumers to obtain information. The details of the survey are provided in the Appendix.

[11] As noted below, information revealed by other means, such as search or signing up to mailing lists, is important in several markets, but this is not the focus of this paper, and is left for future research. In the Appendix we present an illustration of how the information obtained about a customer through purchase is different than the one obtained by the customer signing up to a mailing list. We also present in the Appendix survey results which seem to suggest that the possibility of receiving personalized advertising based on past purchase can be valuable to consumers.

is the price charged by the firm in the second period. The utility if that consumer decides not to buy is $U(\theta_2, \beta; \text{Aware and Not Buy}) = -G(\beta)/\beta$. Note then that a consumer aware chooses to buy the product if and only if $\theta_2 \geq P_2$.

Note that in the real world unaware consumers may also potentially become aware by searching for related products. The analysis below can be seen as the extreme case where the costs of search are high enough such that consumers only become aware of the product if they receive advertising on it. We expect that the main ideas of the results would follow if some search were possible, but such analysis is beyond the scope of this paper. We can also think that unaware consumers are also unaware about how to search for related products.[12]

3. Firm Decisions Given Consumer Information

Consider now the firm's problem in the second period. Let θ_1^* be the threshold valuation in the first period such that a consumer with $\theta_1 \geq \theta_1^*$ bought the product in the first period, and all other consumers did not buy the product in the first period. This will be confirmed below.

The second period price can either be (1) $P_2 > \theta_1^*$, (2) $P_2 = \theta_1^*$, or (3) $P_2 < \theta_1^*$, which has to be checked below at the optimum second-period price. Defining θ^* as the θ that maximizes $\theta[1-F(\theta)]$, the static monopoly pricing problem, we will have that case (1) corresponds to the case in which $\theta_1^* < \theta^*$ and case (2) corresponds to the case in which $\theta_1^* \geq \theta^*$ but θ_1^* is not too large, in a sense that is discussed below, and case (3) corresponds to the case in which $\theta_1^* > \theta^*$ and θ_1^* is large. We present conditions in Section 4.2 under which cases (1), (2) and (3) can occur in equilibrium. Case (1) can occur if $G(\beta)$ is sufficiently convex, $\beta G'(\beta) > G(\beta)$. For case (2), we can also show that we can have either $\theta_1^* = \theta^*$ or $\theta_1^* > \theta^*$. Case (2) occurs when the dis-utility of receiving advertising is not too large. If the firm is more forward-looking than the consumers and/or the correlation of preferences is large enough we have $\theta_1^* > \theta^*$. Otherwise, the equilibrium has $\theta_1^* = \theta^*$. Case (3) occurs when the dis-utility of receiving advertising is sufficiently large and the consumers are sufficiently forward-looking.

3.1. Cases $P_2 > \theta_1^*$ and $P_2 = \theta_1^*$.

For the case in which $P_2 > \theta_1^*$ the second-period problem of the firm can then be written as

$$\max_{P_2,\beta,\phi} \pi_2(P_2, \beta, \phi; \theta_1^*) = (\beta + \phi)p[1 - F(P_2)]P_2 + (\beta + \phi)(1 - p)[1 - F(\theta_1^*)][1 - F(P_2)]P_2 +$$

[12]For a study of advertising and consumer search in a different setting see Mayzlin and Shin (2011).

$$\beta F(\theta_1^*)(1-p)[1-F(P_2)]P_2 - F(\theta_1^*)K(\beta) - [1-F(\theta_1^*)]K(\beta+\phi). \tag{1}$$

The first term in the objective function represents the consumers who bought in the first period and did not change their preferences. The second term represents the set of consumers who bought in the first period but changed their preferences. The third term represents the set of consumers who did not buy a product in the first period but changed their preferences, and could potentially be interested in buying the product in the second period.

The first order condition with respect to price in the problem above is the one from the traditional monopoly problem, $1-F(P_2)-P_2f(P_2)=0$ where $f(P_2)$ is the density marginal probability function of θ_2. Let P_2^* be the solution to this first order condition, which as defined above is $P_2^*=\theta^*$. It is then immediate that this case occurs if $\theta_1^*<\theta^*$. For the case where $F(\theta_2)$ is the uniform distribution on $[0,1]$ we have $P_2^*=\frac{1}{2}$.

From the first order conditions with respect to β and ϕ one obtains

$$(1-p)[1-F(P_2^*)]P_2^* = K'(\beta) \tag{2}$$

$$[1+\frac{pF(\theta_1^*)}{1-F(\theta_1^*)}][1-F(P_2^*)]P_2^* = K'(\beta+\phi). \tag{3}$$

Consider now the case where $P_2=\theta_1^*$. In this case the problem for the firm is the same as (1) with the restriction that $P_2=\theta_1^*$. We can then also obtain the first order conditions with respect to β and ϕ reduce to (2) and (3) with $P_2^*=\theta_1^*$.

From this one can immediately obtain for both case (1) and (2) that $\phi>0$, that is the consumers that purchased in the first period are advertised to more intensively. Furthermore, when the correlation of preferences between the first and second period goes to one, $p\to 1$, there is no base advertising, $\beta\to 0$, and the only advertising is that targeted to the consumers who bought in the first period.

Finally, it is interesting to investigate how the intensity of advertising changes with the extent of demand in the first period. For the case where $P_2^*>\theta_1^*$, totally differentiating (3) one can get that increasing θ_1^* leads to greater advertising to the consumers that purchased in the first period, while β^* is independent of how many consumers bought the product in the first period (independent of θ_1^*). For the case where $P_2^*=\theta_1^*$, one can obtain that the advertising intensity to all consumers (having purchased or not purchased in the first period) is increasing in θ_1^*, and that the increased advertising to the consumers that purchased in the first period is increasing in θ_1^* if the function $K()$ is sufficiently close to a quadratic function. We state these results in the following proposition.

Proposition 1: *Suppose $P_2 \leq \theta_1^*$. The intensity of advertising to the consumers who purchased in the first period is greater the smaller the number of consumers who purchased in the first period. The intensity of advertising to the consumers who did not buy in the first period does not depend on the number of consumers who bought in the first period if $\theta_1^* < \theta^*$, and decreases in how many consumers bought in the first period if $\theta_1^* \geq \theta^*$. Finally, if the function $K()$ is sufficiently close to a quadratic function, the increased advertising to the consumers who bought in the first period (ϕ) decreases in the number of consumers who bought in the first period.*

The result indicates how a lower number of consumers that bought the product in the previous period allows for more effective (and therefore more) advertising to the consumers that have higher valuation. Another force is that the smaller the number of consumers who bought in the first period, the lower the waste of advertising sent to first period buyers who change their preferences in the second period.

3.2. Case $P_2 < \theta_1^*$.

Consider now the case of $P_2 < \theta_1^*$ which, as we discuss in Section 4.2, can occur in equilibrium if the dis-utility of advertising is sufficiently large, if the consumers are sufficiently forward-looking, and if the cost of advertising is not too large. In this case the problem of the firm can be stated as:

$$\max_{P_2,\beta,\phi} \pi_2(P_2,\beta,\phi;\theta_1^*) = (\beta+\phi)p[1-F(\theta_1^*)]P_2 + (\beta+\phi)(1-p)[1-F(\theta_1^*)][1-F(P_2)]P_2 + \beta F(\theta_1^*)(1-p)[1-F(P_2)]P_2 + \beta p[F(\theta_1^*)-F(P_2)]P_2 - F(\theta_1^*)K(\beta) - [1-F(\theta_1^*)]K(\beta+\phi). \quad (4)$$

This objective function can be seen as a function of P_2 that has a term in $[1-F(P_2)]P_2$ and a term in $\phi p[1-F(\theta_1^*)]P_2$, which means that, in this case, the optimal P_2 satisfies $\theta^* < P_2 < \theta_1^*$. One can write the first order condition for P_2 as

$$[\beta + \phi(1-p)[1-F(\theta_1^*)]][1-F(P_2) - P_2 f(P_2)] + \phi p[1-F(\theta_1^*)] = 0. \quad (5)$$

Note that if θ_1^* is too low, the optimal P_2 is above θ_1^*, which contradicts the assumption of this case of $P_2 < \theta_1^*$. So, for this case to occur θ_1^* will have to be relatively large. Differentiating the objective function with respect to β and ϕ one also obtain the conditions for the advertising expenditures as a function of θ_1^*.

4. First Period Decisions

4.1. Consumers' Decisions in the First Period

Consider now the decision-making of the consumers in the first period. A marginal consumer with valuation θ_1 deciding to buy in the first period gets an expected payoff of

$$\theta_1 - P_1 + \delta_C(\beta+\phi)p\max[\theta_1 - P_2^*, 0] + \delta_C(\beta+\phi)(1-p)\int_{P_2^*}^{\overline{\theta}}(\theta_2 - P_2^*)\, dF(\theta_2) - \delta_C G(\beta+\phi). \quad (6)$$

If a consumer in the first period decides not to buy he gets an expected payoff of

$$\delta_C\beta[p\ \max[\theta_1 - P_2^*, 0] + (1-p)\int_{P_2^*}^{\overline{\theta}}(\theta_2 - P_2^*)\, dF(\theta_2)] - \delta_C G(\beta). \quad (7)$$

A consumer indifferent between buying and not buying in the first period is determined by making (6) equal to (7), which determines the threshold valuation θ_1^*. Note that under both cases (1) and (2) ($P_2^* > \theta_1^*$ or $P_2^* = \theta_1^*$) the max terms in both (6) and (7) are equal to zero. Defining $y \equiv \int_{P_2^*}^{\overline{\theta}}(\theta_2 - P_2^*)\, dF(\theta_2)$, one can then obtain that the threshold θ_1^* satisfies

$$\theta_1^* = P_1 - \delta_C\phi^*\{p\max[\theta_1^* - P_2, 0] + (1-p)y\} + \delta_C[G(\beta^*+\phi^*) - G(\beta^*)], \quad (8)$$

and that only the consumers with $\theta_1 \geq \theta_1^*$ purchase the product in the first period.

In order to gain a better intuition for this threshold θ_1^* note that $p\max[\theta_1^* - P_2, 0] + (1-p)y$ is the expected surplus of a marginal consumer receiving advertising, as that consumer gets some surplus when he does not change preferences if $P_2 < \theta_1^*$, and, when he changes preferences, which happens with probability $(1-p)$, his expected surplus is y. In order to understand if advertising can be useful for consumers, we need to compare this expected surplus with the expected marginal disutility of receiving advertising, $G'(\beta)$. We then say that more advertising is beneficial to consumers if $p\max[\theta_1^* - P_2, 0] + (1-p)y > G'(\beta)$, the expected surplus of receiving one additional unit of advertising is greater than the inconvenience of seeing an additional unit of advertising. Putting this together with (8) we can state the following result.

Proposition 2: *If more advertising is beneficial to consumers, $p\max[\theta_1^* - P_2, 0] + (1-p)y > G'(\beta)$ for all β, then the threshold valuation θ_1^* is below the first-period price P_1, some consumers with product valuation below the price charged purchase in the first period.*

The intuition is that consumers anticipate the potential benefit of being advertised to in the second period a product that they potentially value by more than the price charged. From this one can easily obtain that if $G(\beta) = 0$ for all β, receiving advertising does not cause any inconvenience, and the result in the proposition also follows immediately.

From equation (8) note also that if the consumers are not very forward-looking (δ_C is small), or if the costs of advertising are high (which results in ϕ^* small), the threshold θ_1^* is close to P_1. In this case, as we will see below we end up in the case with $P_2 = \theta_1^* > \theta^*$. Note also that when $G(\beta)$ is small, as we have $\theta_1^* < P_1$ we are less likely to be in the case in which $P_2 < \theta_1^*$ as θ_1^* is already low. In fact, as we show below, if $G(\beta)$ is small we will be in the case with $P_2 = \theta_1^*$. On the other hand, if $G(\beta)$ is large, θ_1^* can end up being greater than P_1 and high, and in that case we may end in the case of $P_2 < \theta_1^*$. In that case, consumers refrain from buying in the first period because they know that they will be targeted with advertising in the second period, and advertising creates a large dis-utility.

Checking the effect of the first-period price on the threshold first-period valuation (which is directly related to the demand in the first period) for the case of $P_2^* > \theta_1^*$ one can obtain how the price in the first period affects θ_1^*

Proposition 3: *Suppose that $\theta^* > \theta_1^*$ (such that $P_2^* > \theta_1^*$) and that more advertising is beneficial to consumers, $(1-p)y > G'(\beta)$ for all β. Then demand is less sensitive to price when there is behavioral-based advertising and consumers are forward looking than when either there is no behavioral-based advertising or consumers are myopic.*

When $P_2^* > \theta_1^*$, demand in the first period, at the margin, does not affect the price that will be set in the second period. Then, a higher price in the first period leads to greater difference in advertising between the previous customers and the other customers, which makes it more appealing for consumers to buy in the first period if more advertising is beneficial to consumers.

For the uniform distribution example considered above with $K()$ quadratic and $G()$ linear one can obtain $\beta^* = \frac{1-p}{4\alpha}$, $\phi^*(\theta_1^*) = \frac{p}{4\alpha(1-\theta_1^*)}$, $y = \frac{1}{8}$, and $\theta_1^* = P_1 - \delta_C p \frac{1-p-8\gamma}{32\alpha(1-\theta_1^*)}$. In this case the condition for consumers to benefit from advertising is $1 - p > 8\gamma$. From above we can obtain

$$\theta_1^* = 1 - \frac{1 - P_1 + \sqrt{(1-P_1)^2 + \delta_C p \frac{1-p-8\gamma}{8\alpha}}}{2}, \tag{9}$$

which determines the demand in the first period.

For the case when $P_2^* = \theta_1^*$, we can check the effect of the first-period price on the threshold first-period valuation. In this case, one can find situations where demand is more price sensitive than in the case of myopic consumers, even if advertising is beneficial to consumers. In particular, we have the following result:

PROPOSITION 4: *Suppose that $P_2^* = \theta_1^* \geq \theta^*$, that $G()$ is linear, and that more advertising is beneficial to consumers, $(1-p)y > G'(\beta)$ for all β. Then demand is more sensitive to price when there is behavioral-based advertising and consumers are forward looking than when either there is no behavioral-based advertising or consumers are myopic if and only if $[(1-p)y - G'(\beta^* + \phi^*)]\frac{\partial \phi^*}{\partial \theta_1^*} < (1-p)\phi^*[1 - F(\theta_1^*)]$.*

To see an example of this condition, for the uniform distribution example considered above with $K()$ quadratic and $G()$ linear one can obtain $\beta^* = \frac{1-p}{\alpha}\theta_1^*(1-\theta_1^*)$, $\phi^*(\theta_1^*) = \frac{p}{\alpha}\theta_1^*$, $y = \frac{(1-\theta_1^*)^2}{2}$, and $\theta_1 * = P_1 - \delta_C p(1-p)\frac{\theta_1^*(1-\theta_1^*)^2}{2\alpha} + \delta_C \gamma \theta_1^* \frac{p}{\alpha}$. To check that demand can be more price sensitive than with myopic consumers, (vii) in the Appendix can be greater than one, consider $\gamma = 0$, and then $\frac{\partial \theta_1^*}{\partial P_1} > 1$ if $y\frac{\partial \phi^*}{\partial \theta_1^*} - \phi^*[1 - F(\theta_1^*)] < 0$. This latter condition is equivalent to $\theta_1^* > \frac{1}{3}$.

To gain intuition for this additional effect when $P_2^* = \theta_1^*$, note that now when the price in the first period increases, there will also be an increase in a price of the second period, and realizing this, forward-looking consumers become more price sensitive to the first-period prices.

4.2. Firm's Decision in the First Period.

Consider now the problem of the firm in the first period. The problem of the firm is

$$\max_{P_1} P_1[1 - F(\theta_1^*)] + \delta_F \pi_2(P_2^*, \beta^*(\theta_1^*), \phi^*(\theta_1^*); \theta_1^*). \tag{10}$$

For the case when $P_2^* > \theta_1^*$ the first order condition of this problem, using the envelope theorem, can be written as:

$$1 - F(\theta_1^*) - f(\theta_1^*)\frac{\partial \theta_1^*}{\partial P_1}\{P_1 + \delta_F[\phi^*(\theta_1^*)(1-p)P_2^*[1 - F(P_2^*)] - K(\beta^*(\theta_1^*) + \phi^*(\theta_1^*)) + K(\beta^*(\theta_1^*))]\} = 0. \tag{11}$$

Note that $\phi^*(\theta_1^*)(1-p)P_2^*[1-F(P_2^*)] - K(\beta+\phi) + K(\beta) < 0$ as $(1-p)P_2^*[1-F(P_2^*)] = K'(\beta)$ by (2) and $\phi K'(\beta) < K(\beta+\phi) - K(\beta)$ given that $K'(\beta+x) > K'(\beta)$ for all $x > 0$. This would lead to θ_1^* to be strictly greater than θ^* if $\frac{\partial \theta_1^*}{\partial P_1}$ is close to one, which would mean that $P_2^* > \theta_1^*$ could not be an equilibrium. For the case when consumers are myopic, $\delta_C = 0$, we have that $\theta_1^* = P_1$ and $\frac{\partial \theta_1^*}{\partial P_1} = 1$, such that we obtain that the optimal θ_1^* is greater than the solution to $1 - F(\theta_1^*) - \theta_1^* f(\theta_1^*) = 0$, which represents the case without behavioral-based advertising. Then, in this case, fewer consumers buy the product in the first period than in the case without behavioral-based advertising, and we have that the assumed condition $P_2^* > \theta_1^*$ is violated. We can, however, find conditions under which $P_2^* > \theta_1^*$ occurs in equilibrium. For example, if $K()$ is quadratic, $\delta_F = 0, p = 1$, and $\beta G'(\beta) > G(\beta)$, we can get that $\theta_1^* < \theta^*$ and $P_2^* > \theta_1^*$.

For the uniform distribution example with quadratic $K()$ and linear $G()$, we can re-write (11), and checking the left hand side at $\theta_1^* = \frac{1}{2}$ we can find it to be positive, which indicates that the optimal θ_1^* is strictly greater than $\frac{1}{2}$ which violates the condition that $P_2^* > \theta_1^*$. We summarize these results in the following proposition.

Proposition 5: *If $\frac{\partial \theta_1^*}{\partial P_1}$ is close to one, the market equilibrium cannot have $P_2^* > \theta_1^*$. For the uniform distribution example with quadratic $K()$ and linear $G()$, the equilibrium cannot have $P_2^* > \theta_1^*$.*

Consider now the case when $P_2^* = \theta_1^*$. For this case, the first order condition of the firm's first-period problem can be written as

$$1 - F(\theta_1^*) - \frac{\partial \theta_1^*}{\partial P_1}\{P_1 f(\theta_1^*) - \delta_F[1 - F(\theta_1^*)][\beta^*(\theta_1^*) + \phi^*(\theta_1^*) - \phi^*(\theta_1^*)(1-p)F(\theta_1^*)] +$$

$$\delta_F f(\theta_1^*)[\beta^*(\theta_1^*)\theta_1^* + \phi^*(\theta_1^*)\theta_1^*[p + 2(1-p)(1 - F(\theta_1^*))] + K(\beta^*(\theta_1^*)) - K(\beta^*(\theta_1^*) + \phi^*(\theta_1^*))]\} = 0. \quad (12)$$

Note that if consumers are myopic we can obtain that the left hand side of (12) is positive when evaluated at $\theta_1^* = \theta^*$, that is, that the optimal θ_1^* with the constraint that $\theta_1^* \geq \theta^*$ is strictly greater than θ^*. To see this note that for myopic consumers $\frac{\partial \theta_1^*}{\partial P_1} = 1$, and, therefore, the left hand side of (12) when evaluated at $\theta_1^* = \theta^*$, is equal to

$$f(\theta^*)[-\theta^* \phi^*(\theta^*)(1-p)[1 - F(\theta^*)] + K(\beta^*(\theta^*) + \phi^*(\theta^*)) - K(\beta^*(\theta^*))], \quad (13)$$

as $1 - F(\theta^*) - \theta^* f(\theta^*) = 0$. Given that $K'(\beta^*(\theta^*)) = (1-p)[1 - F(\theta^*)]\theta^*$ and that $K'' > 0$, one can then obtain that (13) is greater than zero.

For the uniform distribution example with quadratic $K()$ and linear $G()$, we can re-write (12) as (ix) in the Appendix.

For the case of $\theta_1^* = \theta^* = \frac{1}{2}$, we can reduce the left hand side (ix) to

$$-\delta_C \frac{p(1-p)}{8\alpha} + \delta_F \frac{p^2}{8\alpha}.$$

For this case we can also see that if the firm is myopic then we have the left hand side of (ix) is negative when evaluated at $\theta_1^* = \theta^*$, which means that in equilibrium $\theta_1^* = \theta^* = \frac{1}{2}$. If the consumers and firms have the same discount factor, we have that the sign of (ix) is equal to the sign of $p(2p-1)$. For this case we then have that if $p < \frac{1}{2}$ the equilibrium has $\theta_1^* = \theta^* = \frac{1}{2}$, and that if $p > \frac{1}{2}$ the equilibrium has $\theta_1^* > \theta^* = \frac{1}{2}$. For $p \to 1$, the case when almost no consumer changes preferences, and γ small, the dis-utility created by advertising is negligible, we can obtain that $\frac{1}{2} < \theta_1^* < \frac{2}{3}$.

Possibility of $P_2^* < \theta_1^*$.

To check for the possible equilibrium, we now only need to check if the case of $P_2^* < \theta_1^*$ is possible. As noted above this case requires θ_1^* to be high. Note first that in this case, a greater θ_1^* lead to lower profits in the second period, as the firm cannot do as good a matching of advertising to the consumers that have a valuation above the second period's price. To see this, differentiating π_2 with respect to θ_1^*, using the envelope theorem, one obtains:

$$\frac{1}{f(P_2)} \frac{\partial \pi_2}{\partial \theta_1^*} = -\phi[1 - F(P_2) + pF(P_2)]P_2 + K(\beta + \phi) - K(\beta) < 0, \tag{14}$$

as $[1 - F(P_2) + pF(P_2)]P_2 = K'(\beta + \phi)$ and $K''() > 0$.

The condition for the first period price can be set as

$$1 - F(\theta_1^*) - P_1 f(\theta_1^*) \frac{\partial \theta_1^*}{\partial P_1} + \delta_F \frac{\partial \pi_2}{\partial \theta_1^*} \frac{\partial \theta_1^*}{\partial P_1} = 0. \tag{15}$$

For $\delta_C \to 0$, we have $\theta_1^* \to P_1$, $\frac{\partial \theta_1^*}{\partial P_1} \to 1$, and therefore, given that $\frac{\partial \pi_2}{\partial \theta_1^*} < 0$, we have that $1 - F(\theta_1^*) - \theta_1^* f(\theta_1^*) > 0$, which means that $\theta_1^* < \theta^*$, which is not possible in this case. Then, when $\delta_C \to 0$ we cannot be in the case of $P_2 < \theta_1^*$ and we have $P_2 = \theta_1^*$. Note also that this is also true for any δ_C with $p \to 0$, as in that case we also have $\theta_1^* \to P_1$. The same holds if the costs of advertising are high, such that the firm chooses low ϕ (for example, in the quadratic $K()$ example, if α is sufficiently high).

However, it is possible to get to a situation where $P_2^* < \theta_1^*$ is an equilibrium. The intuition is that if the advertising dis-utility is sufficiently high, for any given price the only consumers that are willing to purchase have a high valuation (high θ_1^*), such that the firm chooses a second period price below θ_1^*. To see this in the simplest way consider the uniform distribution case, with quadratic $K()$ and linear $G()$, and suppose $p = \delta_C = \delta_F = \alpha = \gamma = 1$. In this case one can get $P_2^* = \frac{3\theta_1^* - \sqrt{9\theta_1^{*2} - 8\theta_1^*}}{4}$, $\phi^* = \frac{P_2^{*2}}{\theta_1^*}$, $\beta^* = P_2^*(1 - P_2^*) - P_2^{*2}\frac{1-\theta_1^*}{\theta_1^*}$. Note that for this case to be possible we need $\theta_1^* \geq \frac{8}{9}$.

To check this possibility consider the case of $\theta_1^* = 8/9$. Then, we would have the price in the first period $P_1 = 1/2$, the advertising to the consumers that did not purchase in the first period $\beta^* = 1/6$, the advertising to the consumers that purchased in the first period $\beta^* + \phi^* = 2/3$, and a price in the second period at $P_2^* = 2/3$. This would then lead to a profit in the first period of $\pi_1 = 1/18$, and a profit in the second period of $\pi_2 = 1/27$, for a total discounted present value of profits of $\pi_1 + \pi_2 = 5/54$.

Consider now the profit that could be obtained if $\theta_1^* < 8/9$, and we were in the case above with $P_2 = \theta_1^*$. In that case, as $\delta_F p > \delta_C(1-p)$ we would be in the case with $\theta_1^* > \theta^*$, and from condition (ix) we can get the marginal valuation $\theta_1^* = 2/3$. We could then obtain that in that case $P_1 = 0, P_2 = 2/3, \beta = 0$, and $\phi = 2/3$. The profit in the first period would then be $\pi_1 = 0$, the profit in the second period would be $\pi_2 = 2/27$, and the present value of profits would then be $\pi_1 + \pi_2 = 2/27$, which is lower than the present value of profits presented above that can be obtained when $\theta_1^* = 8/9$ and we have $P_2^* < \theta_1^*$.[13]

Therefore, this shows that there are parameter values for which we can have $P_2^* < \theta_1^*$ in equilibrium. The intuition is that if the dis-utility of advertising is too high, the first period price that would need to be offered such that in the second period we would have $P_2 = \theta_1^*$ would have to be so low that the firm prefers to offer a higher first period price, understanding that in the second period it will offer a price below the first period marginal valuation consumer. The formal condition to be in this region is too tedious to be presented here.

Figure 1 illustrates the market equilibrium as a function of γ for the case with $p = \delta_F = \delta_C = \alpha = 1$. For γ below a certain threshold (close to .82 in this case), we are in the case of $P_2^* = \theta_1^* > \theta^*$. In this case, as the dis-utility of advertising γ increases, the firm has to decrease the first period price and accept lower demand in the first period (θ_1^* is increasing), which leads to increasing second period price, and increasing advertising to the consumers who purchased in the first period. In this region advertising to the consumers who did not purchase in the first period is set at zero, as no consumers change preferences from the first to the second period ($p = 1$ in this illustration). At some threshold, γ is so high, and the first period price so low, that the firm decides to change where

[13] The optimal θ_1^* for these parameter values is in fact greater than 8/9, but the suboptimal $\theta_1^* = 8/9$ is enough to illustrate this possibility.

$\theta_1^* > P_2^* > \theta^*$, with an increased first period price, a jump upwards in θ_1^* and positive advertising advertising to consumers who did not purchase in the first period, as now some of these consumers may be interested in buying the product given that $P_2^* < \theta_1^*$. Figure 2 shows the different regions of the market equilibrium that may be possible for the uniform distribution case with quadratic $K(.)$ and linear $G(.)$ as a function of p and γ. For p large and γ small we have $P_2^* = \theta_1^* > \theta^*$, in equilibrium, the realized demand in the first period is lower than when there is no behavioral-based advertising, and in the second period, the firm prices as only going after the consumers who purchased in the first period and did not change their preferences. For p and γ small we have $P_2^* = \theta_1^* = \theta^*$, in equilibrium, the realized demand in the first period is the same as when there is no behavioral-based advertising, and in the second period, the firm prices as only going after the consumers who purchased in the first period and did not change their preferences. Finally, for large γ we have $\theta^* < P_2^* < \theta_1^*$, in equilibrium, the realized demand in the first period is lower than when there is no behavioral-based advertising, and in the second period, the firm prices to go also after the consumers who did not purchase in the first period and did not change their preferences. The threshold γ depends on p in a non-monotonic way. When p falls from the high levels the threshold γ goes down as the potential benefit of a lower price decreases with a lower p. But at some point as p keeps on falling the threshold γ starts rising, as the firm does not find as beneficial to advertise much more to the consumers who purchased in the first period than to the consumers who did not purchase in the first period. In fact, we then have that as when $p \to 0$ the threshold γ goes to infinity.

Note that in this case of γ large consumers may receive a negative overall utility in the second period because of the dis-utility of getting advertising. However, consumers may not be able to avoid the advertising dis-utility just by not choosing to buy the product. Consumers respond by choosing to buy or not to buy in each period to maximize their surplus that they can control. Accounting for the first period surplus, the consumer overall surplus can be positive even though negative in the second period. This possibility can be relevant for categories, where consumers receive a positive surplus in an initial purchase, in spite of the high advertising dis-utility that they will receive in later periods. This case of the dis-utility of advertising large may also make more relevant the possible use by consumers (at a cost) of technologies that can potentially block advertising. Studying this possibility is left for future research.

In the remainder of the paper we will now concentrate in the case with γ small (or δ_C small, or α large) such that we are in the case of $P_2 = \theta_1^*$.

4.3. Market Equilibrium Details

4.3.1. Market Coverage as without Behavioral-Based Advertising

Given the analysis above we can now characterize in greater detail the market equilibrium behavior for the case of γ small, focusing on the uniform distribution example with quadratic $K()$ and linear $G()$.

Consider first the case of $\delta_F p - \delta_C(1-p) < 0$, which leads to the case of $\theta_1^* = \frac{1}{2}$, the equilibrium first-period demand being $\frac{1}{2}$. In this case the optimal first period price is $P_1^* = \frac{1}{2} + \delta_C \frac{(1-p)p}{16\alpha} - \delta_C \frac{\gamma p}{2\alpha}$. In comparison to the case with no behavioral-based advertising we have the same market efficiency in the first period, but the first-period consumer surplus can be lower if advertising does not cause too much dis-utility (low γ).

The second-period price is $P_2^* = \frac{1}{2}$, the same price as when there is no behavioral-based advertising, and the base advertising is $\beta^* = \frac{1-p}{4\alpha}$ which is below the base advertising when there is no behavioral-based advertising. The extra advertising to the consumers that purchased the product in the first period is $\phi^* = \frac{p}{2\alpha}$. Note that for this case and example, the total advertising is the same as when there is no behavioral-based advertising, $\frac{1}{4\alpha}$.

Note that even though the first-period realized demand under behavioral-based advertising is the same as with no behavioral-based advertising, the second-period demand is greater under behavioral-based advertising, and equal to $\frac{1+p^2}{8\alpha}$, increasing in p, the index by which preferences are positively correlated across periods. So, if receiving advertising does not cost too much dis-utility to consumers, firms benefit from behavioral-based advertising for two reasons: (1) higher price in the first period, and (2) higher demand in the second period. However, if consumers suffer too much dis-utility from receiving advertising (high γ), firms could be hurt from the possibility of behavioral-based advertising.

The equilibrium present value of profits can be obtained to be $\frac{1}{4} + \frac{\delta_C p(1-p) - 8\delta_C \gamma p + \delta_F(1+p^2)}{32\alpha}$ which can be compared with the profit without behavioral-based advertising, $\frac{1}{4} + \frac{\delta_F}{32\alpha}$. From these we can obtain the following results.

Proposition 6: *Behavioral-based advertising leads to higher profits than without behavioral-based advertising if $\delta_C(1-p) - 8\delta_C\gamma + \delta_F p > 0$. The present value of profits is increasing in the consumer patience, δ_C, if the dis-utility of advertising is not too high, $\gamma < \frac{1-p}{8}$. The present value of profits is increasing in the correlation of preferences through time, p, if $\delta_C(1-2p-8\gamma) + 2p\delta_F > 0$.*

If the dis-utility of advertising is not too high, γ small, the profits increase in the consumer patience, δ_C, as the firm can charge a higher price in the first period. If the dis-utility of advertising is high, then the firm has to offer a lower price in the first period. The condition for the correlation of preferences over time to affect profits positively is interesting. If the consumers and the firm have the same discount factor, the condition is satisfied so long as the dis-utility of advertising is not too high, $\gamma < 1/8$. However, if the firm is myopic, then the present value of profits will be decreasing in p for $p > \frac{1}{2}$. For this case, the firm does not have any benefit from the future profit, and has to reduce prices in the first period with higher p, as the likelihood of getting a bigger surplus in the second period is now lower.

Looking at the expected consumer surplus, one can obtain that it is equal to $\frac{1}{8} + \delta_C \frac{1-p+2p^2}{32\alpha} - \delta_C \frac{\gamma(1-p)}{4\alpha}$. Comparing it with the case of no behavioral-based advertising, $\frac{1}{8} + \frac{\delta_C}{4\alpha}(\frac{1}{8} - \gamma)$, one can obtain that consumer surplus is lower than with no behavioral-based advertising if and only if $\gamma < \frac{1-2p}{8}$ which can only happen if $p < \frac{1}{2}$. The intuition is that, when consumer preferences are likely to change from period to period ($p < \frac{1}{2}$) the benefits of receiving more advertising in the second period are extracted from the consumers with higher prices in the first period. On the other hand, for the case of $p < \frac{1}{2}$, if receiving advertising creates substantial dis-utility to consumers (high γ) then firms have to offer lower prices to the consumers in the first period, which could end up being beneficial to the consumers overall. When $p > \frac{1}{2}$ consumers always benefit from behavioral-based advertising as the preferences remain stable and the consumers receive special advertisements of a product that they are likely to appreciate.

Note also that social welfare is now higher with behavioral-based advertising as advertising is now more targeted at consumers with higher valuation, given $p > 0$.

4.3.2. Market Coverage Less than without Behavioral-Based Advertising

Consider now the case of $\delta_F p - \delta_C(1-p) > 0$, which leads to the case of $\theta_1^* > \frac{1}{2}$. For this case θ_1^* is determined by (12), which for the uniform example can be represented by (ix). In order to simplify the presentation we restrict attention to the case in which the dis-utility of advertising γ is small compared to $1-p$, such that consumers benefit from receiving advertising, preferences in the second period are almost perfectly correlated with preferences in the first period, $p \to 1$, and consumers and firm have the same discount factor, $\delta_C = \delta_F$.

We can then obtain how demand in the first period $(1-\theta_1^*)$, price in the first period, advertising in the second period, and the present value of profits vary with the different parameters. The results on the comparative statics on first period demand and profits are presented in the following proposition.

Proposition 7: *Consider the uniform distribution example with quadratic $K()$ and linear $G()$ and suppose $\delta_F p - \delta_C(1-p) > 0$, $\gamma < \frac{1-p}{18}$, $p \to 1$, and $\delta_F - \delta_C \to 0$. Then demand in the first period $(1-\theta_1^*)$ is increasing in the consumers discount factor δ_C, and in the cost parameter of advertising α, and decreasing in the firm discount factor δ_F, in the probability of the same preferences across periods p, and in the dis-utility of advertising γ. The present value of profits increases in the consumers discount factor δ_C, in the firm discount factor δ_F, and in the probability of the same preferences across periods p, and decreases in the dis-utility of advertising γ, and in the cost parameter of advertising α.*

Let us now provide some intuition on these results. Under the assumption that the dis-utility of advertising is not too large, such that consumers benefit from receiving advertising, we have that an increase in δ_C increases the benefit from the consumers purchasing in the first period which leads to more consumers purchasing in the first period, and a greater price in the first period. More consumers purchasing in the first period reduces the valuation of the marginal consumers purchasing in the first period, which leads to less benefit of advertising to consumers that purchased in the first period. As the marginal consumer buying in the first period is lower, this leads to a lower price in the second period (but still above the monopoly price), which means that there is a greater potential profit from the consumers who did not purchase in the first period, leading to greater advertising to the consumers who did not purchase in the first period. Overall, given the higher valuation to consumers of advertising, the present value of profits increases in the patience of consumers, δ_C.

Consider now the effect of increasing the patience of the firm, δ_F. An increase in δ_F makes the firm care more about second period profits, which can be improved with targeting the more intense advertising to the consumers that most value the product. This can be done having fewer consumers purchase in the first period, which leads to an increase in the valuation of the marginal consumer purchasing in the first period, θ_1^*. To obtain this the firm increases the price in the first period. Having consumers with greater valuation buying in the first period then leads the firm to advertise more intensively to consumers who purchased in the first period, and less intensively to the consumers who did not purchase in the first period. Overall, the effect on the present value of profits is positive, as the firm gives greater value to the second-period profits.

The effect of increasing the probability of the preferences remaining the same (correlation of preferences through time) makes the targeting of advertising more beneficial. Then the firm has a greater advantage of identifying the consumers with a greater valuation, which leads to a greater valuation of the marginal consumer in the first period (greater θ_1^*), lower demand in the first period, which is obtained with a higher price in the first period. Then, with a better identification of the higher valuation consumers, the firm raises its advertising intensity to the consumers who purchased in the first period, and decreases its advertising intensity to the consumers who did not purchase in the first period. These effects naturally lead to a greater present value of profits, resulting from a more effective targeting of advertising.

Increasing the dis-utility of advertising (increasing γ) lowers the benefit to consumers of buying the product in the first period. This results in fewer consumers purchasing the product in the first period (greater θ_1^*). If firms advertise intensively (low α), this can become a big cost for consumers, and the firm also lowers its first period price. Otherwise (high α), the firm just lowers the cost of increased γ reducing the number of consumers that purchase the product in the first period, and increasing the first-period price. At the time when the firm decides on advertising the consumers already made the first period purchasing decisions, therefore the advertising decisions only take into account the composition of consumers who purchased or did not purchase in the first period, leading to more intensive advertising to the consumers who purchased the product in the first period, and less intensive advertising to the consumers who did not purchase the product in the first period. Overall, given the extra cost to consumers, the present value of profits falls.

Finally, increasing the advertising cost parameter reduces the benefits of identifying the consumers with the higher valuation. This then leads to a lower valuation of the marginal consumer purchasing in the first period (lower θ_1^*), there is an increased demand in the first period. In order to increase demand in the first period, the firm then lowers the first-period price. With a greater advertising cost parameter, the firm obviously then chooses to advertise less intensively (to both the consumers who purchased in the first period, and who did not purchase in the first period), and the equilibrium present value of profits also decreases.

4.4. Different Firms in Different Periods

In some important market situations the firm that takes advantage of the information on what consumers purchased is different than the one that sold the initial products to consumers. For example, information on what consumers purchase is sold to advertisers that may be interested in that information for better targeting. In the setting of the model of the previous version, if this information from the first period is sold by the first-period firm at price zero, this means that the firm in the first period only cares about the first-period profits, and the profits in the second period only accrue to the later firm. The equilibrium can then be computed in the same way by making $\delta_F = 0$. We focus the discussion in the uniform distribution case with quadratic $K()$ and linear $G()$.

If we have $\delta_C(1-p) > \delta_F p$, we do not have any change in the market behavior if the the profit in the first and second profit go to two different firms. The equilibrium still has one half of the market served in the first period, with the first period price determined by the consumer's discount factor, and the advertising expenditures determined by the second period optimal behavior.

If we have $\delta_C(1-p) < \delta_F p$, and if the firms selling in the first and second period are different, we then have that the first period price is now lower, because the first period firm does not benefit from the better targeting of advertising in the second period. Demand in the first period increases to one half. This then leads to a lower price in the second period (as the valuation of the marginal consumer is now lower), to lower advertising intensity to the consumers who purchased in the first period (as the consumers who purchased in the first period now have lower product valuations on average), and to increase the advertising intensity to the consumers who did not purchase in the first period. Overall, the total advertising increases, as now more consumers are purchasing the product in the first period.

If the information from the first period is sold at a positive price then the first period firm acts as if in the equilibrium computed above where δ_F now captures the present value of the extent of the gains from information of the second-period profit that the first-period firm can capture with its price for the first-period information. If the first-period firm has all the bargaining power in the sale of information, then we are back in the case considered above.

5. Behavioral-Based Pricing and Behavioral-Based Advertising

In some markets the firm may be able to engage in behavioral-based pricing in addition to behavioral-based advertising. We can investigate this case, focusing on the case of the uniform distribution, quadratic $K()$ linear $G()$, and γ small, for sharper results. In the second period, in addition to choosing advertising to both consumers who purchased in the first period and who did not purchase, the firm can now set a different price for consumers who purchased in the previous period and consumers who did not purchase as well. Let P_p be the price charged in the second period to the consumers who purchased in the first period (previous customers), and let P_n be the price charged in the second period to consumers who did not purchase in the first period (new customers).

Assuming that $P_p \geq \theta_1^*$ (which holds in equilibrium), the second period profit can then be written as

$$\pi_2 = (\beta+\phi)p[1-F(P_p)]P_p + (\beta+\phi)(1-p)[1-F(\theta_1^*)][1-F(P_p)]P_p + \\ \beta F(\theta_1^*)(1-p)[1-F(P_n)]P_n + \beta p[F(\theta_1^*) - F(P_n)]P_n - F(\theta_1^*)K(\beta) - [1-F(\theta_1^*)]K(\beta+\phi). \quad (16)$$

From this we can obtain the first order conditions for the advertising expenditures,

$$[1 - F(P_p) + pF(P_p)]P_p = K'(\beta + \phi), \tag{17}$$

which is the same condition for $\beta + \phi$ as in the model of the previous section for $P_2 = \theta_1^*$, and

$$[1 - F(P_n) - p\frac{F(P_n)}{F(\theta_1^*)}]P_n = K'(\beta) \tag{18}$$

which shows that advertising to the consumers who did not purchase in the first period is now greater, as the firm is now also targeting a low price to the consumers who did not purchase in the first period. For $p \to 1$, when there is no behavioral-based pricing, the firm would choose zero advertising to the consumers who did not purchase in the first period.

PROPOSITION 8: *When there is both behavioral-based advertising and behavioral-based pricing, the firm chooses to advertise more to the consumers who did not purchase in the first period than when there is behavioral-based advertising and there is no behavioral-based pricing.*

The possibility of behavioral-based pricing allows the firm to try to attract the consumers who did not purchase in the first period with a lower price, and therefore, it becomes more likely that those consumers choose to buy the product if exposed to advertising on it. It becomes then more profitable to advertise to those consumers that did not purchase in the first period. As expected, the intensity of advertising to the consumers who did not purchase in the first period is decreasing in the probability of the consumers maintaining the same preferences across periods, p.

For $\theta_1^* \geq \theta^*$ (which occurs in equilibrium), the firm chooses $P_p^* = \theta_1^*$. The price for the new customers solves the first order condition, $F(\theta_1^*) - [F(P_n) + f(P_n)P_n][F(\theta_1^*)(1-p) + p] = 0$, which for the uniform distribution case results in $P_n^* = \frac{\theta_1^*}{2[\theta_1^*(1-p)+p}$. Given the quadratic $K()$ this also leads to $\beta + \phi = \frac{p\theta_1^* + (1-p)(1-\theta_1^*)\theta_1^*}{\alpha}$ and $\beta = \frac{\theta_1^*}{4\alpha[\theta_1^*(1-p)+p]}$. Substituting for these in the second period profit one obtains

$$\pi_2^*(\theta_1^*) = \frac{\theta_1^{*2}(1-\theta_1^*)}{2\alpha}[1 - \theta_1^* + p\theta_1^*]^2 + \frac{\theta_1^{*3}}{32\alpha[\theta_1^*(1-p)+p]^2}, \tag{19}$$

from which one can obtain that the profits in the second period are increasing in p.

Consider now the marginal consumer who purchases in the first period. If that consumer purchases the product, that consumer gets an expected payoff of $\theta_1^* - P_1 + \delta_C(\beta+\phi)\frac{1-p}{2}(1-P_p)^2 - \delta_C\gamma(\beta+\phi)$ (as $P_p^* = \theta_1^*$). If that consumer decides not to purchase in the first period, that consumer would get an expected payoff of $\delta_C\beta[p(\theta_1^* - P_n^*) + \frac{1-p}{2}(1-P_n)^2] - \delta_C\gamma\beta$. The marginal consumer is then obtained by making equal these two terms. That expression is presented for general p in equation (xii) in the Appendix. Assuming that $p = 1$ for a less tedious presentation, the marginal consumer can be obtained to be determined by $\theta_1^* - \delta_C \frac{\theta_1^{*2}}{8\alpha} - \delta_C \gamma \frac{3\theta_1^*}{4\alpha} = P_1$. Using this we can write the present value of profits as a function of θ_1^* as

$$\pi_1 + \delta_F \pi_2 = \theta_1^*(1 - \theta_1^*)(1 - \delta_C \frac{\theta_1^*}{8\alpha} - \delta_C \frac{3\gamma}{4\alpha}) + \delta_F \frac{\theta_1^{*2}}{32\alpha}(16 - 15\theta_1^*). \tag{20}$$

Comparing the optimal θ_1^* in this case with the optimal θ_1^* when there is no behavioral-based pricing we can obtain that θ_1^* is now higher. When there is behavioral-based pricing in addition to behavioral-based advertising the firm chooses to sell to fewer consumers in the first period, as it can sell to the other consumers with a lower price in the second period. This then leads to lower demand in the first period and greater demand in the second period.

Proposition 9: *Suppose that there is behavioral-based advertising and the firm gains the ability to do behavioral-based pricing as well. Then, demand decreases in the first period and increases in the second period.*

With behavioral-based pricing, consumers in the first period know that if they refrain from buying the product, they will be seen as low valuation consumers in the second period, and therefore, receive a better deal in the second period. Then, the consumers with a valuation just above the first-period price do indeed refrain from purchasing in the first period, which results in less overall demand in the first period and greater demand in the second period. This effect of behavioral-based pricing on reduced demand in the first period because of strategic consumers can be seen, for example, in Villas-Boas (2004). Here this effect interacts with the effect of advertising by the firm advertising to the consumers who did not purchase in the first period (but less so than advertising to the consumers who purchased the product in the first period, and who are more profitable because of their higher valuation).

Figure 3 presents the equilibrium as a function of p for the both behavioral-based advertising and pricing case, and compares it with the case of just behavioral-based advertising.

For completeness, it is also interesting to compare these two cases with what happens when there is only behavioral-based pricing, and all consumers are aware of the second period product without the need for advertising. With no need for advertising the second period profits can be obtained to be $(1 - \theta_1^*)\theta_1^* + \frac{\theta_1^{*2}}{4}$, and the present value of profits can be obtained to be

$$(1 - \theta_1^*)\theta_1^*(1 - \frac{\delta_C}{2}) + \delta_F \theta_1^*(1 - \frac{3\theta_1^*}{4}).$$

Comparing this case with the case of behavioral-based advertising alone, we can then obtain that this case leads to a greater θ_1^*, lower demand in the first period, than in the case with behavioral-based advertising alone. Putting this together with the results above we can obtain the following result.

PROPOSITION 10: *With sufficiently patient firm and consumers, behavioral-based advertising can be beneficial to both firm and consumers, while this is not the case for behavioral-based pricing.*

With behavioral-based advertising the firm and consumers can benefit from better matching of advertising without hurting too much the demand in the initial period. With behavioral-based pricing demand is reduced in the initial period, which affects negatively the overall welfare.

Comparing the case of just behavioral-based pricing with the case of both behavioral-based pricing and advertising we can obtain that if advertising is very costly (high α) this leads to a higher θ_1^* than when there is both behavioral-based pricing and advertising. This is because in the advertising case, greater weight is given to the first period, which leads to θ_1^* being closer to θ^*. However, when α is small, and the consumers are myopic, we can obtain that θ_1^* is greater for the case of both behavioral-based pricing and advertising than just behavioral-based pricing (and consumers are fully aware of both products). The reason is that with both behavioral-based pricing and advertising there is an extra benefit of having a high θ_1^* of better targeting of advertising to the higher valuation consumers.

6. Concluding Remarks

This paper considers the effect of firms being able to advertise differently to consumers depending on their purchase history. The paper identifies two effects on consumer behavior: On the one hand consumers realize that by purchasing a product they may receive advertising about products that may generate positive surplus in the future. On the other hand, receiving advertising may cause inconvenience to the consumers. If the former effect dominates, more consumers end up buying the product in the initial periods (some with a valuation lower than the price charged), and the firm ends up increasing its price. As more consumers buy the product, the average valuation is lower, and the firm chooses a lower intensity of the advertising targeted at the consumers who have bought the initial product. If the latter effect dominates, then the firm has to lower the price of its initial product, fewer consumers purchase that initial product, and the firm ends up advertising more intensively to the consumers who purchased the product in the initial period.

Overall, both consumers and firms can benefit from behavioral-based advertising if the dis-utility of advertising is not too high, as the consumers that are most likely to value the products end up receiving advertising about them.

The paper also presents comparative statics of the market equilibrium with respect to the patience of consumers and firm, the degree of correlation of preferences of consumers through time, the dis-utility of advertising, and the cost of advertising.

The paper also discusses what happens when the advertising firm is different than the firm that sells the product in the initial periods. In that case, the initial period demand in equilibrium will never be lower, and can be greater, than in the case when we have the same firm across periods. Having behavioral-based pricing in addition to behavioral-based advertising leads to less sales in the initial period, as some consumers refrain from buying given the potential lower prices to the non-purchasing consumers in the later periods.

In future research it would be interesting to also consider variations on the modeling of advertising as giving information about product attributes rather than about overall awareness of the product. For future research it would also be interesting to consider what happens when behavioral-based advertising can also be based on past search behavior.

APPENDIX

Example of Information Obtained from Past Purchases versus from Subscription to Mailing List:

We present a brief illustrative example how information obtained from past purchases can be different than information obtained from a consumer signing up to a mailing list of information about the company's products. Consider the model described in the text, but suppose that the second-period price P_2 and advertising intensities β and ϕ are fixed and independent of the marginal consumer purchasing the product or subscribing to the mailing list in the first period. In the subscribing to the mailing list case, the firm advertises with intensity β to the consumers who did not subscribe to the mailing list and with intensity $\beta + \phi$ for the consumers who subscribed to the mailing list (which are the advertising intensities based on no-purchase and previous purchase, respectively, for the case when advertising is conditional on purchase history).

Consider first the case when the the consumer information is obtained from the purchase history. Then from (8) we can obtain that the consumers who buy in the first period satisfy

$$\theta_1 > P_1 - \delta_C \phi \{p \max[\theta_1 - P_2, 0] + (1-p)y\} + \delta_C[G(\beta+\phi) - G(\beta)] \tag{i}$$

where y was defined in Section 4.1 as $y \equiv \int_{P_2}^{\overline{\theta}} (\theta_2 - P_2)\, dF(\theta_2)$.

Consider now the decision of a consumer subscribing to the mailing list. Suppose that if the consumer subscribes to the mailing list he pays a transaction cost t, which may also include any privacy concerns beyond what would be revealed by just purchasing the product. Note also that the purchase of the product in the first period does not affect the relative payoff of subscribing to the mailing list as in this case, purchasing in the first period does not affect the advertising intensities that the consumer will face in the future. Then, if the consumer subscribes to the mailing list he gets an expected payoff equal to

$$-t + \delta_C(\beta+\phi) p \max[\theta_1 - P_2, 0] + \delta_C(\beta+\phi)(1-p)y - \delta_C G(\beta+\phi).$$

If the consumer decides not to subscribe to the mailing list the consumer gets an expected payoff of

$$\delta_C \beta [p\ \max[\theta_1 - P_2, 0] + (1-p)y] - \delta_C G(\beta). \tag{ii}$$

The consumer decides to subscribe to the mailing list if

$$t < \delta_C \phi\{p \max[\theta_1 - P_2, 0] + (1-p)y\} - \delta_C[G(\beta + \phi) - G(\beta)]. \tag{iii}$$

From this we can get that if information is obtained from the purchase history, consumers who purchase in the first period have a valuation above a certain threshold, while if information is obtained from subscribing to the mailing list, consumers who subscribe to mailing lists have low transaction costs in comparison to their valuation for the product.

That is, when getting information from the subscription to the mailing list (in comparison with purchase histories), we gain consumers that have low transactions costs and potentially low valuations, but we lose consumers with high transaction costs but with valuations that are above the threshold for purchase, but not too high. This then illustrates that the information provided by the purchase history and subscription to the mailing list is distinct. Note then that this illustrates the different forces at work when comparing information from purchase history versus information from subscription to the mailing list, but it is not a complete analysis of this comparison. Such complete analysis would have to endogeneize the effect of the marginal consumers on the price and advertising intensities in the second period. This analysis for the subscription to mailing lists case is beyond the scope of this paper. This analysis for the purchase history case is presented in the paper.

Description of Survey with respect to Personalized Advertising:

To further understand the phenomena discussed in the paper, we conducted a brief survey in the context of online shopping and behavioral-based advertising, among students in a large university in China. We received answers from 154 respondents, 64.8% of whom shop online at least once a week and 94.4% shop at least once a month. Among the respondents, 76% are graduate students and 24% are undergraduate students. Gender distribution is balanced with 47% males.

The survey mainly constituted of two parts. In one part, we asked the respondents to indicate their degree of agreement with various statements related to information acquisition and behavioral-based advertising on the 1-7 Likert scale (1 means strongly disagree and 7 means strongly agree). The main results can be summarized as follows. The distribution of the ratings scores to the different statements is displayed in Figures A1-A5.

First, most respondents are aware of the practice of behavioral-based advertising: 46% of the respondents agree or strongly agree (18% rated 6 and 28% rated 7) with the statement that "When I purchase something online I realize that this information will be used by the company to send me personalized ads." Less than 20% respondents are unaware of the practice (gave rating 1-3). This result is consistent with behavioral-based advertising being seen as a commonly used practice and that the majority of the consumers in this sample are expecting the use of their purchase information by merchants to some extent.

Second, while signing up for company emails can be an alternative way of getting information about products of potential interest, 38% of the respondents strongly disagree (rated 1) with the notion that "I subscribe to companies' email lists to receive product information." Only 8.5% of the respondents agree, to any extent, with this statement (rated 5-7).

Third, behavioral-based advertising can be interpreted as valuable to consumers: 56.5% of the respondents agree to the some extent that "the products advertised in personalized recommendations are more likely to fit my needs than the random product ads." Similarly, 60% of the respondents tend to agree that "personalized ads can provide information on products that I may like but am not aware of." A proportion of 50% of the respondents agree, to some extent, with the general notion that "I find personalized product recommendations useful."

In another part of the survey, we use conjoint design to solicit respondents' underlying preferences for the possibility of receiving personalized advertising based on their past purchase. We ask respondents to rate on a scale from 1 to 100 their tendency to shop at a given fashion web site with varying attributes. The web site profiles differ in terms of product variety, price, whether it is recommended by friends, and whether a customer will receive personalized email ads in the future based on purchase history. Each attribute has two levels and our key interest is the attribute of whether the web site offers personalized ads. We use orthogonal design and 8 profiles are generated.

Analysis of the individual responses reveals that the part-worth utility of receiving personalized email ads is positive for 75% of the respondents. In other words, for three quarters of the respondents, the possibility of receiving personalized ads of products they might be interested in the future increases their tendency to shop at a given online store. From the aggregate analysis, the attribute of receiving personalized ads is positive and significant, and claims 15% weight in respondents' decision to favor a store. While the weight needs to be interpreted with caution as it can be affected by the number of attributes included in the study, the feature of personalized ads based on one's past purchase may offer positive utility to a significant proportion of the respondents.

In summary, the survey results provide some suggestive evidence that behavioral-based advertising seems to be a commonly used business practice in online retailing, and that, at least some consumers are aware of it. In addition, behavioral-based advertising and personal recommendations may be valuable to consumers. This is a brief survey, and the results should be interpreted with care. A full-fledged survey to study these questions is beyond the scope of this paper.

Proof of Proposition 1: For the case where $P_2^* > \theta_1^*$, totally differentiating (3) one can get

$$\frac{\partial \phi^*}{\partial \theta_1^*} = p \frac{[1-F(P_2^*)]P_2^*}{K''(\beta^*+\phi^*)} \frac{f(\theta_1^*)}{[1-F(\theta_1^*)]^2} > 0. \tag{iv}$$

where ϕ^* represents the optimal value of ϕ, and β^* the optimal β. Note also that for this case the base advertising level β^* is independent of how many consumers bought the product in the first period (independent of θ_1^*).

For the case where $P_2^* = \theta_1^*$, one can obtain $\frac{\partial(\beta^*+\phi^*)}{\partial \theta_1^*} > 0$, and

$$\frac{\partial \phi^*}{\partial \theta_1^*} = \frac{p}{K''(\beta^*)} + [\frac{1}{K''(\beta^*+\phi^*)} - \frac{1}{K''(\beta^*)}][1-(1-p)F(\theta_1^*) - (1-p)f(\theta_1^*)\theta_1^*] \tag{v}$$

which is positive if the function $K()$ is sufficiently close to a quadratic function. Note that for this case the optimal base advertising level β^* increases in how many consumers bought the product in the first period.

Proof of Proposition 3: By totally differentiating (8) one can obtain

$$\frac{\partial \theta_1^*}{\partial P_1} = \frac{1}{1+\delta_C[(1-p)y - G'(\beta^*+\phi^*)]\frac{\partial \phi^*}{\partial \theta_1^*}} \tag{vi}$$

where $\frac{\partial \phi^*}{\partial \theta_1^*}$ is presented in (iv). From this one can immediately obtain the results in the proposition.

Proof of Proposition 4: When $P_2^* = \theta_1^*$ we can obtain

$$\frac{\partial \theta_1^*}{\partial P_1} = \frac{1}{1+\delta_C(1-p)[y\frac{\partial \phi^*}{\partial \theta_1^*} - \phi^*[1-F(\theta_1^*)]] - \delta_C[G'(\beta^*+\phi^*)(\frac{\partial \beta^*}{\partial \theta_1^*} + \frac{\partial \phi^*}{\partial \theta_1^*}) - G'(\beta^*)\frac{\partial \beta^*}{\partial \theta_1^*}]}, \tag{vii}$$

from which we can immediately obtain the results in the proposition.

Proof of Proposition 5: The proof of the result for a general distribution function is presented in the text. For the uniform distribution example with quadratic $K()$ and linear $G()$, we can re-write (11) to obtain

$$1-\theta_1^* - \frac{4(1-\theta_1^*)^2}{4(1-\theta_1^*)^2 + \frac{\delta_C p}{8\alpha}(1-p-8\gamma)}\{\theta_1^* + \frac{p}{32\alpha(1-\theta_1^*)}[\delta_C(1-p-8\gamma) - \delta_F\frac{p}{1-\theta_1^*}]\} = 0. \tag{viii}$$

Checking the left hand side of (viii) at $\theta_1^* = \frac{1}{2}$ one can obtain that the sign of the left hand side is equal to the sign of $\frac{\delta_F p^2}{16\alpha} > 0$. That means that the optimal θ_1^* is strictly greater than $\frac{1}{2}$ which violates the condition that $P_2^* > \theta_1^*$.

First order condition in first period for uniform distribution example with quadratic $K()$ and linear $G()$ for $P_2^* = \theta_1^*$: Substituting for this example in equation (12), we obtain

$$1 - 2\theta_1^* + \delta_C(1 - 2\theta_1^*)\frac{p}{\alpha}[(1-p)\frac{(1-\theta_1^*)^2}{2} - \gamma] - \delta_C\frac{p}{\alpha}\theta_1^*(1-\theta_1^*)^2(1-p) +$$
$$\delta_F\frac{1-\theta_1^*}{\alpha}[(1-p)\theta_1^*(1-\theta_1^*) + p\theta_1^* - p(1-p)\theta_1^{*2}] - \delta_F\frac{\theta_1^{*2}}{\alpha}[(1-\theta_1^*)(1-p^2) + \frac{p^2}{2}] = 0. \qquad \text{(ix)}$$

Proof of Proposition 7: In order to see the effect of the different parameters on the demand in the first period we can see the effect of the different parameters on θ_1^*. To do that we can differentiate (ix) with respect to each of the parameters. For δ_C, differentiating the left hand side of (ix) with respect to δ_C one obtains

$$(1-2\theta_1^*)\frac{p}{\alpha}[(1-p)\frac{(1-\theta_1^*)^2}{2} - \gamma] - \frac{p}{\alpha}\theta_1^*(1-p)(1-\theta_1)^2$$

which is negative given the condition on the proposition on γ and given that $\theta_1 * > \frac{1}{2}$, which yields $\frac{\partial\theta_1^*}{\partial\delta_C} < 0$. Differentiating the left hand side of (ix) with respect to δ_F one can also obtain that $\frac{\partial\theta_1^*}{\partial\delta_F} > 0$.

For the case of p, differentiating the left hand side of (ix) with respect to p, and then making $p \to 1, \delta_F - \delta_C \to 0$, and $\gamma \to 0$, one obtains $-(1-\theta_1^*)^2\frac{1-3\theta_1^*}{\alpha} + \theta_1^*(2 - \theta_1^* - 2\theta_1^{*2} > 0$, which yields $\frac{\partial\theta_1^*}{\partial p} > 0$. Similarly, we can also obtain $\frac{\partial\theta_1^*}{\partial\gamma} > 0$ and $\frac{\partial\theta_1^*}{\partial\alpha} < 0$.

Consider now the effect of the different parameters on P_1. Using (8) we can obtain for the conditions of the proposition, $P_1 = \theta_1^* + \delta_C p(1-p)\frac{\theta_1^*(1-\theta_1^*)^2}{2\alpha} - \delta_C\gamma\theta_1^*\frac{p}{\alpha}$. For the cases of δ_C and δ_F, using $p \to 1$ and $\gamma \to 0$, we can obtain $\frac{\partial P_1}{\partial\delta_C} = \frac{\partial\theta_1^*}{\partial\delta_C} < 0$, and $\frac{\partial P_1}{\partial\delta_F} = \frac{\partial\theta_1^*}{\partial\delta_F} > 0$.

Under these conditions for the case of p, one can obtain $\frac{\partial P_1}{\partial p} = \frac{\partial\theta_1^*}{\partial p} - \delta_C\frac{\theta_1^*(1-\theta_1^*)^2}{2\alpha}$. Noting that the second derivative of (10) is $-2 + \frac{\delta_F}{\alpha}(1 - 3\theta_1^*)$, when $p \to 1$, and $\gamma \to 0$, we can obtain that the sign of $\frac{\partial P_1}{\partial p}$ is equal to the sign of

$$(3\theta_1^* - 1)(1-\theta_1^*)^2 + \theta_1^*(2 - \theta_1^* - 2\theta_1^{*2}) - [2 + \frac{\delta_F}{\alpha}(3\theta_1^* - 1)]\delta_C\frac{\theta_1^*(1-\theta_1^*)^2}{2\alpha}. \qquad \text{(x)}$$

For the interior value of advertising expenditures, $\phi^* < 1$, we need $\alpha > \theta_1^*$. Making $\delta_F - \delta_C \to 1$, and checking α at its lowest possible value θ_1^* we can then obtain that (x) is positive, which means that $\frac{\partial P_1}{\partial p} > 0$.

Consider now the effect of γ on P_1. For $p \to 1$ and $\gamma \to 0$ we can obtain the sign of $\frac{\partial P_1}{\partial \gamma}$ as equal to the sign of $(3\theta_1^* - 1)(1-\theta_1^*)^2 + \theta_1^*(2 - \theta_1^* - 2\theta_1^{*2}) - (2 + \frac{\delta_F}{\gamma}(3\theta_1^* - 1))\frac{\theta_1^*}{\alpha}$, which is greater than zero for α large, and is less than zero for α small (close to θ_1^*). Finally, for $p \to 1$ and $\gamma \to 0$, we can obtain $\frac{\partial P_1}{\partial \alpha} = \frac{\partial P_1}{\partial \alpha} < 0$.

For the effects on advertising expenditures, we can immediately obtain for $p \to 1$ and $\gamma \to 0$ that $\frac{\partial \beta}{\partial p} = -\frac{\theta_1^*(1-\theta_1^*)}{\alpha} < 0$, $\frac{\partial \beta}{\partial \alpha} < 0$, $\frac{\partial \beta}{\partial \delta_C} = \frac{(1-p)(1-2\theta_1^*)}{\alpha}\frac{\partial \theta_1^*}{\partial \delta_C} > 0$, $\frac{\partial \beta}{\partial \delta_F} = \frac{(1-p)(1-2\theta_1^*)}{\alpha}\frac{\partial \theta_1^*}{\partial \delta_F} < 0$, and $\frac{\partial \beta}{\partial \gamma} = \frac{(1-p)(1-2\theta_1^*)}{\alpha}\frac{\partial \theta_1^*}{\partial \gamma} < 0$. Similarly, we can obtain $\frac{\partial(\beta+\phi)}{\partial p} = \frac{\theta_1^{*2}}{\alpha} + \frac{p}{\alpha}\frac{\partial \theta_1^*}{\partial p} > 0$, $\frac{\partial(\beta+\phi)}{\partial \alpha} < 0$, $\frac{\partial(\beta+\phi)}{\partial \delta_C} = \frac{p}{\alpha}\frac{\partial \theta_1^*}{\partial \delta_C} < 0$, $\frac{\partial(\beta+\phi)}{\partial \delta_F} = \frac{p}{\alpha}\frac{\partial \theta_1^*}{\partial \delta_F} > 0$, and $\frac{\partial(\beta+\phi)}{\partial \gamma} = \frac{p}{\alpha}\frac{\partial \theta_1^*}{\partial \gamma} > 0$.

For the effect on the present value of profits, note that the present value of profits can be written as

$$\begin{aligned}\pi_1 + \delta_F \pi_2 = {} & (1-\theta_1)(\theta_1 + \delta_C p(1-p)\frac{\theta_1(1-\theta_1)^2}{2\alpha} - \delta_C \gamma \theta_1 \frac{p}{\alpha}) \\ & + \delta_F \frac{\theta_1^2(1-\theta_1)}{2\alpha}[[(1-p)(1-\theta_1) + p]^2 + \theta_1(1-\theta_1)(1-p)^2]. \end{aligned} \tag{xi}$$

Using the envelope theorem we know that to check the effect of each of the parameters on the present value of profits we just need to take the partial derivative of the present value of profits with respect to each parameter as $\frac{\partial(\pi_1 + \delta_F \pi_2)}{\partial \theta_1} = 0$, at the optimum. Using also the condition on γ small compared to $(1-p)$, $p \to 1, \gamma \to 0$, and $\delta_F - \delta_C \to 0$, we can obtain that the present value of profits is increasing in δ_C, δ_F, and p, and decreasing in γ and α.

Marginal Consumer for Both Behavioral-Based Advertising and Pricing:

For the case of both behavioral-based advertising and pricing, substituting for β^*, ϕ^*, and P_n^*, in the equality of the consumer surplus after purchasing in the first period with the consumer surplus of waithing for the second period, we can obtain:

$$\begin{aligned}\theta_1^* - P_1 + \delta_C \frac{\theta_1^*(1-\theta_1^*)}{2\alpha}(1-p)[1 - \theta_1^* + p\theta_1^*] = {} & \delta_C \frac{\theta_1^*}{32\alpha[\theta_1^*(1-p)+p]^3}[4p\theta_1^*[2[\theta_1^*(1-p) + \\ & p] - 1][\theta_1^*(1-p)+p] + (1-p)[2[\theta_1^*(1-p)+p] - \theta_1^*]^2] + \delta_C \gamma \frac{\theta_1^*}{\alpha}[1 - \theta_1^* + p\theta_1^* \\ & - \frac{1}{4[\theta_1^*(1-p)+p]}]. \end{aligned} \tag{xii}$$

REFERENCES

Armstrong, M., and J. Zhou (2015), “Search Deterrence," *Review of Economic Studies,* forthcoming.

Bergemann, D., and A. Bonatti (2011), “Targeting in Advertising Markets: Implications for Offline vs. Online Media," *RAND Journal of Economics,* **42**, 417-443.

Chen, Y. (1997), “Paying Customers to Switch," *Journal of Economics and Management Strategy,* **6**, 877-897.

Chen, Y., and J.A. Pearcy (2010), “Dynamic Pricing: When to Entice Brand Switching and When to Reward Consumer Loyalty," *RAND Journal of Economics,* **41**, 674-685.

Chen, Y., and Z.J. Zhang (2009), “Dynamic Targeted Pricing with Strategic Consumers," *International Journal of Industrial Organization,* **27**, 43-50.

Dias, M.B., D. Locher, M. Li, W. El-Deredy, and P.J. Lisboa (2008), “The Value of Personalized Recommender Systems to E-Business: a Case Study," *Proceedings of the 2008 ACM Conference on Recommender Systems,* RecSys ’08, New York, NY, USA, pp.291-294

Esteves, R.B. (2009), “Customer Poaching and Advertising," *Journal of Industrial Economics,* **57**, 112-146.

Esteves, R.B. (2010), “Pricing under Customer Recognition," *International Journal of Industrial Organization,* **28**, 669-681.

Fudenberg, D., and J. Tirole (2000), “Customer Poaching and Brand Switching," *RAND Journal of Economics,* **31**, 634-657.

Goldfarb, A., and C.E. Tucker (2011), “Privacy Regulation and Online Advertising," *Management Science,* **57**, 57-71.

Hauser, J.R., G.L. Urban, G. Liberali, and M. Braun (2009), “Website Morphing," *Marketing Science,* **28**, 202-223.

Iijima, R., and Y. Kamada (2016), “Social Distance and Network Structures," *Theoretical Economics,* forthcoming.

Iyer, G., D. Soberman, and J.M. Villas-Boas (2005), “The Targeting of Advertising," *Marketing Science,* **24**, 461-476.

Joe, H. (1997), *Multivariate Models and Dependence Concepts,* London: Chapman & Hall.

Lambrecht, A., and C.E. Tucker (2013), “When Does Retargeting Work? Information Specificity in Online Advertising," *Working Paper,* London Business School and MIT Sloan School of Management.

Mahanjan, V., and E. Muller (1986), “Advertising Pulsing Policies for Generating Awareness for New Products," , **5**, 89-106.

Malthouse, E., and R. Elsner (2006), "Customisation with Crossed-Basis Sub-Segmentation," *The Journal of Database Marketing & Customer Strategy Management,* **14**, 4-50.

Mayzlin, D., and J. Shin (2011), "Uninformative Advertising as an Invitation to Search," *Marketing Science,* **30**, 666-685.

Narasimhan, C. (1988), "Competitive Promotional Strategies," *Journal of Business,* **61**, 427-449.

Nelson, P. (1974), "Advertising as Information," *Journal of Political Economy,* **82**, 729-754.

Pazgal, A., and D. Soberman (2008), "Behavior-Based Discrimination: Is It a Winning Play and If So, When?," *Marketing Science,* **27**, 977-994.

Roy, S. (2000), "Strategic Segmentation of a Market," *International Journal of Industrial Organization,* **18**, 1279-1290.

Shin, J., and K. Sudhir (2010), "A Customer Management Dilemma: When Is It Profitable to Reward Existing Customers," *Marketing Science,* **29**, 671-689.

Stegeman, M. (1991), "Advertising in Competitive Markets," *American Economic Review,* **81**, 210-223.

Taylor, C.R. (2003), "Supplier Surfing: Price-Discrimination in Markets with Repeat Purchases," *RAND Journal of Economics,* **34**, 223-246.

Villas-Boas, J.M. (1999), "Dynamic Competition with Customer Recognition," *RAND Journal of Economics,* **30**, 604-631.

Villas-Boas, J.M. (2004), "Price Cycles in Markets with Customer Recognition," *RAND Journal of Economics,* **35**, 486-501.

Zhang, J. (2011), "The Perils of Behavior-Based Personalization," *Marketing Science,* **30**, 170-186.

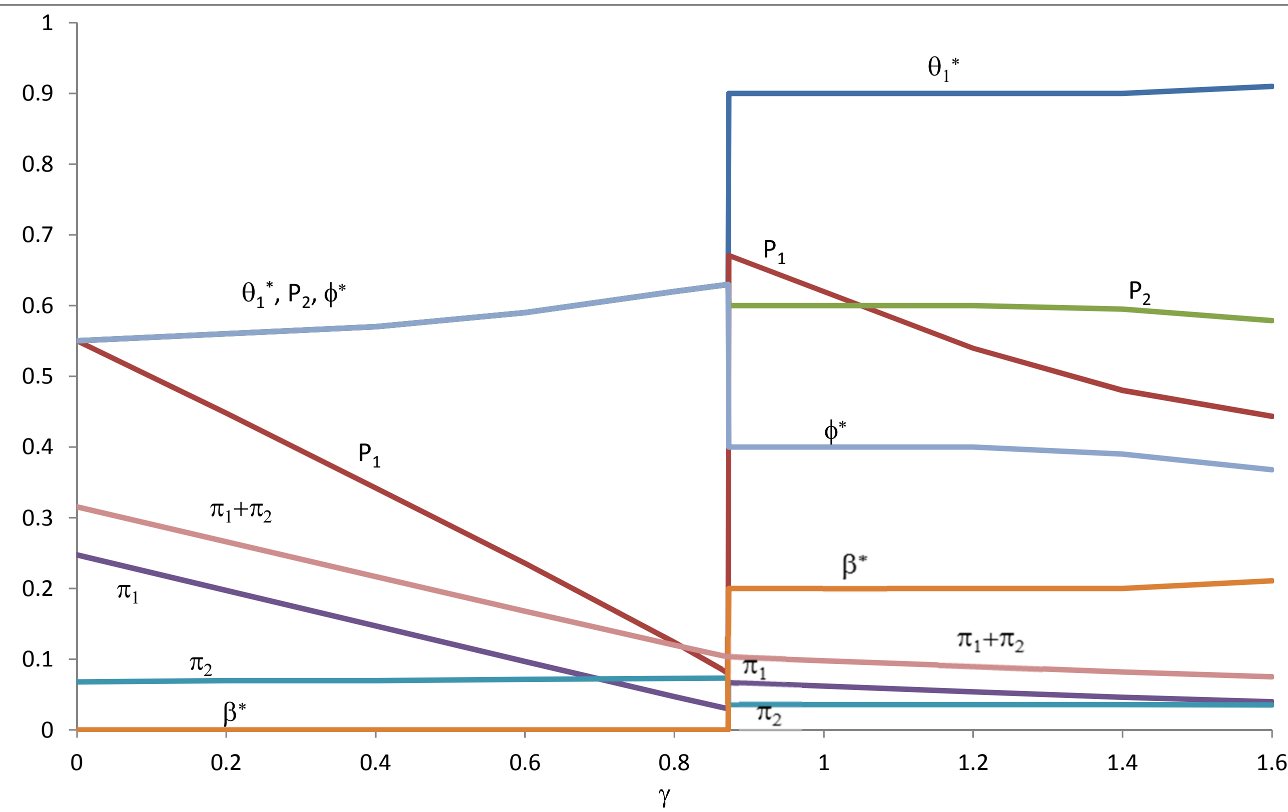

Figure 1: Evolution of the equilibrium θ_1^*, P_1, P_2, β^*, ϕ^*, π_1, π_2, and $\pi_1+\pi_2$ as a function of γ, for $p=\delta_C=\delta_F=\alpha=1$.

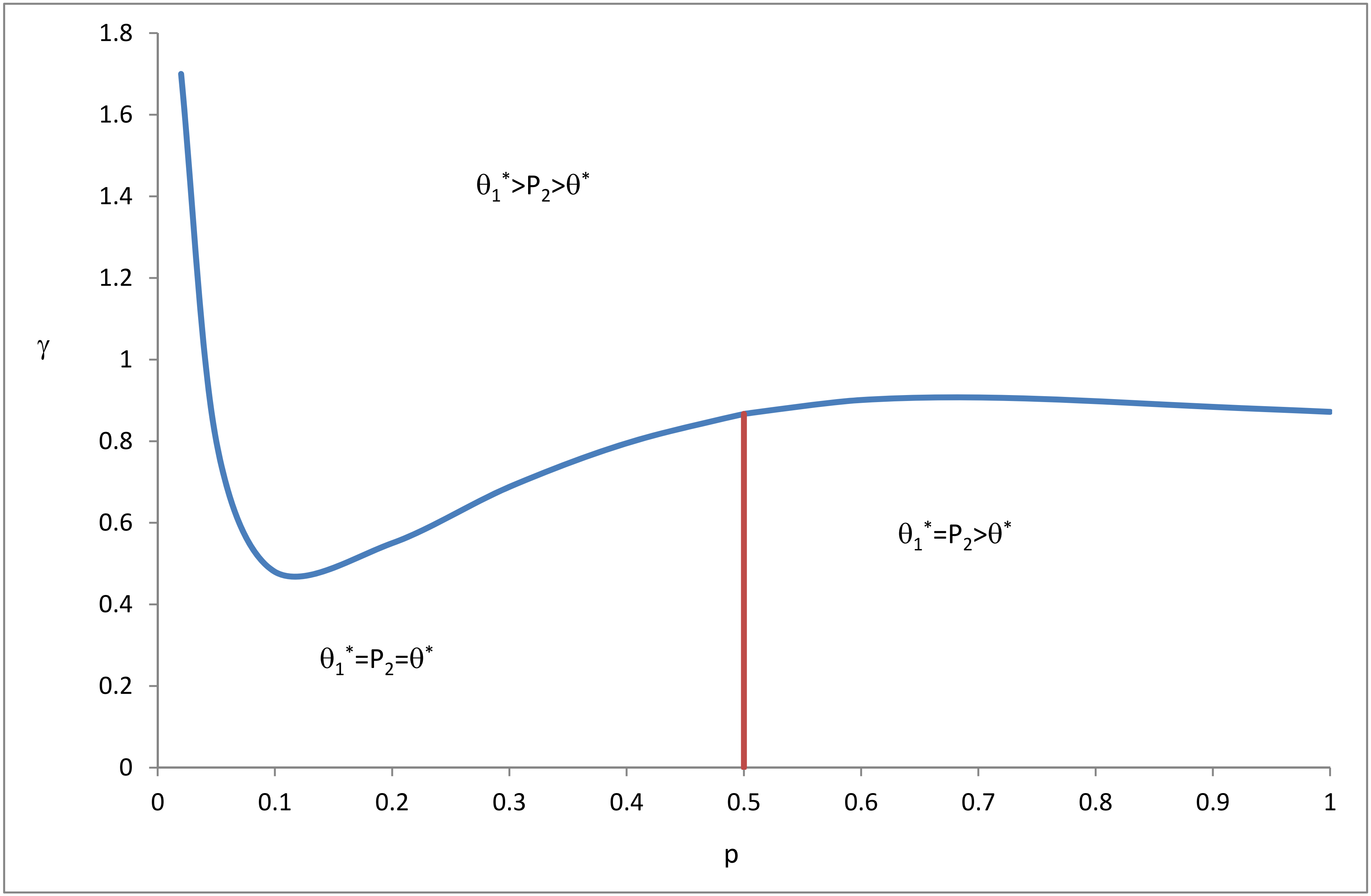

Figure 2: Areas of different equilibrium cases as a function of p and γ for $\alpha=\delta_C=\delta_F=1$.

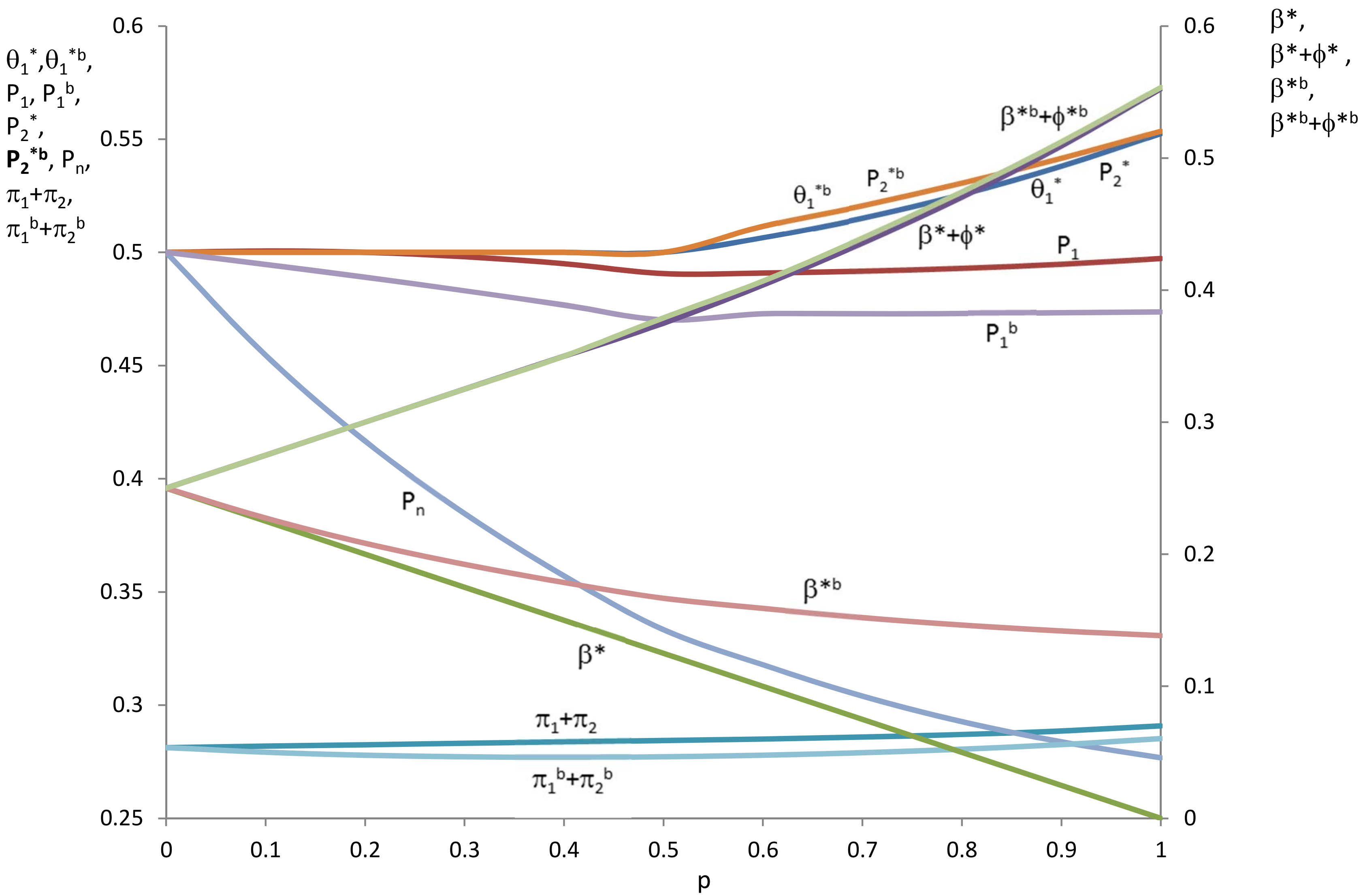

Figure 3: Comparison of equilibrium outcomes with and without behavioral-based pricing as a function of p for $\delta_C=\delta_F=\alpha=1$ and $\gamma=.1$. The superscript b means the case with behavioral-based pricing.

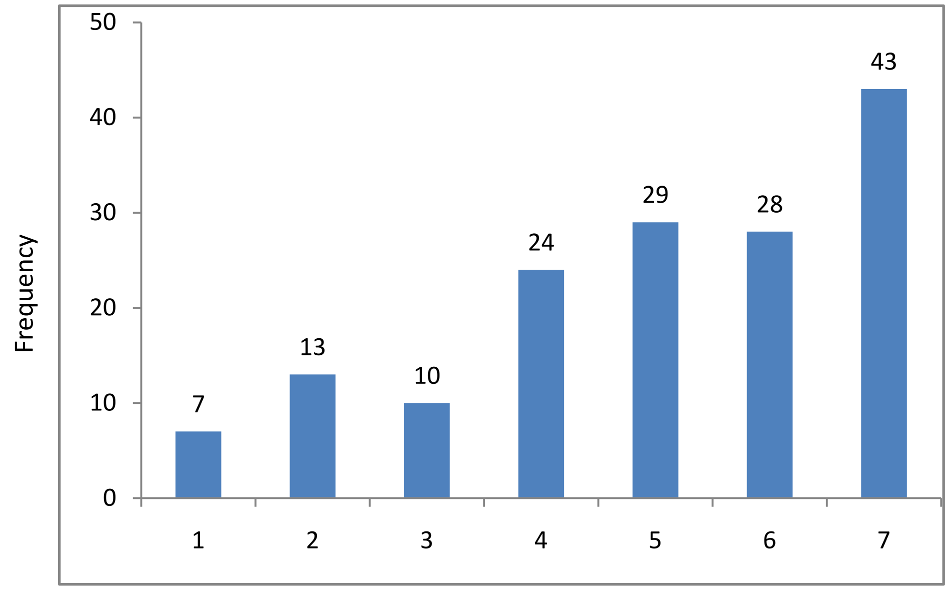

Figure A1: Rating distribution for question "When I purchase something online I realize that this information will be used by the company to send me personalized ads."

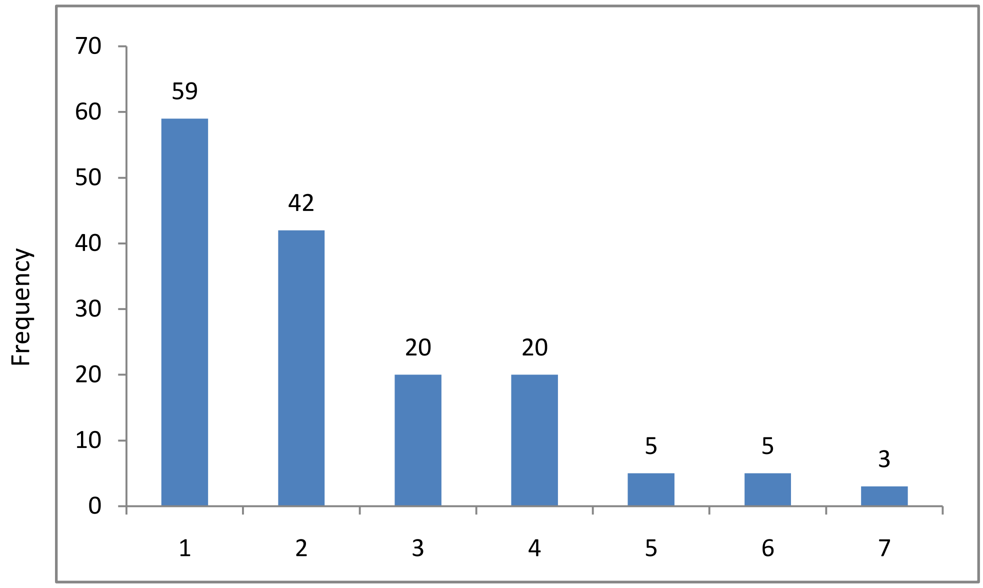

Figure A2: Rating distribution for question "I subscribe to companies' email lists to receive product information."

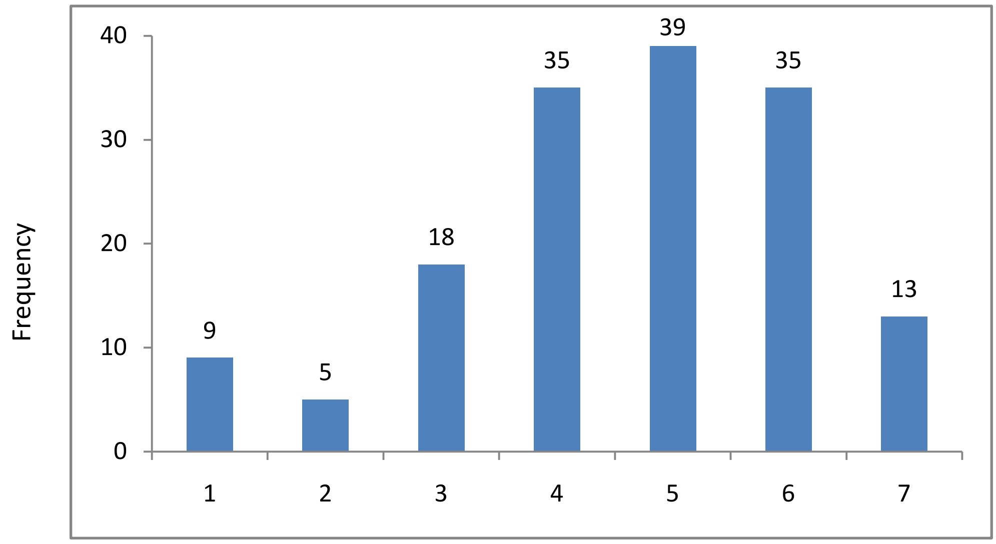

Figure A3: Rating distribution for question "The products advertised in personalized recommendations are more likely to fit my needs than random product ads."

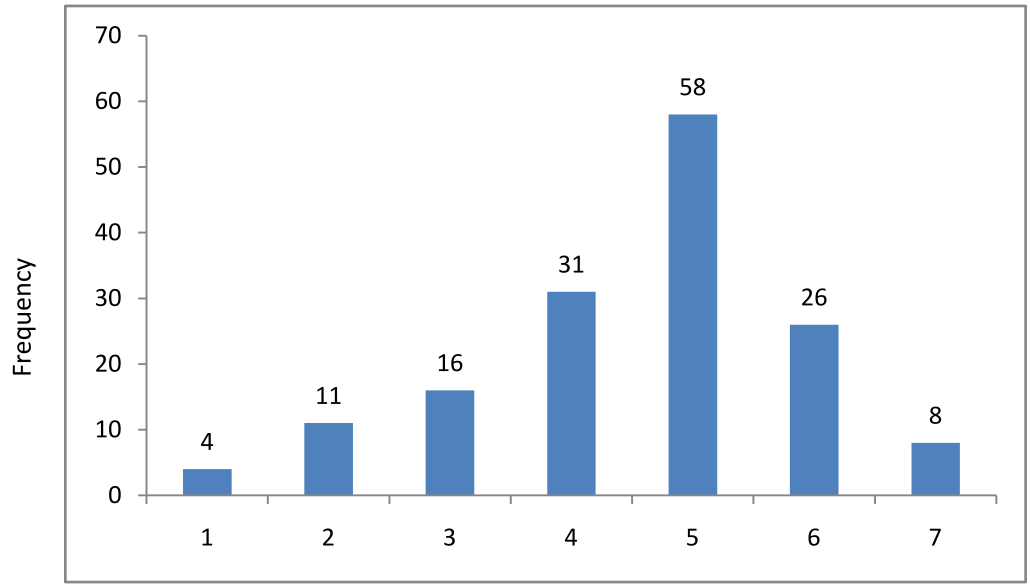

Figure A4: Rating distribution for question "personalized ads can provide information on products that I may like but am not aware of."

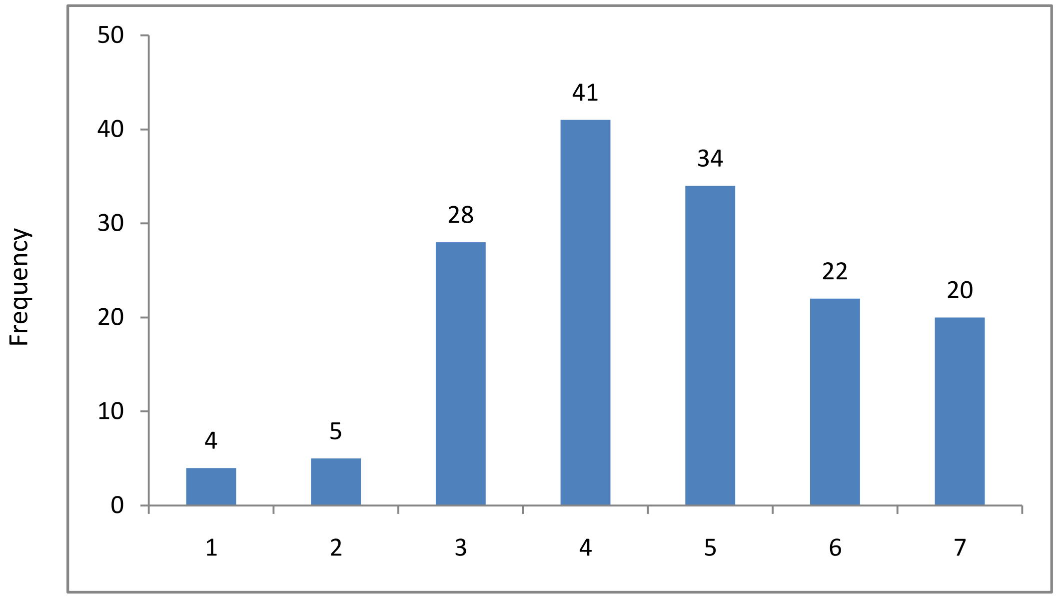

Figure A5: Rating distribution for question "I find personalized product recommendations useful."