Expertise, Attribution, and Ad Blocking in the World of Online Marketing

by
Stylianos Despotakis

A dissertation submitted to the Tepper School of Business in partial fulfillment of the requirements for the degree of Doctor of Philosophy

at

Carnegie Mellon University

May 2018

Dissertation Committee:
R. Ravi (Chair)
Kannan Srinivasan
Vibhanshu Abhishek
Isa Hafalir
Kaifu Zhang

Abstract

In this dissertation, we model and provide insights to some of the main challenges the world of online marketing currently faces. In the first chapter, we study the role of information asymmetry introduced by the presence of experts in online marketplaces and how it affects the strategic decisions of different parties in these markets. In the second chapter, we study the attribution problem in online advertising and examine optimal ways for advertisers to allocate their marketing budget across channels. In the third chapter, we explore the effects of modern ad blockers on users and online platforms.

In the first chapter, we examine the effect of the presence of expert buyers on other buyers, the platform, and the sellers in online markets. We model buyer expertise as the ability to accurately predict the quality, or condition, of an item, modeled as its common value. We show that nonexperts may bid more aggressively, even above their expected valuation, to compensate for their lack of information. As a consequence, we obtain two interesting implications. First, auctions with a "hard close" may generate higher revenue than those with a "soft close". Second, contrary to the linkage principle, an auction platform may obtain a higher revenue by hiding the item's common-value information from the buyers. We also consider markets where both auctions and posted prices are available and show that the presence of experts allows the sellers of high quality items to signal their quality by choosing to sell via auctions.

In the second chapter, we study the problem of attributing credit for customer acquisition to different components of a digital marketing campaign using an analytical model. We investigate attribution contracts through which an advertiser tries to incentivize two publishers that affect customer acquisition. We situate such contracts in a two-stage marketing funnel, where the publishers should coordinate their efforts to drive conversions. First, we analyze the popular class of multi-touch contracts where the principal splits the attribution among publishers using fixed weights depending on their position. Our first result shows the following counterintuitive property of optimal multi-touch contracts: higher credit is given to the portion of the funnel where the existing baseline conversion rate is higher. Next, we show that social welfare maximizing contracts can sometimes have even higher conversion rate than optimal multi-touch contracts, highlighting a prisoners' dilemma effect in the equilibrium for the multi-touch contract. While multi-touch attribution is not globally optimal, there are linear contracts that "coordinate the funnel" to achieve optimal revenue. However, such optimal-revenue contracts require knowledge of the baseline conversion rates by the principal. When this information is not available, we propose a new class of 'reinforcement' contracts and show that for a large range of model parameters these contracts yield better revenue than multi-touch.

In the third chapter, we study the effects of ad blockers in online advertising. While online advertising is the lifeline of many internet content platforms, the usage of ad blockers has surged in recent years presenting a challenge to platforms dependent on ad revenue. In this chapter, using a simple analytical model with two competing platforms, we show that the presence of ad blockers can actually benefit platforms. In particular, there are conditions under which the optimal equilibrium strategy for the platforms is to allow the use of ad blockers (rather than using an adblock wall, or charging a fee for viewing ad-free content). The key insight is that allowing ad blockers serves to differentiate platform users based on their disutility to viewing ads. This allows platforms to increase their ad intensity on those that do not use the ad blockers and achieve higher returns than in a world without ad blockers. We show robustness of these results when we allow a larger combination of platform strategies, as well as by explaining how ad whitelisting schemes offered by modern ad blockers can add value. Our study provides general guidelines for what strategy a platform should follow based on the heterogeneity in the ad sensitivity of their user base.

Keywords: competitive strategy; segmentation; pricing; sniping; online auctions; signaling; attribution; advertising; Shapley value; multi-touch; game theory; marketing funnel; ad blocking; ad sensitivity; ad intensity

Acknowledgements

I would like to start by expressing my deepest gratitude to R. Ravi for his immense help throughout my time at CMU. I am truly grateful for the invaluable advice and support that he generously offered to me all these years.

I would also like to thank Isa Hafalir, Amin Sayedi, Vibs Abhishek, and Kannan Srinivasan for their tremendous help and contributions to the projects of this dissertation. The discussions we had and their feedback motivated me to work harder and encouraged me to be more passionate about my work. I also thank Kaifu Zhang for being part of my dissertation committee and providing some insightful comments and suggestions.

I also want to express my appreciation to Lawrence Rapp and Laila Lee for taking care of all the administrative needs at Tepper and making life for PhD students easier.

Special thanks go to all my friends and fellow students at CMU. I learned a lot from our work-related discussions, while our non-work related activities helped in making this journey more fun and enjoyable.

Most of the things I know today are thanks to my many teachers throughout my life, before and during my studies at CMU. I will be forever indebted to them for everything they provided to me.

Last but not least, I would like to thank my family. Their constant and unconditional love and support has benefited me enormously and I feel truly blessed for having them in my life.

Contents

Chapter 1

Expertise in Online Markets[1]

In this chapter, we examine the effect of the presence of expert buyers on other buyers, the platform, and the sellers in online markets. We model buyer expertise as the ability to accurately predict the quality, or condition, of an item, modeled as its common value. We show that nonexperts may bid more aggressively, even above their expected valuation, to compensate for their lack of information. As a consequence, we obtain two interesting implications. First, auctions with a "hard close" may generate higher revenue than those with a "soft close". Second, contrary to the linkage principle, an auction platform may obtain a higher revenue by hiding the item's common-value information from the buyers. We also consider markets where both auctions and posted prices are available and show that the presence of experts allows the sellers of high quality items to signal their quality by choosing to sell via auctions.

1.1 Introduction

The advent of online auctions such as those in eBay led to the first massive-scale deployment of simple second-price auction mechanisms for consumer products. Even though eBay started as a platform for consumer-to-consumer auctions for selling items out of one's garage, it is now a large selling platform enabling over $200 billion commerce volume and reaching over 200 million users annually.[2] The addition of posted-price sales has fueled this growth by allowing it to serve as a competitor to other online retail sites. The growth of this new segment of online markets that combine auctions with posted prices raises important new questions about the optimal strategies for buyers and sellers as well as questions about the best design of the platform.

The eBay auction format enforces a "hard close" or ending time at which the item is sold to the highest (winning) bid. In the hours leading up to closing time, the auction is open and simulates the open outcry English auction. If all bidders had only private values, traditional auction theory dictates that the dominant strategy for every bidder is to bid up to his true value. To enable this, eBay offers a proxy bidding tool that allows a bidder to specify

[1]Based on joint work with Isa Hafalir, R. Ravi, and Amin Sayedi

[2]http://venturebeat.com/2013/10/16/ebay-earnings-sales-up-21-revenue-up-14-and-double-digit-paypal-user-growth/ (accessed January 2016)

his maximum value, and the tool automatically bids the minimum bid increment above the current highest bid (as long as it is below the bidder-specified value). Thus, it was something of a paradox when a majority of eBay auctions exhibited sniping – the phenomenon where a bidder submits his only bid in the last few seconds of the auction, thus avoiding any response from other bidders.

While several explanations for this behavior have been advanced, one of the most intuitive and accepted ones is that of experienced bidders (Wilcox, 2000) or dealers/experts (Roth and Ockenfels, 2002). For example, Roth and Ockenfels (2002, p. 1095) argue, and provide empirical evidence, that the existence of sniping in online markets is partly due to buyers' heterogeneity in their experience with online markets and their expertise in the product category: "[T]here may be bidders who are dealers/experts and who are better able to identify high-value antiques. These well-informed bidders...may wish to bid late because other bidders will recognize that their bid is a signal that the object is unusually valuable."

In this line of reasoning, the item auctioned off is assumed to have a common value which these experts have a better knowledge of, and submitting a sniping bid is a way for experts to withhold this information to reap the advantage of this information asymmetry in the resulting price. While several papers have subsequently built upon and refined this explanation of sniping (Bajari and Hortacsu, 2003; Rasmusen, 2006; Ockenfels and Roth, 2006; Hossain, 2008; Ely and Hossain, 2009), all of them have examined the phenomenon only from the bidders' perspective. More broadly, to best of our knowledge, no other paper has studied the strategic impact of buyers' heterogeneity in expertise (which causes the sniping behavior) on the platform and sellers' strategies in online markets. In this paper, we examine the effect of the existence of expert buyers on *all* of the stakeholders in online markets: the expert and nonexpert buyers, the sellers, and the platform. We discuss the following research questions:

1. How do nonexpert buyers adjust their strategies to compete with experts?
2. How does the presence of experts affect the platform revenue?
3. How does the presence of experts affect the sellers' strategies in online markets?

1.1.1 Our Contributions

First, we show that the presence of experts encourages the nonexperts to bid more aggressively. In particular, we show that because of the sniping strategy of the expert buyers in hard-close auctions, nonexpert buyers have to bid more than their expected value; otherwise they only win items of low quality against the expert buyers. Quantifying this, we show in Proposition 1.1 that the higher the proportion of experts among the bidders, the more aggressively the nonexperts bid above their expected value for the item.

Next, we consider the impact of the presence of experts on the platform's strategies. In particular, should the platform maintain the hard-close format for the auction, which allows the experts to snipe, rather than switch to the "soft-close" format? Also, if the platform knows the quality value of the item and can credibly reveal it to the buyers, should it commit to sharing this information with them? We find interesting answers to these questions. Regarding the first question, at the outset, it appears that the hard-close format may hurt platform revenue since without the sniping behavior of experts, nonexpert buyers could respond to bids of experts, and the item would sell at a higher price. Since the platform's fee is usually a fixed fraction of the selling price, the platform would then have an incentive to favor the soft-close format.[3] Contrary to this expectation, we show that the aggressive bidding behavior of the nonexperts that we describe above implies that the platform's overall revenue *increases* in the hard-close format for a wide range of parameter values (Proposition 1.2). This is a potential new explanation as to why online auction companies such as eBay[4] retain the hard-close auction format from a revenue perspective. We note, however, that the strategic choice of soft- versus hard-close format is a complex decision affected by competition among auction platforms as well as a variety of other bidder considerations such as the avoidance of potentially costly bidding wars in hard-close auctions. Our observation above exposes a new facet in a variety of such potential explanations for the popularity of this format.

This result has another important and interesting implication regarding the second question: the platform can benefit from committing to withholding the quality information (Corollary 1.1). This is in contrast to the celebrated *linkage principle*[5] (Milgrom and Weber, 1982), and is driven by buyers' heterogeneity in their level of expertise. Proposition 1.1 can also be interpreted as a reverse *winner's curse.* In auctions with common values, bidders bid lower than their valuation to avoid the winner's curse. However, our result shows that when bidders are heterogeneous in their level of information, non-informed bidders bid more than their valuation to make up for their lack of information.

Finally, we consider the impact of the presence of expert buyers on the sellers' strategies. In particular, we investigate the choice of selling mechanisms between the auction and a posted price sale when they are both available (as is common in most online auction-houses). In the presence of expert buyers, under certain conditions, we show that by selling in an auction, a seller can credibly signal[6] the quality value of his item (Proposition 1.3). By selling in an auction, the seller shows that he can rely on the *market* (specifically, on the expert buyers) to decide the value of the item. This is a risk that a seller with a low quality-value item cannot take. Furthermore, this signaling is possible only if there are enough experts, who know the value of the item, in the market. Otherwise, the seller of a high quality-value item will not be able to separate himself from the seller of a low-value item. In other words, the existence of experts in the market allows the sellers of high-quality products to separate themselves by selling in auctions. This finding is in line with auction houses' claim that auctions increase buyers' confidence. For example, Fraise Auction[7] argues that one of the benefits of selling in auction is that the "competitive bidding format creates confidence among the buyers when they see other people willing to pay a similar amount for the property." To best of our knowledge, this result is a new explanation for the popularity

[3] In fact, some auction platforms such as the now defunct Amazon Auctions and Trademe, removed sniping by implementing a soft close that automatically extended the auction time whenever a bid is submitted.

[4] EZsniper.com provides an extensive list of auction sites with a hard close.

[5] The linkage principle argues that the auction house always benefits from committing to revealing all available information.

[6] Note that the signal that we discuss here is the seller's choice of the selling mechanism. This is different from bids by other bidders, which can also be signals of the quality of the product.

[7] `http://fraiseauction.com/why-auction/` (accessed January 2016)

of auctions in certain product categories. We reiterate that the strategic choice of auction versus posted-price is a complex decision affected by several factors. Our observation above proposes a new explanation for why some sellers may choose to use auctions.

Taken together, we initiate the first comprehensive study of the effect of the presence of expert buyers in online markets featuring auctions with a hard close and posted prices, and establish the following results.

1. Nonexpert buyers must adjust their strategies in response to experts' sniping, and, under certain conditions, have to bid more than their expected value in hard-close auctions in equilibrium.

2. As a consequence, the platform revenue is higher in the hard-close auction than in the soft-close format for a wide range of parameter values.

3. Finally, the presence of experts in markets with hard-close auctions and posted prices allows the seller of high-quality items to credibly signal the quality of the item by selling in the auction and separating himself from sellers of low-quality items who sell using posted prices, under certain conditions.

Note that despite the explosive growth of auctions particularly in the consumer-to-consumer arena, our findings are relevant mainly to items with a significant common value component (such as collectibles, antiques, art, and used items of uncertain quality).

In what follows, we review related literature. Section 1.2 introduces the main model, Section 1.3 solves the equilibria of the model with a hard close, and Section 1.4 compares them with the corresponding equilibria of the auction with a soft close, which does not allow for sniping. In Section 1.5, we analyze the sellers' game of choosing among selling formats. We conclude the paper in Section 1.6. All proofs and further details are relegated to the Appendices.

1.1.2 Related Literature

Our work relates to the literature on online auctions with common values and a hard close, intermediaries' incentives to reveal product quality information, sellers' strategies to signal product quality, and the advantages and disadvantages of auctions versus posted prices. In the following, we review the related literature on each topic.

Bajari and Hortacsu (2003) argue that last-minute bidding is an equilibrium in a stylized model of eBay auctions with common values. They develop and estimate a structural econometric model of bidding in eBay auctions with common value and endogenous entry. Wilcox (2000) and Rasmusen (2006) use common values to model sniping and bidders' behavior on Ebay auctions. Wilcox (2000) shows that sniping increases as buyers' experience increases. Furthermore, the increase in the sniping behavior of the more experienced bidders is more pronounced for the type of items that are more likely to have a common value component. Similarly, a model with no common value as in Yoganarasimhan (2013) demonstrates no sniping behavior. Rasmusen (2006) considers a model where bidders incur a cost for learning the common value of the item. As a result, those who acquire the information snipe to hide their information from other bidders. Similar to the previous literature,[8] sniping emerges as an equilibrium strategy in our model as well. However, our focus is the effect of the presence of experts on nonexperts', sellers', and the platform's strategies and revenues, which is crucially missing in the earlier literature. Glover and Raviv (2012) show that when sellers can choose between hard-close and soft-close formats, soft close leads to a higher revenue, and experienced sellers are more likely to choose soft close. We discuss their result in Section 1.5, and show that soft close emerges as the unique pooling equilibrium if sellers can choose the closing format. Our result provides a new theoretical explanation for their empirical findings. In contrast to earlier work by Ockenfels and Roth (2006), who show an example in which seller revenue is lower at the equilibrium for hard-close than in the soft-close case, in our model, we show that the hard-close format increases revenue compared to the soft-close format. More specifically, we provide an explanation as to why online auction companies such as eBay retain the auction format that allows for sniping from a revenue perspective that takes into account the aggressive bidding behavior of the nonexperts.

In this paper, we show that an intermediary could benefit from withholding information about the quality of the items in an auction. This is in contrast with the well-known linkage principle by Milgrom and Weber (1982). The linkage principle argues that the auction house always benefits from committing to reveal all available information. The intuition behind the principle is that revealing the information can mitigate the winner's curse and motivates the buyers to bid more aggressively. We arrive at the contrast due to buyers' *heterogeneity* in terms of their information about the quality value of the item, as modeled by their expert status. More specifically, the result of Milgrom and Weber (1982) is established when valuation of bidders depend *symmetrically* on the unobserved signals of the other bidders, a condition that is not satisfied in our setup.[9] Withholding information, under certain circumstances, has also been shown to increase social welfare, by Zhang (2013), in the context of product labeling. Gal-Or et al. (2007) show that, under certain conditions, a buyer benefits from withholding information in procurement schemes.

Many researchers in marketing have studied signaling unobserved quality under information asymmetry. Moorthy and Srinivasan (1995) and Soberman (2003) show that sellers can use warranties such as money-back guarantees to signal the quality of their items. Bhardwaj and Balasubramanian (2005) show that by letting the customers request information about an item, rather than revealing it without solicitation, a seller can signal the quality of his item. Mayzlin and Shin (2011) show that uninformative advertising, as an invitation for search, can be used to signal product quality. Li et al. (2009) investigate auction features such as pictures and reserve price that enable sellers to reveal more information about their credibility and product quality, and empirically examine how different types of indicators help alleviate uncertainty. Finally, Subramanian and Rao (2016) show that, by displaying daily deal sales, a platform can leverage its sales to experienced customers to signal its type and attract new customers. This is relevant to our result as in both Subramanian and Rao's paper and our paper, the existence of experts (or experienced customers) can help the sell-

[8] The literature on trying to explain sniping in online auctions is vast. Other than previously mentioned papers, see also Ockenfels and Roth (2006), Hossain (2008), Wintr (2008), and Ely and Hossain (2009).

[9] Failure of the linkage principle has also been argued in a few other papers in the auction theory literature. For example, Perry and Reny (1999), Krishna (2009, Chapter 8.1), and Fang and Parreiras (2003) show the failure in setups with multiple items, ex-ante asymmetries, and budget constraints, respectively.

ers to extract more revenue from the nonexpert customers. However, the higher revenue is achieved using very different tools, displaying daily deal sales versus selling in auctions, in the two papers. Compared to the previous literature, we introduce a new dimension for sellers to signal the quality of their items. In particular, for product categories with a common value component where assessing the common value needs expertise (e.g., in the antiques category), we show that selling via auction can signal that the item has a high common value.

Finally, we review the related literature that compares auctions to posted price selling mechanisms. Einav et al. (2013) propose a model to explain the shift from Internet auctions to posted prices and consider two hypotheses: a shift in buyer demand away from auctions, and general narrowing of seller margins that favors posted prices. By using eBay data, they find that the former is more important. There is a significant economics literature that compares auctions to posted price mechanisms. Notably, Wang (1993) compares auctions with posted prices and shows that auctions become preferable when buyers' valuations are more dispersed. In another important paper, Bulow and Klemperer (1996) have shown that the additional revenue one can obtain by attracting one more bidder in an auction without reserve price is greater than the additional revenue by setting the optimal reserve price, hence in a sense establishing that "value of negotiating skills is small relative to value of additional competition" (p. 180). In an empirical work, Bajari et al. (2008) conclude that the choice of sales mechanism may be influenced by the characteristics of the product being sold. To the best of our knowledge, our paper is the first work that considers the signaling effects of the choice of the mechanism on buyers' beliefs. Specifically, we show that the choice of selling mechanism can be used by sellers of high-quality items as a signal of their item's quality.

1.2 Model

We consider a model with two buyers and one item. We assume that there are two types of buyers, *experts* and *nonexperts*, and each buyer is an expert with probability p. Given anonymity of online marketplaces, we assume that each buyer does not know whether his opponent is an expert or not.[10]

In our model, the items sold in online auctions have differing levels of "quality value," which may reflect the condition of a used good or the relative efficacy of a product among its competitors. Note that this value is similar to a common value in that its benefit accrues equally to both expert bidders (who can accurately predict quality value) and nonexpert bidders (who do not know the quality value). We assume that the quality value, denoted by a binary random variable C with realizations 0 and $c > 0$, is known only by experts and is the same for both experts and nonexperts (therefore it can be described as a common value). Moreover, the items sold in online auctions also have differing levels of "private value," which may reflect bidders' private tastes for the items, or whether they have immediate needs for the items. Each bidder may have a different private value. We assume that the private value, denoted by a binary random variable V with realizations 0 and $v > 0$, is learned privately by both experts and nonexperts.

[10]On eBay and most other auction platforms, identities of bidders are revealed only after an auction ends. Furthermore, bidders can easily hide their type by creating and using a new account online.

The total value of the item for a bidder is the sum of the quality value and an additional private value component. More specifically, we assume that C has a binary distribution: $\Pr[C = c] = q$ (high common value) and $\Pr[C = 0] = 1 - q$ (low common value), also V (for each bidder) has a binary distribution: $\Pr[V = v] = r$ (high private value) and $\Pr[V = 0] = 1 - r$ (low private value). We assume that c, v, p, q and r are common knowledge. Moreover, buyers' private value types are privately known by all buyers, and the realization of C is privately known only by experts (nonexperts know only the prior probability distribution). The total value of the item for each bidder is simply $C + V$, where C is the quality value of the item and V is the buyer's specific private value.

We model the online auction with a hard close as a two-stage bidding game where the second stage represents the very last opportunity to submit a bid (the sniping window), while the first stage represents the whole window of time preceding the close. Even though in practice the period before the sniping window is a dynamic game, we model it (Stage 1) by allowing each bidder to submit a single bid: to reconcile this with reality, we can think of the highest bid that a bidder submitted before the sniping window as the first-stage bid. Bidders can observe competitors' bids of Stage 1 and respond to them in Stage 2; however, they do not have enough time to respond to competitors' bids of Stage 2. It is worth mentioning that we can derive all of our results with a more realistic dynamic game model of the first stage.[11] However, though it is a bit more involved, it does not add any further insight to our analysis, so we use the simpler two-stage formulation here.

Motivated by the fact that bidding in the sniping window has the risk of losing the bid due to erratic internet traffic, we assume that a bid in stage 2 goes through only with probability $1 - \delta$ for sufficiently small $\delta \geq 0$. Throughout the paper, we assume that $0 \leq \delta \leq \bar{\delta}$ where $\bar{\delta}$ is defined in Section 1.A.3. This assumption implies that the risk of the bid not going through, due to δ, is not large enough to outweigh the benefit of sniping for experts. We provide an example of equilibrium structure when $\delta > \bar{\delta}$ in Section 1.B.2. The assumption of small δ is also consistent with industry numbers that show that the rate of failure of sniping bids is less than 1%.[12]

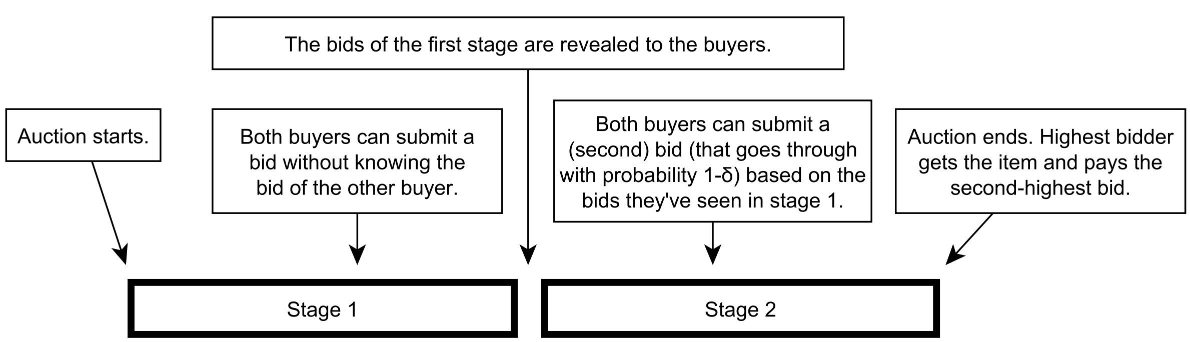


Figure 1.1: Timeline of the game

The timing of the model is as follows (see also Figure 1.1). Before Stage 1, each buyer knows his own type (expert or nonexpert), but not the type of the other buyer. If a buyer

[11] We can consider a dynamic auction in the time interval $[0, 1)$ and sniping at time 1.

[12] For example, see https://www.quicksnipe.com/faq.php (accessed January 2016).

is an expert, he also knows the common value (whether $C = 0$ or $C = c$). All buyers also know their buyer-specific private values (whether $V = 0$ or $V = v$). In Stage 1, both buyers simultaneously submit their bids. After Stage 1 and before Stage 2, both buyers observe the other buyer's bid, and may be able to infer their opponent's type (and values). In Stage 2, both buyers simultaneously decide if they want to increase their bid from Stage 1, and if so by how much. In other words, bids of Stage 2 have to be greater than or equal to bids of Stage 1. Stage 2 bids are received by the auctioneer with probability $1 - \delta$. If the bid of Stage 2 is lost for a bidder (with probability δ), the auctioneer continues to use the bid of Stage 1 for that bidder. After Stage 2, the item is given to the buyer with the highest bid at the price of the second-highest bid. If there is a tie between two bidders of different values, then the item goes to the one of higher value; if both have the same value but are of different types, the tie is broken in favor of the nonexpert; if both bidders have the same value and type, the tie is broken randomly.[13]

In auctions with a soft close, there are possibly an infinite number of stages. If a bid is submitted at any stage, bidders can submit another bid in the next stage. The game ends when no bid is submitted in some stage. We also consider posted prices in Section 1.4. In this game, the seller posts a price z and the bidders then decide whether to buy at this price. The trade takes place at the posted price z if and only if at least one bidder is interested in the item. If both bidders want the item, each of them gets the item with probability $\frac{1}{2}$. Finally, in both types of auctions, soft and hard close, and in posted price, we assume that the platform fee is a constant fraction ξ of the selling price and is paid by the seller.

1.3 Effect of Experts on Buyer Strategies

In this section, we describe the equilibria of the auction game (a formal complete treatment is in Section 1.A.1). We derive conditions under which experts use sniping, in equilibrium, to protect their information about the common value of the item. Furthermore, we show that, under certain conditions, nonexperts with high private value bid aggressively—even above their expected valuation—to compete with experts.

We call an expert/nonexpert with high/low private value a *high/low expert/nonexpert*. Our main lemma characterizing the equilibrium (Lemma 1.2 in Appendix 1.A) splits the values of v into nine ranges depending on the relative values of c, v, p, r, and q. Our characterization labels the strategies for each of the four types of players as one of five different behaviors: (i) a *sniping* strategy is adopted only by experts and involves mimicking the nonexperts in the first stage and bidding their true value only in the second stage; (ii) a *truthful* strategy involves bidding the truthful (expected) value and revising it in the second stage under any additional relevant information; (iii) an *aggressive* strategy is adopted only by high nonexperts and involves bidding over the expected value to have a chance of winning against the experts – we discuss this strategy in detail in subsection 1.3.2; (iv) a *mixed* strategy is a mixed version of the truthful and aggressive strategies; (v) an *underbidding* strategy is used only by low nonexperts, where they bid lower than their expected value for

[13] For a full description and motivation of the tie-breaking rule, please see Section 1.B.3. We demonstrate that our results continue to hold if we change the rule to break the tie in favor of experts rather than nonexperts.

the item.

1.3.1 Experts Induce Sniping

Lemma 1.2 presents necessary and sufficient conditions for each of the above strategies to emerge in equilibrium for each type of bidder. In particular, we show that low experts use the sniping strategy if and only if $v \le c \cdot \frac{(1-p)(1-q)r}{2pq(1-r)+(1-p)r}$, while high experts always use a sniping strategy.

Note that the expression $c \cdot \frac{(1-p)(1-q)r}{2pq(1-r)+(1-p)r}$ is decreasing in q and p, and increasing in r and c. In other words, a low expert's incentive to snipe increases as p or q decrease, and as r or c increase.

To see why, first note that a low expert snipes only if the common value is high. A low value of p (i.e., there are few experts in the market), a low value of q (i.e., there are few high quality items in the market), or a high value of c (i.e., quality difference between low-quality and high-quality items is large), all indicate that the low expert's information, that the common value is high, is *valuable*. This motivates the low expert to snipe and hide this information. Therefore, as p decreases, q decreases, or c increases, the threshold on v for the low expert to snipe increases. Moreover, a high value of r indicates that the opponent is likely to have a high private value. Therefore, as r increases, the probability that the low expert would win the item *without* sniping decreases, which increases his motivation to snipe. As a result, as r increases, the threshold on v for the low expert to snipe increases.

1.3.2 Impact of Experts on Nonexperts' Strategy

A high nonexpert's optimal strategy depends on the value of v. If v is sufficiently high ($cq + v \ge c$), a high nonexpert's expected value for the item is higher than c. In this case, high nonexperts always win the competition against low experts. For smaller values of v, the situation is more interesting. By bidding their expected value against experts, high nonexperts win only when the common value is low. Therefore, high nonexperts have to bid higher than their expected value (aggressive strategy and mixed strategy) to win a high-common-value item against low experts. Note that bidding above the expected value does not necessarily mean that they have to pay more than their expected value, because the auction format is second price. The only risk is that if two high nonexperts compete with each other, they may both bid above their expected value and end up paying more than their expected value. In this case, a nonexpert's payoff could be negative. Our first proposition discusses the conditions under which nonexperts bid more than their expected value.

Proposition 1.1. *If the expected value of a high nonexpert for the item is less than the common value of the item (i.e., $cq + v < c$), the high nonexpert may bid more than his valuation for the item in equilibrium. Moreover, the probability of overbidding increases as the fraction of experts in the market (i.e., p) increases.*

Proposition 1.1 shows that if the value of v is high enough, nonexperts always take the risk of over paying, and bid above their expected value in order to win against experts. However, if v is not sufficiently large, a nonexpert over bids only with some probability (depicted in Figure 1.2). This mixed strategy allows the nonexperts to mitigate the risk of over paying

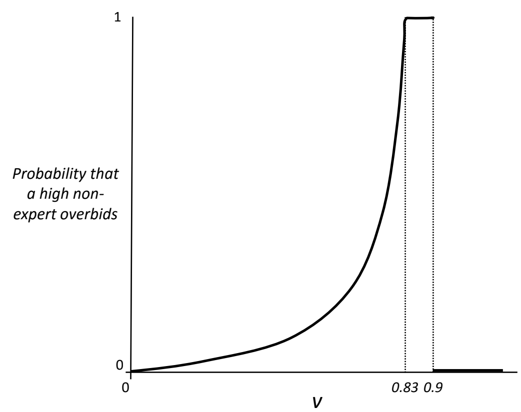


Figure 1.2: Probability that a high nonexpert overbids as v increases for $p = 0.3$, $r = 0.5$, $q = 0.1$, and $c = 1$.

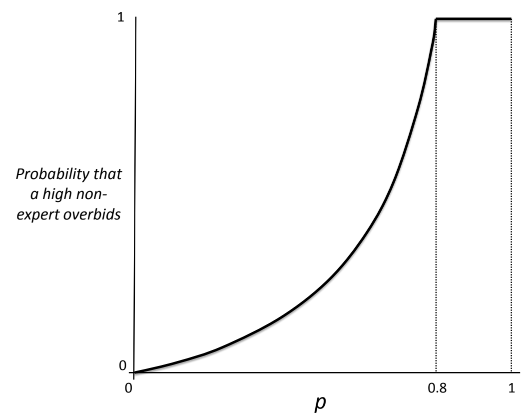


Figure 1.3: Probability that a high nonexpert overbids as p increases for $v = 0.5$, $r = 0.5$, $q = 0.1$, and $c = 1$.

due to competition with another nonexpert. Furthermore, Proposition 1.1 shows that as the probability p that the opponent is an expert increases, a nonexpert's willingness to take the risk and bid above his expected value increases (depicted in Figure 1.3).

1.4 Effect of Experts on Platform Strategies

An important assumption in Proposition 1.1 is that experts can hide their information by sniping. The platform can eliminate sniping by extending the duration of the auction whenever a bid is submitted (this is the soft-close auction format). In this case, nonexperts always have enough time to respond to experts' bids and, therefore, do not have to bid above their expected valuation.

We show that, under certain conditions, nonexperts' aggressive behavior leads to higher revenue for the platform to the extent that the platform benefits from allowing sniping (by enforcing a hard close). In other words, experts' ability to hide their information forces the nonexperts to bid more aggressively, and ultimately leads to higher revenue for sellers and for the platform. This result also relates to platform strategies regarding the revelation of information. In Section 1.4.3, we show the breakdown of the *linkage principle* by showing that the platform may benefit from withholding quality information from the buyers when the buyers are heterogeneous in their level of expertise.

1.4.1 An Auction with a Soft Close

We now consider a model in which sniping is not possible. One way to prevent sniping is by extending the duration of the auction by a few minutes every time there is a bid near the current end time of the auction. This auction is called an auction with a soft close and was used by the now defunct Amazon Auctions. A way to model this is by starting with a game that has only one stage and every time there is a bid during the current stage, the auction extends for one more stage. In other words, every time someone makes a bid, the other buyers can see it and respond to it. In Section 1.4.2, we first characterize the equilibrium for a model of soft-close auctions—the details are in Lemma 1.3 in Section 1.A.2. Then we compare seller's revenue and the platform's revenue across the two models. The goal is to see which ending rule results in better revenues for the sellers (and therefore for the platform).

1.4.2 Effect of Experts on Platform Revenue

Here we summarize the key implications of Lemma 1.3 that appears in Appendix 1.A: when the soft-close format is used, high nonexperts bid their expected value. If they see a bid of c, they infer that the opponent is a low expert and the common value is high. In that case, they increase their bid to c to win the item at price c. On the other hand, with soft close, experts always reveal the value of a high-common-value item to nonexperts. This increases the nonexperts willingness to pay and in some cases leads to higher revenue for the seller. However, when there is a soft close, nonexperts do not have to bid above their valuation. This reduces the competition and can hurt sellers' revenue as well as the platform's revenue. In Lemma 1.1 we see that sellers can benefit from a hard close under certain conditions. We use this lemma to analyze the platform's incentive in having a hard close.

Lemma 1.1. *When $cq + v < c$, the seller of an item with low common value always has higher expected revenue in a hard close than in a soft close, whereas the seller of an item with high common value has higher revenue in hard than soft close if and only if p is sufficiently large.*

Lemma 1.1 shows that the seller of an item with low common value always benefits from a hard close. This is intuitive because a hard close causes sniping, which prevents the flow of information from experts to nonexperts. Therefore, when there is a hard close, nonexperts are more likely to overpay for an item with low common value. The interesting part is that even the seller of an item with high common value benefits from a hard close if p is high enough. This is because when there is a hard close, nonexperts know that they will not be able to infer the common value, and therefore, have to bid more aggressively to win the item. As we observe in Proposition 1.1, this aggressive bidding behavior increases as p increases. If p is sufficiently large, the positive effect of this aggressive bidding behavior on seller's revenue can dominate the negative effect of the lack of information flow, and result in higher revenues for the seller of a high-quality item with a hard close than with a soft close. Using the same argument, we can see that the platform can also benefit from a hard close when p is sufficiently large. This result is formalized in Proposition 1.2.

Proposition 1.2. *If the expected value of the high nonexperts for the item is less than the common value of the item (i.e., $cq + v < c$), and the fraction of experts in the market (i.e., p) is sufficiently large, the platform's revenue from a hard close is higher than that from a soft close.*

A graphical illustration of Proposition 1.2 is depicted in Figure 1.4. When $cq + v < c$ ($v/c < 0.9$ in the figure), the region where a hard close provides higher revenue appears when v is sufficiently larger than c, and p is sufficiently large. This is because higher v and

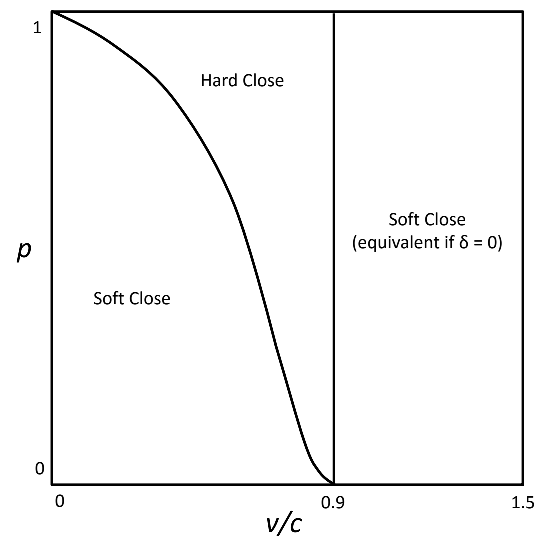


Figure 1.4: The regions are labeled with the format that provides higher revenue for the platform (for $r = 0.5$ and $q = 0.1$). Note that $0.9 = 1 - q$.

higher p both lead to nonexperts' aggressive bidding, as we saw in Figures 1.2 and 1.3 and Proposition 1.1.

Proposition 1.2 shows that for some items the platform's revenue is higher in a hard close, while for other items the revenue is higher in a soft close. Ideally, the optimal strategy for a platform would be to use different policies for different items. However, in practice, platforms may have to use the same policy for all items for other reasons (e.g. consistent user experience). Therefore, the optimal policy will depend on the distribution of the items and the volume of the transactions across the parameter space.

1.4.3 Experts and the Breakdown of the Linkage Principle

Finally, we discuss the connection between the hard-close format and revelation of information in the marketplace. Note that a hard close allows the experts to protect their information about the value of the item. We know that the platform sometimes benefits from a hard close. This could suggest that the platform may also benefit from *withholding information* about the value of the item. This is an important implication because it is in contrast with the well-known "linkage principle" in auction theory (Milgrom and Weber, 1982).

The linkage principle states that auction platforms (e.g., auction houses) benefit from committing to reveal all available information about an item, positive or negative. The platform revealing the information reduces the downside risk of winning the item, also known as the winner's curse. But we show that there is also a downside in revealing the information in the presence of heterogeneous bidders, and the platform may sometimes benefit from committing to not revealing the information.

Our result shows that when bidders are *asymmetric* in terms of their information about the value of the item, bidders with less information have to bid more aggressively, otherwise, they only win the item when bidders with more information do not want the item (i.e., the common value is low). This aggressive behavior incentivizes the platform to withhold any information about the quality value of an item. This result is formalized in the following corollary.

Corollary 1.1. *In auctions with hard close, for medium values of* p *and* $\frac{v}{c}$, *committing to reveal the common value to the buyers decreases platform's revenue.*

We should note that the region in Figure 1.4 where the hard-close format provides higher revenue is the same as the region in Corollary 1.1 in which the platform prefers to withhold the common value information.

Our model is different from the model in Milgrom and Weber (1982) in several aspects. However, the breakdown of the linkage principle is due to only two differences in modeling assumptions. First, we allow the bidders to be heterogenous in terms of their information about the value of the item. Second, bidders do not know how much information other bidders have in this regard. We can show that even in a sealed bid second price auction, a special case of the model in Milgrom and Weber (1982), introducing these two aspects can lead to the breakdown of the linkage principle. Furthermore, both of these aspects are required for the linkage principle to break down. In particular, if bidders are asymmetric in terms of how much information they have about the value of the item, but they *know* how much information other bidders have (e.g., whether the opponent is an expert or not), Campbell and Levin (2000) establish that the linkage principle still holds.

Finally, note that Corollary 1.1 applies only to settings in which the platform has access to some valuable information about the item that is not easily available to all the bidders. For example, using historical market data, eBay provides a quality score for used items in certain categories. Another example is the free vehicle history reports that eBay provided for some time but later discontinued.[14]

So far we have discussed the effect of the existence of experts on nonexperts' and the platform's decisions. In Section 1.5, we analyze the effect of experts on sellers' choice of selling mechanism. In particular, we show that the existence of experts can help the sellers of items with a high common value to signal the value of their items to nonexperts.

1.5 Effect of Experts on Seller Strategies

In this section, we show that the existence of experts in the market could help the sellers to signal the quality/common value of their item to nonexperts. We look at sellers' choice of selling mechanism between an auction and a posted price sale.[15] We call the seller of an item with high common value a high-type seller, and the seller of an item with low common value a low-type seller. A seller is high-type with probability q where q is common knowledge. A seller naturally knows his own type; experts also know the seller's type (since they know the common value of items being offered). But nonexperts do not know the seller's type. We investigate whether a seller can signal his type using the selling mechanism (auction

[14] `http://announcements.ebay.com/2009/11/free-vehicle-history-reports-on-ebay-motors/` (accessed January 2016)

[15] In Section 1.B.5, we further consider the seller's choice of closing format (hard versus soft) as a signaling mechanism.

versus posted price). In particular, we derive conditions for the existence of a separating equilibrium. We show that existence of enough experts in the market is a necessary condition for a separating equilibrium to exist; furthermore, when the fraction of experts in the market, p, is sufficiently large, a separating equilibrium exists only for moderate values of $\frac{v}{c}$.

A seller sets his selling mechanism M (posted price or auction). In case of posted price, M also includes the price. For a mechanism M, we assume that all nonexperts have the same belief about a seller who uses M. In general, nonexperts' belief about a mechanism is the probability that they think a seller using that mechanism is a high type. However, since we consider only pure strategy Nash equilibria of the game, the nonexperts' belief about a mechanism is limited to three possibilities: Low (L), High (H), and Unknown (X). In belief L, nonexperts believe that a seller using mechanism M is always a low-type seller. In belief H, nonexperts believe that a seller using mechanism M is always a high-type seller. Finally, in belief X, nonexperts cannot infer anything about the seller's type and believe that the seller is high-type with probability q.

Nonexperts have beliefs about each mechanism M. In equilibrium, the beliefs must be consistent with the sellers' strategies. In particular, if both types of sellers use the same mechanism in (a pooling) equilibrium, the nonexperts' belief for that mechanism must be X. If the two types of sellers use different mechanisms in (a separating) equilibrium, the nonexperts' belief for the mechanism used by the low-type seller must be L and for the mechanism used by the high-type seller must be H. Furthermore, in an equilibrium, given the nonexperts' beliefs, sellers should not be able to benefit from changing their strategies.

Note that sniping is relevant only when the buyers' belief about some mechanism M is X. Therefore, in a separating equilibrium, the platform's decision on whether to use a soft or hard close does not affect buyers' equilibrium behavior or sellers' strategies. In other words, the following analysis applies to both soft- and hard-close cases.

In general, signaling games can have infinitely many equilibria, supported by different out-of-equilibrium beliefs in the game. Therefore, proving just the existence of an equilibrium with certain characteristics may not be a strong result. To further strengthen the support for our result that selling in auction can be used by high-type sellers as a signal of quality, we show that, under certain conditions, such an equilibrium is the only separating equilibrium that survives the "Intuitive Criterion" refinement. The Intuitive Criterion, introduced by Cho and Kreps (1987), is an equilibrium refinement that requires out-of-equilibrium beliefs to place zero weight on types that can never gain from deviating from a fixed equilibrium outcome. The Intuitive Criterion has been used in various signaling papers in the marketing literature including, but not limited to, Simester (1995), Desai and Srinivasan (1995) and Jiang et al. (2011).

Proposition 1.3 below shows that when the fraction of experts in the market is sufficiently large and the value of $\frac{v}{c}$ is moderate, there exists a unique separating equilibrium in which a high-type seller chooses an auction and a low-type seller chooses posted price as their respective selling mechanisms. A proof and related analysis are provided in Section 1.A.4. Figure 1.5 shows the regions in which this separating equilibrium exists and is unique as a function of p and v/c.

Let us define

$$\nu_1 = \min\left(\frac{(1-p)(1-p(1-2r(1-r)))}{2r(1-r)}, \frac{(1-p)^2}{2(1-p(1-p))(1-r)r}\right),$$
$$\nu_2 = \min\left(\frac{(1-pr)^2}{r(p(2-pr)-r)}, \frac{1}{2r(1-r)}\right),$$
$$\nu_3 = \min\left(\frac{1-r}{2r}, \frac{(1-p)(2-r(1-p))}{r\left(4+(2-p(2-p))r^2-2r(3-p)\right)}\right).$$

Proposition 1.3. *If* $\frac{v}{c} \in [\nu_1, \nu_2]$*, there exists a separating equilibrium in which a high-type seller uses an auction and a low-type seller uses a posted price* v*. Furthermore, if* $\frac{v}{c} \in (\nu_1, \nu_3)$*, this is the only separating equilibrium that survives the Intuitive Criterion refinement. Finally, there exists no separating equilibrium in which a low-type seller uses an auction.*

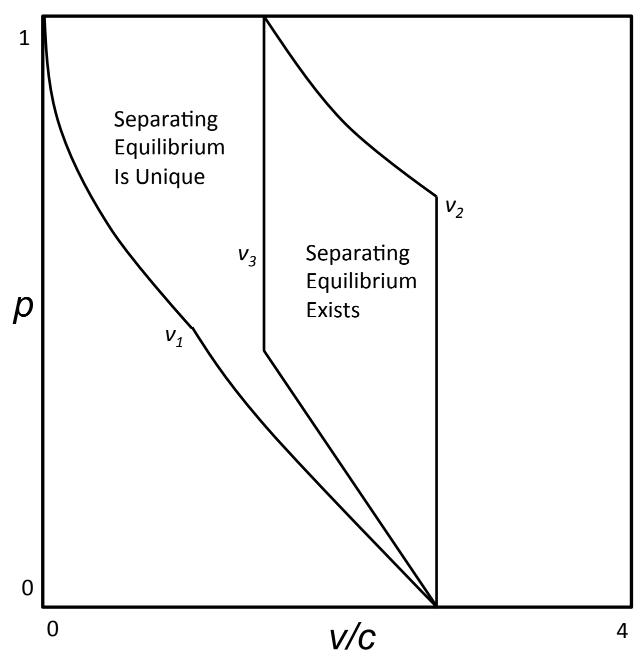


Figure 1.5: The graph shows the existence and uniqueness of a separating equilibrium in which the high-type seller uses auction and the low-type seller uses posted price, assuming $r = \frac{1}{4}$.

The proof and a more elaborate discussion of Proposition 1.3 are relegated to Appendix 1.A. The intuition behind the proof of Proposition 1.3 is as follows. First, note that in general, an auction is more favorable to a high-type than a low-type seller. This is because, in auctions, the price is determined by bidders, and expert bidders do not bid high when the seller is low-type. This allows the high-type seller to separate himself from the low-type seller by selling in an auction. But for this separating equilibrium to exist, the low-type seller's incentive to mimic has to be sufficiently low and the high-type seller's incentive to separate has to be sufficiently high. These two forces give us the thresholds ν_1 and ν_2 for existence (and ν_3 for uniqueness under IC refinement) of this equilibrium.

In a separating equilibrium, even nonexperts know that the low-type seller is a low type. Hence, nonexperts are willing to pay at most v for the item sold by the low-type seller. Therefore, the low-type seller's incentive to mimic increases as v or p decreases. If p and v are sufficiently small, since the low-type seller's incentive to mimic is sufficiently large, a separating equilibrium does not exist. This is captured by condition $\frac{v}{c} \geq \nu_1$ in Proposition 1.3, and is represented by the left contour in Figure 1.5.

On the other hand, as $\frac{v}{c}$ increases, the common value matters less, and the high-type seller's incentive to signal his type (and to separate himself) decreases. When $\frac{v}{c}$ is large enough, we show that the high-type seller chooses to sell via an auction only if p is sufficiently small. This gives us the second condition for existence of this separating equilibrium, namely, $\frac{v}{c} \leq \nu_2$. The condition for uniqueness of the equilibrium, $\frac{v}{c} \leq \nu_3$, follows a similar intuition.

It is interesting to note that the seller's strategy in a separating equilibrium, and the conditions for existence of this equilibrium, do not depend on q. Intuitively, this is because buyers can always infer the seller's type in a separating equilibrium; therefore, when considering the seller's strategy and possible out-of-equilibrium deviations, the ex-ante probability that the seller is high type does not matter.

A Note on Hard- vs. Soft-close Formats. In this section, motivated by eBay's platform, we studied sellers' choice of auction versus posted price. It is theoretically interesting to know what happens, when limited to using auctions, if sellers can choose between hard-close and soft-close formats.[16] This is the mechanism that was employed by the now defunct Yahoo Auctions. In Section 1.B.5, we show that if sellers can choose between soft-close and hard-close formats, the only equilibrium that survives D1 criterion refinement[17] is the one in which both types of sellers use the soft-close format (as a pure strategy pooling equilibrium). Furthermore, nonexperts' belief in the hard-close format will be low. This implies that sellers who choose the hard-close format (out of equilibrium) will earn less revenue in expectation. Our results are consistent with the empirical findings of Glover and Raviv (2012) that show that the soft-close format leads to higher revenue than the hard-close format, and that sellers with less experience are more likely to use the hard-close format. Our explanation, however, is different from theirs, as we attribute the revenue difference to buyers' beliefs and the underlying signaling mechanism as opposed to sniping.

1.6 Conclusion

In this paper, we examined important questions for the buyers, sellers, and the platform of an online market supporting auctions and posted prices. We answered questions about optimal behavior for each of them using the well-documented presence of expertise among the bidders as the key underlying assumption. In particular, we studied the impact of the

[16] We are grateful to an anonymous referee for suggesting this question.

[17] Intuitively, D1 equilibrium refinement requires out-of-equilibrium beliefs to be supported on types that have the most to gain from deviating from a fixed equilibrium. For an extended discussion, see Fudenberg and Tirole (1991, Section 11.2).

presence of expert bidders in online markets using a simple model of auctions with a hard close and posted prices. Motivated by large number of *used* items sold in online markets such as eBay.com, we supposed that items have differing levels of "quality" (which we model as common values), and different bidders have different capacities (which we model as expertise) to predict the quality. Bidders with low expertise may be affected by bids earlier in the auction, as these can be interpreted as signals for the quality of the item. In our model, sniping emerges as an equilibrium strategy for experts to hide their information about the quality of the item in hard-close auctions.

Our results provide several important managerial implications.

- We show that, as a consequence of sniping behavior in equilibrium by the experts in hard-close auctions, nonexpert buyers with less information have to bid aggressively, i.e., more than their expected value. This result highlights the compensatory behavior adopted by the large majority of bidders (nonexperts) that arises endogenously in these common marketplaces.

- Surprisingly, given the aggressive behavior of nonexperts, the platform's revenue can be *higher* in hard-close auctions (where sniping is prevalent) than in soft-close auctions (where sniping cannot happen). This is a new, as-yet unexplored addition to the variety of explanations of why many online auction sites use the hard-close rather than the soft-close format.

- Another interesting implication of nonexperts' aggressive behavior is that the platform can benefit in its revenue from committing to hide the information. This result has important managerial implications, as it suggests that when buyers are heterogeneous in terms of their information about the value of the item, the linkage principle does not always hold.

- When sellers can choose between auction and posted-price formats, a seller may be able to signal the high quality (or authenticity) of his item to the buyers by selling in an auction and thus separate himself from low-quality-item sellers as long as there are enough experts in the market. This provides useful guidance to vendors in such markets, where the magnitude and extent of these decisions can be moderated based on the degree and extent of the presence of expert buyers in the mix. This result also provides a new explanation for the success of auctions in categories such as antiques, art, and collectibles, where common value and therefore expertise are important.

Collectively, our work sheds light on the important differences that arise when knowledgeable or expert buyers are introduced to online marketplaces, and leads to useful guidelines for all participants in such markets.

1.A Appendix

In this appendix, we present detailed explanations of the results. First, we discuss the analyses and proofs of Sections 1.3 and 1.4, in Sections 1.A.1 and 1.A.2, respectively. Then, we provide details of the role of the parameter δ in our model in Section 1.A.3. Finally, in

Section 1.A.4, we detail the results of Section 1.5. Some of the proofs and longer discussions are relegated to Appendix 1.B.

1.A.1 Analyses and Proofs of Section 1.3

In this section, we formally characterize the equilibria of the auction game.

Based on the relation of the parameters c, v, p, r, and q, we split the set of possible parameter values into nine mutually exclusive and collectively exhaustive ranges. In the first four ranges, we have that $cq + v < c$ and $v < cq$; in the next two, we have $cq + v < c$ and $v \geq cq$, in the next two, we have $cq + v \geq c$ and $v < cq$, and in the last range, we have $cq + v \geq c$ and $v \geq cq$.

Consider the function

$$f(c, p, r, q) = c \cdot \frac{(1-p)(1-q)r}{2pq(1-r) + (1-p)r}.$$

Let $m_1 = f(c, p, r, q), m_2 = f(c, p, 1-r, 1-q), M_1 = f(c, p, 1, q) = c \cdot (1-q)$, and $M_2 = f(c, p, 1, 1-q) = c \cdot q$. It is easy to verify that $m_1 \leq M_1$ and $m_2 \leq M_2$. We consider nine different cases as follows: $v \in [0, \min\{m_1, m_2\})$, $v \in [m_1, \min\{m_2, M_1\})$, $v \in [m_2, \min\{m_1, M_2\})$, $v \in [\max\{m_1, m_2\}, \min\{M_1, M_2\})$, $v \in [M_2, m_1)$, $v \in [\max\{m_1, M_2\}, M_1)$, $v \in [M_1, m_2)$, $v \in [\max\{m_2, M_1\}, M_2)$, and $v \in [\max\{M_1, M_2\}, +\infty)$.

To describe an equilibrium, we use the notation (s_1, s_2, s_3, s_4), which means that a high expert follows the strategy s_1, a low expert follows the strategy s_2, a high nonexpert the strategy s_3, and a low nonexpert the strategy s_4. For the bidding strategies of each type we use the following notation:

- For a high expert, consider the following strategies:
  - s_1^{HE}: If $C = 0$, he bids v in the first stage and does nothing in the second stage. If $C = c$, he bids $cq + v$ in the first stage and bids $c + v$ in the second stage (sniping strategy).
  - s_2^{HE}: If $C = 0$, he bids v in the first stage and does nothing in the second stage. If $C = c$, he bids c in the first stage and bids $c + v$ in the second stage (sniping strategy).
- For a low expert, consider the following strategies:
  - s^{LE}: If $C = 0$, he does nothing. If $C = c$, he bids $cq + v$ in the first stage and c in the second stage (sniping strategy).
  - t^{LE}: If $C = 0$, he does nothing. If $C = c$, he bids c in the first stage and nothing in the second stage (truthful strategy).
- For a high nonexpert, consider the following strategies:
  - x^{HNE}: He bids $cq + v$ in the first stage. If he sees a bid other than $0, v, cq$, or $cq + v$ in the first stage, he bids $c + v$ in the second stage. Otherwise, he bids c in the second stage with probability $1 - a$, where $a = 1 - \frac{2p(1-r)qv}{(1-p)r(c-(cq+v))}$ (mixed strategy).

- o^{HNE}: He bids c in the first stage. If he sees a bid other than $0, v, cq$, or c in the first stage, he bids $c+v$ in the second stage. Otherwise, he does nothing in the second stage (aggressive strategy).
- t^{HNE}: He bids $cq+v$ in the first stage. If he sees a bid other than $0, v, cq, c$, or $cq+v$ in the first stage, he bids $c+v$ in the second stage (truthful strategy).

- For a low nonexpert, consider the following strategies:
  - x^{LNE}: He bids v in the first stage. He bids cq in the second stage with probability $1-g$, where $g = \frac{\frac{2pr(1-q)v}{(1-p)(1-r)(cq-v)}-\delta}{1-\delta}$ (mixed strategy).
  - u^{LNE}: He bids v in the first stage and nothing in the second stage (underbidding strategy).
  - t^{LNE}: He bids cq in the first stage and nothing in the second stage (truthful strategy).

We describe equilibrium bidding strategies for buyers in the nine cases in the following lemma.

Lemma 1.2. *For the auction model described in Section 1.2, the buyers' equilibrium bidding strategies are given below.*

1. *If $v \in [0, \min\{m_1, m_2\})$, the set of strategies $(s_1^{HE}, s^{LE}, x^{HNE}, x^{LNE})$ forms an equilibrium.*
2. *If $v \in [m_1, \min\{m_2, M_1\})$, the set of strategies $(s_2^{HE}, t^{LE}, o^{HNE}, x^{LNE})$ forms an equilibrium.*
3. *If $v \in [m_2, \min\{m_1, M_2\})$, the set of strategies $(s_1^{HE}, s^{LE}, x^{HNE}, u^{LNE})$ forms an equilibrium.*
4. *If $v \in [\max\{m_1, m_2\}, \min\{M_1, M_2\})$, the set of strategies $(s_2^{HE}, t^{LE}, o^{HNE}, u^{LNE})$ forms an equilibrium.*
5. *If $v \in [M_2, m_1)$, the set of strategies $(s_1^{HE}, s^{LE}, x^{HNE}, t^{LNE})$ forms an equilibrium.*
6. *If $v \in [\max\{m_1, M_2\}, M_1)$, the set of strategies $(s_2^{HE}, t^{LE}, o^{HNE}, t^{LNE})$ forms an equilibrium.*
7. *If $v \in [M_1, m_2)$, the set of strategies $(s_1^{HE}, t^{LE}, t^{HNE}, x^{LNE})$ forms an equilibrium.*
8. *If $v \in [\max\{m_2, M_1\}, M_2)$, the set of strategies $(s_1^{HE}, t^{LE}, t^{HNE}, u^{LNE})$ forms an equilibrium.*
9. *If $v \in [\max\{M_1, M_2\}, +\infty)$, the set of strategies $(s_1^{HE}, t^{LE}, t^{HNE}, t^{LNE})$ forms an equilibrium.*

The proof of Lemma 1.2 is relegated to Section 1.B.1.

Proof of Proposition 1.1. This result comes directly from Lemma 1.2. We can see that when $m_1 \leq v < M_1$, nonexperts overbid all the time, and when $v < m_1$, they overbid with some probability. We can check in the proof of Lemma 1.2 that the probability of over bidding is $1 - a = \frac{2p(1-r)qv}{(1-p)r(c-(cq+v))}$. It is easy to see that this is an increasing function on p. $\square$

1.A.2 Analyses and Proofs of Section 1.4

Expert strategies for soft-close auctions

As before, for the bidding strategies of each type of buyer, we use the following notation:

- For a high expert, consider the following strategy:
  - t'^{HE}: If $C = 0$, he bids v in the first stage and nothing later. If $C = c$, he bids $c + v$ in the first stage and nothing later (truthful strategy).
- For a low expert, consider the following strategy:
  - t'^{LE}: If $C = 0$, he does nothing. If $C = c$, he bids c in the first stage and nothing later (truthful strategy).
- For a high nonexpert, consider the following strategy:
  - t'^{HNE}: He bids $cq + v$ in the first stage. If he sees a bid of c or $c + v$ at some point and $cq + v < c$, he bids c in the next stage (truthful strategy).
- For a low nonexpert, consider the following strategies:
  - x'^{LNE}: He bids v in the first stage. In the second stage, he bids cq with probability $1 - w$, where $w = \frac{2pr(1-q)v}{(1-p)(1-r)(cq-v)}$, and nothing later (mixed strategy).
  - u'^{LNE}: He bids v in the first stage and nothing later (underbidding strategy).
  - t'^{LNE}: He bids cq in the first stage and nothing later (truthful strategy).

Lemma 1.3. *In a platform with soft close,*

1. *if $v \in [0, m_2)$, the set of strategies $(t'^{HE}, t'^{LE}, t'^{HNE}, x'^{LNE})$ forms an equilibrium;*
2. *if $v \in [m_2, M_2)$, the set of strategies $(t'^{HE}, t'^{LE}, t'^{HNE}, u'^{LNE})$ forms an equilibrium;*
3. *if $v \in [M_2, +\infty)$, the set of strategies $(t'^{HE}, t'^{LE}, t'^{HNE}, t'^{LNE})$ forms an equilibrium.*

Proof. With soft close, an expert is going to bid his true valuation at some point, because anything less than the true valuation will result in a lower payoff. If there is a nonexpert opponent he is going to respond to that; therefore the expert may as well bid truthfully from the first stage. More specifically, the strategies for the experts will be as follows:

- High Expert: If $C = 0$, bids v in the first stage and nothing later. If $C = c$, bids $c + v$ in the first stage and nothing later (strategy t'^{HE}).

- Low Expert: If $C = 0$, does nothing. If $C = c$, he bids c in the first stage and nothing later (strategy t'^{LE}).

For the high nonexpert, the strategy is simple as well. He will bid his expected valuation in the first stage, which is $cq + v$. If the opponent bids c or $c + v$ in the first stage (or at some later point), he will understand that he is an expert and that $C = c$, therefore if $cq + v < c$ he will bid c in the next stage (the minimum possible bid that maximizes his payoff). This is strategy t'^{HNE}.

If $cq \leq v$ (i.e. $v \geq M_2$), then a low nonexpert will bid his expected valuation in the first stage, which is cq, and then he will not do anything (strategy t'^{LNE}). Because, even if, for example, he sees a bid of c and realizes that the common value is high, by bidding c and winning the item, his payoff is still 0.

If $v < cq$ (i.e. $v < M_2$), then a low nonexpert doesn't want to bid cq from the beginning because if the opponent is a high expert and $C = 0$, he will end up with negative payoff. So, he bids v in the first stage, i.e., the maximum he can without the risk above, and waits. If he sees a bid other than v from the opponent, he will lose anyway, so it doesn't matter what strategy he will follow next, and we assume he will follow the same strategy as if he sees a bid of v. If he sees a bid of v, then he bids cq in the second stage with probability $1 - w$. No matter what happens in the second stage, he does nothing in the third stage. We need now to calculate the probability w.

First of all, if he does nothing in the second stage and he sees a bid of cq, he realizes that the opponent is another low nonexpert, but there is no reason to bid something higher because his expected payoff will be 0. If the opponent doesn't bid as well, then the auction ends, and there is no third stage. Therefore, his payoff if he sees a bid of v in the first stage and he does nothing in the second, is

$$\underbrace{\frac{pr(1-q)}{pr(1-q)+(1-p)(1-r)}}_{\text{opponent is high expert and } C=0}(0) + \underbrace{\frac{(1-p)(1-r)}{pr(1-q)+(1-p)(1-r)}}_{\text{opponent is low nonexpert}}(w\frac{cq-v}{2} + (1-w)0).$$

If he bids cq in the second stage, his payoff is

$$\underbrace{\frac{pr(1-q)}{pr(1-q)+(1-p)(1-r)}}_{\text{opponent is high expert and } C=0}(-v) + \underbrace{\frac{(1-p)(1-r)}{pr(1-q)+(1-p)(1-r)}}_{\text{opponent is low nonexpert}}(w(cq-v) + (1-w)0).$$

We need these two expressions to be equal, from which we get

$$w = \frac{2pr(1-q)v}{(1-p)(1-r)(cq-v)}.$$

This is always non-negative, and it is < 1 iff

$$v < \frac{c(1-p)(1-r)q}{2pr(1-q)+(1-p)(1-r)} = m_2.$$

Therefore, if $v < m_2$, the low nonexpert follows the strategy x'^{LNE}.

If $v \geq \frac{c(1-p)(1-r)q}{2pr(1-q)+(1-p)(1-r)} = m_2$ (and $v < M_2$), then it is sub optimal to bid cq, therefore we set $w = 1$ (strategy u'^{LNE}). $\square$

Proof of Lemma 1.1. For a low seller, a hard close is always better, because the bid of every bidder is greater than or equal to his bid when there is a soft close.

For a high seller, we know from Proposition 1.1 that as p increases, high nonexperts bid more and more aggressively. This makes the revenue higher as p increases, in the hard-close format. Therefore, to show the result, it is enough to show that for $p \approx 1$ the revenue with hard close is better than the revenue with soft close.

When $p \approx 1$, it holds that $m_1 \approx m_2 \approx 0$; therefore there are only two relevant equilibria in Lemma 1.2 (cases 4 and 6, since it is also $v < M_1$) and two in Lemma 1.3 (cases 2 and 3). Case 4 of Lemma 1.2 corresponds to case 2 of Lemma 1.3 and case 6 of Lemma 1.2 corresponds to case 3 of Lemma 1.3. We can see that all bids are the same in both models except the bids of the high nonexpert, which are higher with a hard close (the high nonexpert is overbidding in the equilibria 4 and 6 of Lemma 1.2). Therefore, overall the expected revenue is higher for a high seller with the hard-close format.

This is also illustrated in Figure 1.A.1, which shows which policy gives higher revenue to the high seller in different regions of the parameter space. Notice that this is slightly different from Figure 1.4, which refers to the platform's revenue.

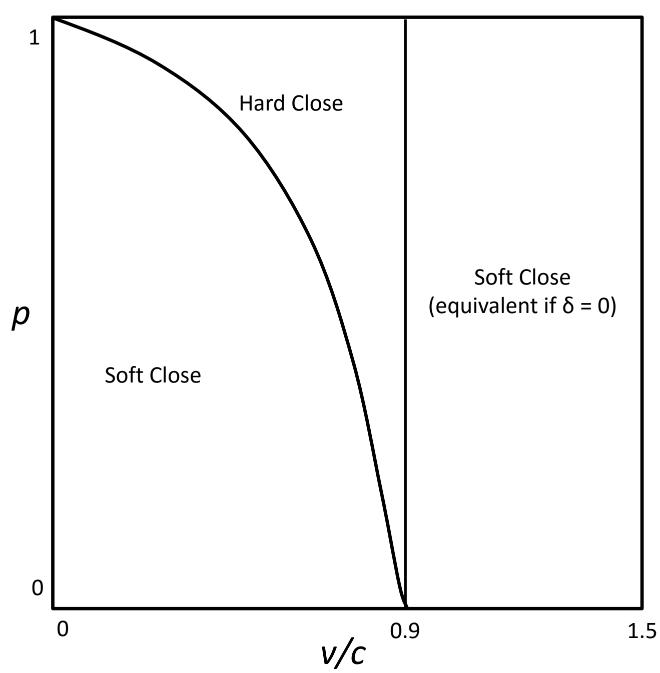
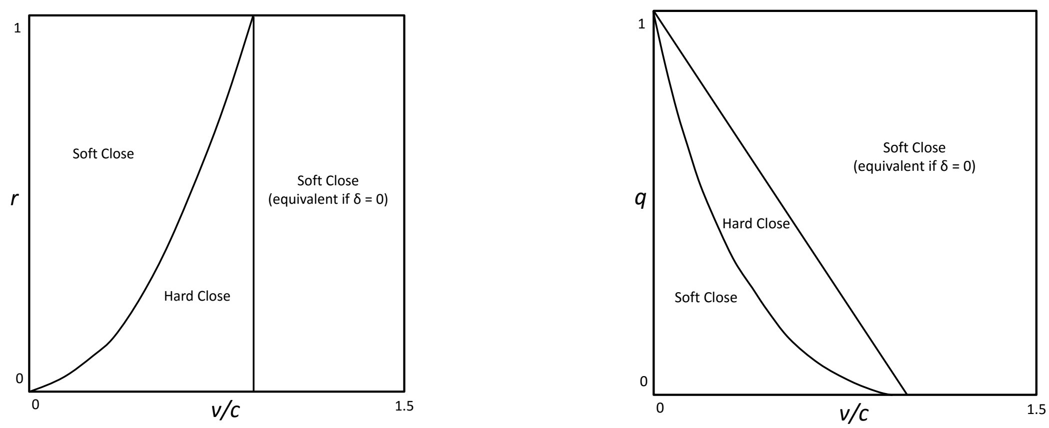


Figure 1.A.1: The regions show whether a hard close provides higher revenue for a high seller (for $r = 0.5$ and $q = 0.1$). This figure is slightly different from Figure 1.4 in that this compares formats that provide higher revenue for a high seller versus the earlier figure that does the same for the overall platform revenue.

$\square$

Proof of Proposition 1.2. This result follows directly from Lemma 1.1. Since a low seller always benefits from a hard close, and a high seller benefits for large p, the expected platform's revenue is better with a hard close for sufficiently large p. $\square$

The analogue of Figure 1.4 where the format that provides the higher revenue is labeled as a function of other parameters in the model is presented in Figure 1.A.2. In particular, in Figure 1.A.2a we can see that as r (the probability that a bidder has high private value) increases, the region where a hard close provides higher revenue becomes smaller. This is because from the perspective of a high nonexpert, high r means higher probability that the other bidder is a high nonexpert too, which in turn means lower willingness to bid aggressively in the hard-close format. This results in lower revenue for a hard close when r is large.



(a) For $p = 0.5$ and $q = 0.1$.

(b) For $r = 0.5$ and $p = 0.5$.

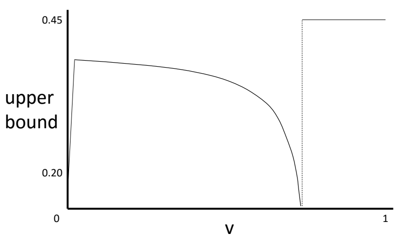
Figure 1.A.2: The regions are labeled with the format that provides higher revenue for the platform. This figure is an analogue of Figure 1.4, presenting the same result for other parameter variations.

Proof of Corollary 1.1. When the platform reveals the common value to everyone, all bidders bid their true valuation. Therefore, in the region in which the aggressive bidding of high nonexperts makes hard close better than soft close for the platform (the middle region in Figure 1.4), the platform prefers to hide the common value so that the high nonexperts keep bidding higher than their true valuation. $\square$

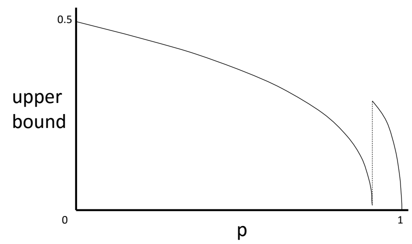
1.A.3 Upper-bound Condition on δ

In our model, we assume that δ is sufficiently small, i.e., $\delta \leq \bar{\delta}$. This upper-bound condition is calculated as the minimum of at most three different thresholds coming from the indifference conditions for the three of the types of players: high experts, low experts, and low nonexperts. These are the conditions that reflect the relations between the parameter values at which the current set of strategies are no longer in equilibrium. Intuitively, when $\delta > \bar{\delta}$, the cost of sniping (i.e., the risk that the bid does not go through) outweighs its benefits. Therefore, some types of bidders decide not to snipe. Since other types of bidders know this, they also have to update their strategies. As a result, we get different (and several cases of) equilibrium structures for $\delta > \bar{\delta}$. We provide an example of this in Section 1.B.2.

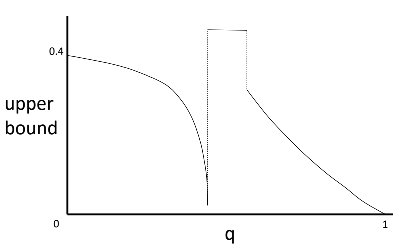
A thorough discussion and calculation of the thresholds for $\bar{\delta}$ is deferred to Section 1.B.2. The exact definition of $\bar{\delta}$ is given in Lemma 1.4. To provide some intuition, in Figure 1.A.3 we present plots of $\bar{\delta}$ as a function of v, of p, of q, and of r.



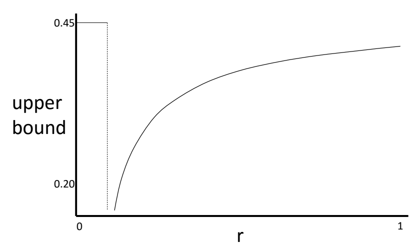
(a) As a function of v, for $c = 1$, $q = 0.1$, $r = 0.5$, $p = 0.5$.



(b) As a function of p, for $c = 1$, $q = 0.1$, $r = 0.5$, $v = 0.3$.



(c) As a function of q, for $c = 1$, $r = 0.5$, $p = 0.5$, $v = 0.3$.



(d) As a function of r, for $c = 1$, $q = 0.1$, $p = 0.5$, $v = 0.3$.

Figure 1.A.3: Plots of the upper-bound $\bar{\delta}$ as a function of v, of p, of q, and of r.

1.A.4 Analyses and Proofs of Section 1.5

We use the following notation to explain the results of this section: Let $\pi_T^B(M)$, where $T \in \{L, H\}$ and $B \in \{L, H, X\}$ denote the expected profit of a seller who uses mechanism $M \in \{A, (B, z)\}$ (where A denotes auction, and (B, z) denotes posted price where the price is z), has type T, and nonexperts believe has type B. Let M^{pool} be the mechanism that both types of sellers use in a pooling equilibrium.

The revenue of a high- or low-type seller in an auction, where nonexperts have belief high or low, is given in the following formulas. Recall that p is the probability of being expert, and r is the probability of having high value.

$$\pi_H^H(A) = c + r^2 v;$$

$$\pi_L^L(A) = r^2 v;$$

$$\pi_H^L(A) = \begin{cases} cp^2 + rv(2(1-p)p(1-r) + r) & \text{if } v \leq c, \\ cp\left(2(1-p)r + p - 2(1-p)r^2\right) + r^2 v & \text{if } v > c; \end{cases}$$

$$\pi_L^H(A) = \begin{cases} c(1-p)^2 + rv(2(1-p)p(1-r) + r) & \text{if } v \leq c, \\ c(1-p)(p(2(1-r)r - 1) + 1) + r^2 v & \text{if } v > c; \end{cases}$$

Similarly, the revenue of a high- or low-type seller using posted price with price z, in each of the four cases, is

$$\pi_H^H(B, z) = \begin{cases} z & \text{if } z \leq c, \\ (2r - r^2)z & \text{if } c < z \leq c + v, \\ 0 & \text{otherwise;} \end{cases}$$

$$\pi_H^L(B, z) = \begin{cases} (1 - (1-p)^2(1-r)^2)z & \text{if } v \leq c \text{ and } z \leq v, \\ (2p - p^2)z & \text{if } v \leq c \text{ and } v < z \leq c, \\ (2pr - p^2 r^2)z & \text{if } v \leq c \text{ and } c < z \leq c + v, \\ (1 - (1-p)^2(1-r)^2)z & \text{if } v > c \text{ and } z \leq c, \\ (2r - r^2)z & \text{if } v > c \text{ and } c < z \leq v, \\ (2pr - p^2 r^2)z & \text{if } v > c \text{ and } v < z \leq c + v, \\ 0 & \text{otherwise;} \end{cases}$$

$$\pi_L^H(B, z) = \begin{cases} (1 - p^2(1-r)^2)z & \text{if } v \leq c \text{ and } z \leq v, \\ (2(1-p) - (1-p)^2)z & \text{if } v \leq c \text{ and } v < z \leq c, \\ (2r(1-p) - r^2(1-p)^2)z & \text{if } v \leq c \text{ and } c < z \leq c + v, \\ (1 - p^2(1-r)^2)z & \text{if } v > c \text{ and } z \leq c, \\ (2r - r^2)z & \text{if } v > c \text{ and } c < z \leq v, \\ (2r(1-p) - r^2(1-p)^2)z & \text{if } v > c \text{ and } v < z \leq c + v, \\ 0 & \text{otherwise;} \end{cases}$$

$$\pi_L^L(B, z) = \begin{cases} (2r - r^2)z & \text{if } z \leq v, \\ 0 & \text{otherwise.} \end{cases}$$

Proof of Proposition 1.3. We prove the proposition in three parts. In part A, we show that there is no separating equilibrium in which the high-type seller uses posted price and the low-type seller uses auction. In part B, we show that when $v \in [\nu_1, \nu_2]$, there exists a separating equilibrium in which the high-type seller uses auction and the low-type seller uses posted price v. Finally, in part C, we show that for $v \in (\nu_1, \nu_3)$, this is the only equilibrium that survives the Intuitive Criterion refinement.

Part A. Note that $\pi_L^L(A) < \pi_L^L(B, v)$, which means that conditioned on the type of sellers being revealed, the low-type seller always prefers posted price v to auction. Therefore, the low-type seller never uses an auction in a separating equilibrium.

Part B. Note that for a separating equilibrium in which the high-type seller uses auction and the low-type seller uses posted price to exist, the following two conditions are necessary and sufficient:

$$\pi_L^L(B, z) \leq \pi_L^H(A),$$

$$\pi_H^H(A) \geq \pi_H^L(B, z).$$

The first condition guarantees that the low-type seller cannot benefit from deviating and the second condition guarantees that the high-type seller cannot benefit from deviating. $\pi_L^L(B, z)$ is optimized at $z = v$, and is equal to $(2r - r^2)v$. Having this less than or equal to $\pi_L^H(A)$, and using basic calculus, gives us the condition $\frac{v}{c} \geq \nu_1$. Similarly, solving the second inequality for v gives us condition $\frac{v}{c} \leq \nu_2$. If nonexpert buyers' beliefs are L for posted prices and H for auction, then $\nu_1 \leq \frac{v}{c} \leq \nu_2$ is also sufficient for existence of this equilibrium.

Part C. Finally, we show that if $\nu_1 \leq \frac{v}{c} \leq \nu_3$, the separating equilibrium in which the high-type seller uses auction and the low-type seller uses posted price is the only pure strategy separating Nash equilibrium that survives the Intuitive Criterion refinement. Assume for sake of contradiction that there exists another separating equilibrium. We already know from Part A of this proof that the low-type seller cannot be using auction. Therefore, both types must be using posted price (with different prices) in this equilibrium. Using the same argument as in Part B of the proof, we know that the low-type seller must be using posted price v. Suppose that the high-type seller is using posted price ζ. For this to be a separating equilibrium, the low-type seller should not benefit from deviating and mimicking the high-type seller: $\pi_L^L(B, v) \geq \pi_L^H(B, \zeta)$. Using basic calculus, we can show that this implies the following condition on ζ. We must have $\zeta \leq \frac{(r-2)v}{(p-1)(pr-r+2)}$. Let $\pi^* = \pi_H^H(B, \zeta)$ be the profit of the high-type seller (in the hypothetical separating equilibrium) subject to this constraint.

If $\pi_H^H(A) > \pi^*$, then the high-type seller benefits from deviating to auction unless nonexperts' belief about auction is L. But note that if $\frac{v}{c} > \nu_1$, nonexperts' belief about auction cannot be L according to the Intuitive Criterion refinement. Specifically, since the high-type seller benefits from deviating to auction and the low-type seller never benefits from deviating to auction even if buyers' belief in auction is H, according to the Intuitive Criterion refinement, buyers' belief in auction should be H. Therefore, if $\pi_H^H(A) > \pi^*$ the high-type seller benefits from deviating to auction and the hypothetical equilibrium cannot exist. Using basic calculus, the condition $\pi_H^H(A) > \pi^*$ reduces to $\frac{v}{c} \leq \nu_3$. Therefore, for $\frac{v}{c} \in (\nu_1, \nu_3)$, the separating equilibrium in which the high-type seller uses auction and the low-type seller uses posted price is the only pure strategy separating Nash equilibrium that survives the Intuitive Criterion refinement. □

1.B Additional Appendix

In this appendix, we first give a complete proof of the equilibrium-characterization lemma 1.2 in Section 1.B.1. In Section 1.B.2, we calculate the explicit upper bound $\bar{\delta}$ on the value of

δ alluded to in Section 1.A.3. We then provide a robustness check on the choice of our tie-breaking rule in Section 1.B.3 by showing that the analog of the main lemma 1.2 continues to hold even if we invert the tie-breaking rule to favor experts instead of nonexperts. In the next Section 1.B.4, we show an extrapolation of our main result for the case when the private value distribution has a support of more than two values, thus lending support for our main observations in the limiting continuous case. Finally, in Section 1.B.5, we consider equilibria when sellers sell in an auction but can choose between hard-close versus soft-close formats.

1.B.1 Proof of Lemma 1.2

Proof. We will group the nine equilibria into four cases. These are $v < \min\{M_1, M_2\}$ (equilibria 1, 2, 3 and 4), $M_2 \leq v < M_1$ (equilibria 5 and 6), $\max\{M_2, M_1\} \leq v$ (equilibrium 9), and $M_1 \leq v < M_2$ (equilibria 7 and 8).

Case 1. First, assume that $v < \min\{M_1, M_2\}$. This means that $cq + v < c$ and $v < cq$. Consider the following general set of strategies:

- High Expert: If $C = 0$, bids v in the first stage and does nothing in the second stage. If $C = c$, bids $cq + v$ in the first stage and bids $c + v$ in the second stage.
- Low Expert: If $C = 0$, does nothing. If $C = c$, he bids $cq + v$ in the first stage and c in the second stage.
- High nonexpert: Bids $cq + v$ in the first stage. If he sees a bid other than $0, v, cq$, or $cq + v$ in the first stage, he bids $c + v$ in the second stage. Otherwise, he bids c in the second stage with probability $1 - a$.
- Low nonexpert: Bids cq in the first stage with probability b and v with probability $1 - b$. If his bid was v, he bids cq in the second stage with probability $1 - g$.

The probabilities a, b, g are as yet undetermined . For now, we just assume that $a > 0$. We will examine if anyone has an incentive to change his strategy and at the same time try to determine the probabilities and the conditions for which the above is an equilibrium. These conditions will give us the proof that equilibria 1 and 3 are correct. Later, we will relax the assumption on a and examine what happens when $a = 0$; this will lead us to conditions that equilibria 2 and 4 are correct, and will conclude the proof of the four equilibria in the first case.

— High Expert with $C = 0$: His valuation is v and now he bids v in the first stage. If he does nothing in the first stage and he bids v in the second stage, then there is some probability that his bid will not go through with a payoff of 0, and in the case it goes through, his payoff would be the same in all cases as if he had bid v in the first stage (against a low nonexpert, his payoff is 0 in both cases). Therefore, it is optimal for him to follow this strategy.

— High Expert with $C = c$: His valuation is $c+v$ and now he bids $cq+v$ in the first stage and bids $c + v$ in the second stage. We consider three alternative strategies which dominate all the rest, and we prove that he doesn't have any incentive to deviate to any of them.

One strategy is to bid v in the first stage and $c+v$ in the second stage. This strategy has a different result for him only if his bid in the second stage doesn't go through. In that case, by having a bid of v instead of a bid of $cq+v$ can only decrease his payoff.

Another strategy is to bid 0 in the first stage and $c+v$ in the second. The behavior of the rest of the bidders will not change, but his payoff will decrease because he can lose in some cases whereas the bid of $cq+v$ would give him a positive payoff.

The last strategy is to bid $c+v$ (or c) in the first stage and nothing (or $c+v$) in the second. However, if we assume that his bid will go through in the second stage, with the alternative strategy the result would be the same in all cases except in the case he faces a high nonexpert, where his payoff strictly decreases. Therefore, since δ is sufficiently small[18], it is better to bid in the second stage.

— Low Expert with $C=0$: His value is 0 and he does nothing, which is optimal for him.

— Low Expert with $C=c$: His value is c. The payoff if he bids c in the first stage is

$$\begin{aligned}A(\delta) = \quad & \underbrace{pr((1-\delta)0+\delta(c-(cq+v)))}_{\text{opponent is high expert}} \\ & + \underbrace{p(1-r)((1-\delta)0+\delta(c-(cq+v)))}_{\text{opponent is low expert}} \\ & + \underbrace{(1-p)r((1-\delta)0+\delta(c-(cq+v)))}_{\text{opponent is high nonexpert}} \\ & + \underbrace{(1-p)(1-r)(b(c-cq)+(1-b)(g(c-v)+(1-g)(1-\delta)(c-cq)+(1-g)\delta(c-v)))}_{\text{opponent is low nonexpert}}.\end{aligned}$$

The payoff with the current strategy is

$$\begin{aligned}B(\delta) = \quad & \underbrace{(1-\delta)}_{\text{bid goes through}} \cdot \Bigg[\underbrace{pr((1-\delta)0+\delta(c-(cq+v)))}_{\text{opponent is high expert}} \\ & + \underbrace{p(1-r)((1-\delta)0+\delta(c-(cq+v)))}_{\text{opponent is low expert}} \\ & + \underbrace{(1-p)r(a(c-(cq+v))+(1-a)(1-\delta)0+(1-a)\delta(c-(cq+v)))}_{\text{opponent is high nonexpert}} \\ & + \underbrace{(1-p)(1-r)(b(c-cq)+(1-b)(g(c-v)+(1-g)(1-\delta)(c-cq)+(1-g)\delta(c-v)))}_{\text{opponent is low nonexpert}}\Bigg] \\ & + \underbrace{\delta}_{\text{bid doesn't go through}} \cdot \Bigg[\underbrace{pr((1-\delta)0+\delta 0)}_{\text{opponent is high expert}} \\ & + \underbrace{p(1-r)\left((1-\delta)0+\delta\frac{c-(cq+v)}{2}\right)}_{\text{opponent is low expert}} \\ & + \underbrace{(1-p)r(0)}_{\text{opponent is high nonexpert}} \\ & + \underbrace{(1-p)(1-r)(b(c-cq)+(1-b)(g(c-v)+(1-g)(1-\delta)(c-cq)+(1-g)\delta(c-v)))}_{\text{opponent is low nonexpert}}\Bigg].\end{aligned}$$

It holds that $B(0)-A(0)=(1-p)ra(c-(cq+v))>0$ (for $a>0$), which means that for sufficiently small δ, $B(\delta)>A(\delta)$, i.e. the current strategy is better.

[18] Formally, this condition means $\delta<\bar{\delta}$, which we discuss in Section 1.B.2.

The alternative is to bid something else in the first stage other than $cq+v$ or c, and c in the second, but this doesn't increase the payoff.

— High nonexpert: His expected valuation is $cq+v$. Bidding something else other than $cq+v$ in the first stage will not change the bidding behavior of the opponent to something better for him, therefore he prefers to bid $cq+v$ in the first stage rather than wait.

In the second stage, it doesn't matter what they do if they see a bid of 0 or v or cq, since the result cannot change. If they see a bid other than $0, v, cq$, or $cq+v$ (like c or $c+v$), something that doesn't happen in the equilibrium, they assume that the common value is high which means that their valuation is $c+v$, so they bid $c+v$. The reason is that the only one who might have incentive to deviate from the current strategies is an expert with $C=c$ who tries to bluff in some way to hide the common value.

If they see a bid of $cq+v$, then they know that their opponent is a high expert with $C=c$, or a low expert with $C=c$, or a high nonexpert. Their payoff by doing nothing in the second stage is

$$\begin{aligned}
A_2 = \quad & \underbrace{\frac{prq}{prq+p(1-r)q+(1-p)r}((1-\delta)0+\delta(c+v-(cq+v)))}_{\text{opponent is high expert and } C=c} \\
& + \underbrace{\frac{p(1-r)q}{prq+p(1-r)q+(1-p)r}((1-\delta)0+\delta(c+v-(cq+v)))}_{\text{opponent is low expert and } C=c} \\
& + \underbrace{\frac{(1-p)r}{prq+p(1-r)q+(1-p)r}(a0+(1-a)(1-\delta)0+(1-a)\delta 0)}_{\text{opponent is high nonexpert}},
\end{aligned}$$

while their payoff by bidding c is

$$\begin{aligned}
B_2 = \quad & \underbrace{(1-\delta)}_{\text{bid goes through}} \cdot \Bigg[\underbrace{\frac{prq}{prq+p(1-r)q+(1-p)r}((1-\delta)0+\delta(c+v-(cq+v)))}_{\text{opponent is high expert and } C=c} \\
& + \underbrace{\frac{p(1-r)q}{prq+p(1-r)q+(1-p)r}((1-\delta)v+\delta(c+v-(cq+v)))}_{\text{opponent is low expert and } C=c} \\
& + \underbrace{\frac{(1-p)r}{prq+p(1-r)q+(1-p)r}(a0+(1-a)(1-\delta)\frac{cq+v-c}{2}+(1-a)\delta 0)}_{\text{opponent is high nonexpert}} \Bigg] \\
& + \underbrace{\delta}_{\text{bid doesn't go through}} \cdot A_2.
\end{aligned}$$

By bidding $c+\epsilon$ for some small $\epsilon>0$, his payoff can only decrease. By bidding $c-\epsilon$, the payoff is the same as if they stay with the bid of $cq+v$ (according to the tie-breaking rule, if two bidders are both high, the nonexpert wins). It holds that

$$B_2 - A_2 = (1-\delta)\left[\frac{p(1-r)q}{prq+p(1-r)q+(1-p)r}(1-\delta)v + \frac{(1-p)r}{prq+p(1-r)q+(1-p)r}(1-a)(1-\delta)\frac{cq+v-c}{2}\right],$$

and we want this to be equal to 0 to permit mixing these strategies, which will give us an expression for the mixing probability a. This is

$$a = 1 - \frac{2p(1-r)qv}{(1-p)r(c-(cq+v))}.$$

This is always ≤ 1. We assumed also that $a > 0$, which is equivalent to $v < \frac{c(1-p)r(1-q)}{2p(1-r)q+(1-p)r} = m_1$. Therefore, we need this condition to have an equilibrium in this case.

If $1 - \frac{2p(1-r)qv}{(1-p)r(c-(cq+v))} \leq 0$, which is equivalent to $v \geq m_1$ and corresponds to $a = 0$, we need a different set of strategies and we consider this case later.

— Low nonexpert: His expected valuation is cq. His payoff if he bids cq in the first stage is

$$\begin{aligned} A_3 = & \underbrace{pr(q0 + (1-q)(-v))}_{\text{opponent is high expert}} \\ & + \underbrace{p(1-r)(q0 + (1-q)0)}_{\text{opponent is low expert}} \\ & + \underbrace{(1-p)r(0)}_{\text{opponent is high nonexpert}} \\ & + \underbrace{(1-p)(1-r)(b0 + (1-b)(g(cq-v) + (1-g)(1-\delta)0 + (1-g)\delta(cq-v)))}_{\text{opponent is low nonexpert}} \end{aligned}$$

His payoff if he bids v in the first stage and follows the current strategy in the second stage is

$$\begin{aligned} B_3 = & \underbrace{pr(q0 + (1-q)(g0 + (1-g)(1-\delta)(-v) + (1-g)\delta 0))}_{\text{opponent is high expert}} \\ & + \underbrace{p(1-r)(q0 + (1-q)0)}_{\text{opponent is low expert}} \\ & + \underbrace{(1-p)r(0)}_{\text{opponent is high nonexpert}} \\ & + (1-p)(1-r)\Bigg[\underbrace{b0 + (1-b)(g^2(\frac{cq-v}{2}) + (1-g)g((1-\delta)(cq-v) + \delta\frac{cq-v}{2})}_{\text{opponent is low nonexpert}} \\ & + g(1-g)((1-\delta)0 + \delta\frac{cq-v}{2}) + (1-g)^2((1-\delta)^2 0 + (1-\delta)\delta(cq-v) + \delta(1-\delta)0 + \delta^2\frac{cq-v}{2}))\Bigg]. \end{aligned}$$

Now, in the second stage, if a low nonexpert with a bid of v sees any bid other than v from the opponent, bidding cq or nothing in the second stage doesn't affect his payoff. If he sees a bid of v, then he knows that the opponent is either a high expert with $C = 0$ or a low nonexpert. If he does nothing in the second stage, his payoff is

$$\begin{aligned} A_4 = & \underbrace{\frac{pr(1-q)}{pr(1-q) + (1-p)(1-r)(1-b)}(0)}_{\text{opponent is high expert and } C=0} \\ & + \underbrace{\frac{(1-p)(1-r)(1-b)}{pr(1-q) + (1-p)(1-r)(1-b)}(g\frac{cq-v}{2} + (1-g)(1-\delta)0 + (1-g)\delta\frac{cq-v}{2})}_{\text{opponent is low nonexpert}}, \end{aligned}$$

while if he bids cq, the payoff is

$$\begin{aligned} B_4 = & \underbrace{(1-\delta)}_{\text{bid goes through}} \Bigg[\underbrace{\frac{pr(1-q)}{pr(1-q) + (1-p)(1-r)(1-b)}(-v)}_{\text{opponent is high expert and } C=0} \\ & + \underbrace{\frac{(1-p)(1-r)(1-b)}{pr(1-q) + (1-p)(1-r)(1-b)}(g(cq-v) + (1-g)(1-\delta)0 + (1-g)\delta(cq-v))}_{\text{opponent is low nonexpert}}\Bigg] \\ & + \underbrace{\delta A_4}_{\text{bid doesn't go through}}. \end{aligned}$$

It must hold that $A_4 = B_4$ to permit mixing these strategies, from which we get an expression for the mixing probability g which is

$$g = \frac{\frac{2pr(1-q)v}{(1-p)(1-r)(1-b)(cq-v)} - \delta}{1-\delta}.$$

This expression is non-negative for sufficiently small δ and it is < 1 iff

$$v < \frac{c(1-p)(1-r)(1-b)q}{2pr(1-q) + (1-p)(1-r)(1-b)}.$$

For $b = 0$ and the corresponding g, we get $A_3 \leq B_3$ (for $g < 1$), therefore the current strategy of the low expert is optimal and we get an equilibrium. For this reason, we set $b = 0$. The above condition then becomes

$$v < \frac{c(1-p)(1-r)q}{2pr(1-q) + (1-p)(1-r)} = m_2.$$

If $v \geq m_2$, then we set $g = 1$ (which corresponds to strategy u^{LNE}).

This ends the proof for equilibria 1 and 3.

When $a = 0$, the strategy for the low expert we considered above is not always optimal. This happens when $v \geq m_1$. More specifically, since he knows that the high nonexpert will bid c in the second stage for sure, he has no reason to wait until the second stage to bid, and bids c from the first stage. With the same logic, since a high nonexpert knows for sure that he will bid c in the second stage, it is even better to bid c from the first stage. Moreover, when a high nonexpert sees a bid of c in the first stage, he doesn't know for sure what the opponent is, so he doesn't increase his bid. This will change also the strategy for the high expert with $C = c$. In the first stage, he prefers to bid c instead of $cq + v$, because a bid of $cq + v$ would reveal that he is a high expert and $C = c$. So, the equilibrium when $v \geq \frac{c(1-p)r(1-q)}{2p(1-r)q+(1-p)r} = m_1$ is as follows.

- High Expert: If $C = 0$, bids v in the first stage and does nothing in the second stage. If $C = c$, bids c in the first stage and bids $c + v$ in the second stage.

- Low Expert: If $C = 0$, does nothing. If $C = c$, he bids c in the first stage and nothing in the second stage.

- High nonexpert: Bids c in the first stage. If he sees a bid other than $0, v, cq$ or c in the first stage, he bids $c + v$ in the second stage. Otherwise, he does nothing in the second stage.

- Low nonexpert: Bids v in the first stage. He bids cq in the second stage with probability $1 - g$.

The proofs for the high expert, the low expert and the low nonexpert are the same. We need to check if the high nonexpert has any reason to change strategy. An alternative strategy for him would be the one he had before, i.e. to bid $cq + v$ in the first stage and c in the second with some probability. So, suppose that he had bidden $cq + v$ in the first stage and he sees a bid of c. His payoff by doing nothing in the second is 0, while the payoff to bid c in the second stage is

$$\begin{aligned}
B' = \quad & \underbrace{(1-\delta)}_{\text{bid goes through}} \cdot \Bigg[\underbrace{\frac{prq}{prq + p(1-r)q + (1-p)r}((1-\delta)0 + \delta(c+v-(c)))}_{\text{opponent is high expert and } C=c} \\
& + \underbrace{\frac{p(1-r)q}{prq + p(1-r)q + (1-p)r}(c+v-c)}_{\text{opponent is low expert and } C=c} \\
& + \underbrace{\frac{(1-p)r}{prq + p(1-r)q + (1-p)r}(\frac{cq+v-c}{2})}_{\text{opponent is high nonexpert}} \Bigg] \\
& + \underbrace{\delta}_{\text{bid doesn't go through}} \cdot 0.
\end{aligned}$$

This is ≥ 0 for $v \geq \frac{c(1-p)r(1-q)}{2prq\delta + 2p(1-r)q + (1-p)r}$, which is true since

$$v > \frac{c(1-p)r(1-q)}{2p(1-r)q + (1-p)r} \geq \frac{c(1-p)r(1-q)}{2prq\delta + 2p(1-r)q + (1-p)r}.$$

Therefore, he is better off by bidding c rather than 0 in the second stage. This means that by bidding in the first stage he can increase his payoff. All other possible strategies are trivially dominated by those we considered above.

This ends the proof for equilibria 2 and 4.

Summarizing the first case, when $a > 0$ (i.e. $v < m_1$) and $g < 1$ (i.e. $v < m_2$), we get the equilibrium $(s_1^{HE}, s^{LE}, x^{HNE}, x^{LNE})$, when $a = 0$ (i.e. $v \geq m_1$) and $g < 1$ (i.e. $v < m_2$), we get the equilibrium $(s_2^{HE}, t^{LE}, o^{HNE}, x^{LNE})$, when $a > 0$ (i.e. $v < m_1$) and $g = 1$ (i.e. $v \geq m_2$), we get the equilibrium $(s_1^{HE}, s^{LE}, x^{HNE}, u^{LNE})$, and when $a = 0$ (i.e. $v \geq m_1$) and $g = 1$ (i.e. $v \geq m_2$), we get the equilibrium $(s_2^{HE}, t^{LE}, o^{HNE}, u^{LNE})$.

Case 2. Assume now that $M_2 \leq v < M_1$. This means that $cq \leq v$ and $cq + v < c$. Consider the following set of strategies:

- High Expert: If $C = 0$, bids v in the first stage and does nothing in the second stage. If $C = c$, bids $cq + v$ in the first stage and bids $c + v$ in the second stage.

- Low Expert: If $C = 0$, does nothing. If $C = c$, he bids $cq + v$ in the first stage and c in the second stage.

- High nonexpert: Bids $cq + v$ in the first stage. If he sees a bid other than $0, v, cq$, or $cq + v$ in the first stage, he bids $c + v$ in the second stage. Otherwise, he bids c in the second stage with probability $1 - a$.

- Low nonexpert: Bids cq in the first stage and nothing in the second.

We now investigate if anyone has incentive to change strategy. For $a > 0$, the arguments for all types of bidders are the same as in the previous case except for the low nonexpert.

The expected valuation of a low nonexpert is cq. Now he bids cq in the first stage and his expected payoff is 0. The only way to get the item is only if he faces another low nonexpert, in which case they both bid cq and there is a tie. But even in this case he has to pay cq, so his payoff is 0. He cannot achieve a better payoff, since it is never optimal to bid something above his expected valuation.

This ends the proof for equilibrium 5.

Similarly as in the previous case, the equilibrium when $v \geq \frac{c(1-p)r(1-q)}{2p(1-r)q+(1-p)r} = m_1$ (which means $a = 0$) is as follows.

- High Expert: If $C = 0$, bids v in the first stage and does nothing in the second stage. If $C = c$, bids c in the first stage and bids $c + v$ in the second stage.
- Low Expert: If $C = 0$, does nothing. If $C = c$, he bids c in the first stage and nothing in the second stage.
- High nonexpert: Bids c in the first stage. If he sees a bid other than $0, v, cq$, or c in the first stage, he bids $c + v$ in the second stage. Otherwise, he does nothing in the second stage.
- Low nonexpert: Bids cq in the first stage and nothing in the second.

This ends the proof for equilibrium 6.

Summarising the second case, when $a > 0$ (i.e. $v < m_1$), we get the equilibrium $(s_1^{HE}, s^{LE}, x^{HNE}, t^{LNE})$, and when $a = 0$ (i.e. $v \geq m_1$), we get the equilibrium $(s_2^{HE}, t^{LE}, o^{HNE}, t^{LNE})$.

Case 3. Next, suppose that $\max\{M_2, M_1\} \leq v$. This means that $\max\{cq, c(1-q)\} \leq v$. We consider the following set of strategies:

- High Expert: If $C = 0$, bids v in the first stage and does nothing in the second stage. If $C = c$, bids $cq + v$ in the first stage and bids $c + v$ in the second stage.
- Low Expert: If $C = 0$, does nothing. If $C = c$, he bids c in the first stage and nothing in the second stage.
- High nonexpert: Bids $cq + v$ in the first stage. If he sees a bid other than $0, v, cq, c,$, or $cq + v$ in the first stage, he bids $c + v$ in the second stage.
- Low nonexpert: Bids cq in the first stage and nothing in the second.

This is the simplest case. Both high and low nonexperts have nothing to lose by bidding their expected valuation, therefore they do so from the first stage. The low nonexpert has no reason to hide his identity, therefore he bids his valuation from the first stage. The same is true for a high expert with $C = 0$. Finally, the high expert with $C = c$ bids the highest possible he can in the first stage without revealing that he is a high expert, which is a bid of $cq + v$, and then he bids $c + v$ in the second stage. If he bids $c + v$ from the first stage, then his payoff strictly decreases because of the possibility that the opponent is a high nonexpert.

This ends the proof for equilibrium 9.

Summarizing the third case, we get the equilibrium $(s_1^{HE}, t^{LE}, t^{HNE}, t^{LNE})$.

Case 4. Finally, suppose that $M_1 \leq v < M_2$. This means that $c(1-q) \leq v < cq$. We consider two cases:

- If $v < \frac{c(1-p)(1-r)q}{2pr(1-q)+(1-p)(1-r)} = m_2$, the following is an equilibrium.
  - High Expert: If $C = 0$, bids v in the first stage and does nothing in the second stage. If $C = c$, bids $cq + v$ in the first stage and bids $c + v$ in the second stage.
  - Low Expert: If $C = 0$, does nothing. If $C = c$, he bids c in the first stage and nothing in the second stage.
  - High nonexpert: Bids $cq+v$ in the first stage. If he sees a bid other than $0, v, cq, c,$, or $cq + v$ in the first stage, he bids $c + v$ in the second stage.
  - Low nonexpert: Bids v in the first stage. He bids cq in the second stage with probability $1 - g$, where $g = \frac{\frac{2pr(1-q)v}{(1-p)(1-r)(cq-v)} - \delta}{1-\delta}$.
- If $v \geq \frac{c(1-p)(1-r)q}{2pr(1-q)+(1-p)(1-r)} = m_2$, the following is an equilibrium.
  - High Expert: If $C = 0$, bids v in the first stage and does nothing in the second stage. If $C = c$, bids $cq + v$ in the first stage and bids $c + v$ in the second stage.
  - Low Expert: If $C = 0$, does nothing. If $C = c$, he bids c in the first stage and nothing in the second stage.
  - High nonexpert: Bids $cq+v$ in the first stage. If he sees a bid other than $0, v, cq, c,$, or $cq + v$ in the first stage, he bids $c + v$ in the second stage.
  - Low nonexpert: Bids v in the first stage and nothing in the second.

For the experts and the high nonexpert, the proofs are similar to the previous case. For the low nonexpert, the proof is similar to the second case.

This ends the proof for equilibria 7 and 8.

Summarizing the fourth case, when $g < 1$ (i.e. $v < m_2$), we get the equilibrium $(s_1^{HE}, t^{LE}, t^{HNE}, x^{LNE})$, and when $g = 1$ (i.e. $v \geq m_2$), we get the equilibrium $(s_1^{HE}, t^{LE}, t^{HNE}, u^{LNE})$. □

1.B.2 More on δ

Recall that δ is the probability that a bid in the second stage does not go through (due to network or other technical difficulties). For the result of Lemma 1.2 to hold, in the model section, we assumed that $\delta \leq \bar{\delta}$. In this section, we elaborate on how to calculate the value of $\bar{\delta}$. We also briefly discuss how the equilibrium structure changes when $\delta > \bar{\delta}$.

We start with the first case of Lemma 1.2, i.e. when $v \leq \min\{m_1, m_2\}$. In that case, the set of strategies $(s_1^{HE}, s_1^{LE}, x^{HNE}, x^{LNE})$ is an equilibrium for sufficiently small δ.

More specifically, there are three threshold values τ_1, τ_2, τ_3, and the case 1 of Lemma 1.2 holds if $\delta \leq \min\{\tau_1, \tau_2, \tau_3\}$. The first threshold, τ_1, corresponds to the strategy of the high expert. When δ exceeds this threshold, a high expert with $C = c$ prefers to bid $c + v$ in the first stage instead of waiting to bid in the second stage (i.e. instead of following strategy s_1^{HE}).

The second threshold, τ_2, corresponds to the strategy of the low expert. When δ exceeds this threshold, a low expert with $C = c$ prefers to bid c in the first stage instead of following the strategy s_1^{LE}. To compute τ_2, we have to find the minimum δ for which $B(\delta) \geq A(\delta)$ in the proof of Lemma 1.2, or equivalently solve the equation $B(\delta) = A(\delta)$ for δ.

The third threshold, τ_3, corresponds to the strategy of the low nonexpert. When δ exceeds this threshold, a low nonexpert prefers to bid cq in the first round, i.e. prefers to follow the strategy t^{LNE} instead of the strategy x^{LNE}. To compute τ_3, we have to find the minimum δ for which the probability g in the proof of Lemma 1.2 is non-negative, or equivalently solve the equation $g(\delta) = 0$ for δ.

The closed-form expressions for the three thresholds are given below.

$$\tau_1 = \frac{-\sqrt{2c^2(p-2)(p-1)(q-1)^2r^2 + 4c(p-2)p(q-1)q(r-1)rv + p^2(r-1)^2v^2} + 2c(p-1)(q-1)r + p(4q-1)(r-1)v}{cp(q-1)r + 2p(2q-1)(r-1)v},$$

$$\begin{aligned}\tau_2 = \max\Bigg\{ &\frac{1}{\frac{(p(r-1)-2r)(c(q-1)+v)}{\sqrt{2}\sqrt{(p(r-1)-2r)(c(q-1)+v)(c(p-1)(q-1)r+v(p(2q(r-1)+r)-r))}} + 1}, \\ &\frac{1}{-\frac{(p(r-1)-2r)(c(q-1)+v)}{\sqrt{2}\sqrt{(p(r-1)-2r)(c(q-1)+v)(c(p-1)(q-1)r+v(p(2q(r-1)+r)-r))}} + 1}\Bigg\} = \\ = &\frac{1}{\frac{(p(r-1)-2r)(c(q-1)+v)}{\sqrt{2}\sqrt{(p(r-1)-2r)(c(q-1)+v)(c(p-1)(q-1)r+v(p(2q(r-1)+r)-r))}} + 1},\end{aligned}$$

$$\tau_3 = \frac{2pr(1-q)v}{(1-p)(1-r)(cq+v)}.$$

The plot in Figure 1.B.1 shows how platform's revenue changes as δ increases in the interval $[0, \min\{\tau_1, \tau_2, \tau_3\}]$.

We can see that as δ increases, platform's revenue decreases. The reason for this is that the equilibrium remains the same, i.e. the bids of the buyers are the same, but the probability that some bids don't go through increases, therefore the expected final price of the item is lower.

We continue with the second case of Lemma 1.2, where the equilibrium is $(s_2^{HE}, t^{LE}, o^{HNE}, x^{LNE})$. Here we need only two bounds for δ, one for the high expert and one for the low nonexpert, since the low expert bids only in the first round independently of the value of δ.

The threshold for the low nonexpert remains the same as in the previous case, i.e. it is τ_3. However, the bound for the high expert will change due to the change in the strategies of the other bidders.

Consider the following strategy for the high expert:

- t^{HE}: If $C = 0$, he bids v in the first stage and does nothing in the second stage. If $C = c$, he bids $c + v$ in the first stage and nothing in the second stage (truthful strategy).

To find the new threshold for the high expert, we need to compare his payoff when he uses the strategy s_2^{HE}, his payoff when he uses the strategy t^{HE}, and see when the first is

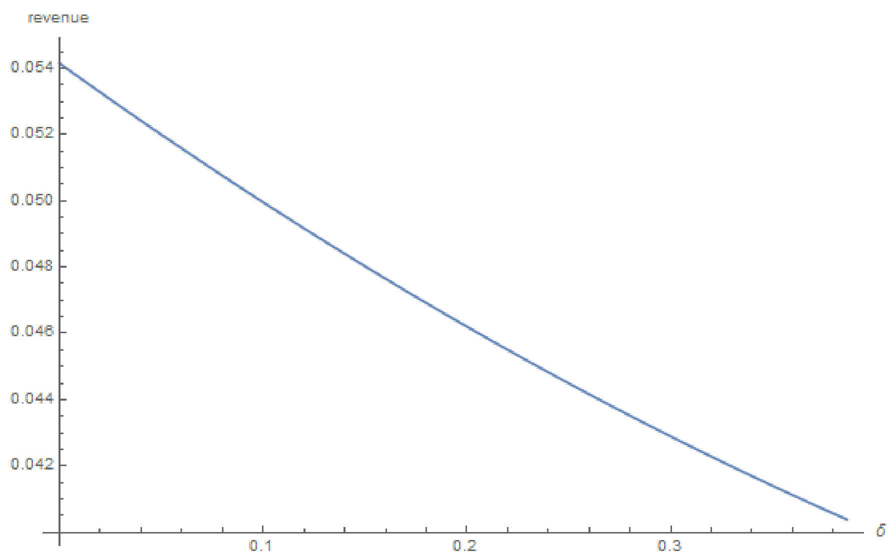


Figure 1.B.1: Platform's revenue as δ increases, for $v = 0.03$, $c = 1$, $p = 0.5$, $r = 0.5$, and $q = 0.1$. It is $\tau_1 = 0.445287$, $\tau_2 = 0.386605$, and $\tau_3 = 0.415385$.

larger than the second, which will make $(s_2^{HE}, t^{LE}, o^{HNE}, x^{LNE})$ an equilibrium. The new threshold is

$$\sigma_1 = \max\left\{\frac{2p + \sqrt{2}\sqrt{(p-2)(p-1)} - 2}{p}, \frac{2p - \sqrt{2}\sqrt{(p-2)(p-1)} - 2}{p}\right\} =$$
$$= \frac{2p + \sqrt{2}\sqrt{(p-2)(p-1)} - 2}{p}.$$

Therefore, the necessary and sufficient condition for the second case of Lemma 1.2 is $\delta \leq \min\{\sigma_1, \tau_3\}$. Similarly for every case of the lemma, we can find the bound for δ. The following result summarizes all the cases.

Lemma 1.4. *The necessary and sufficient condition for δ in Lemma 1.2 is $\delta \leq \bar{\delta}$, where*

$$\bar{\delta} = \begin{cases} \min\{\tau_1, \tau_2, \tau_3\}, & \text{if } v \in [0, \min\{m_1, m_2\}) \text{ (case 1)}, \\ \min\{\sigma_1, \tau_3\}, & \text{if } v \in [m_1, m_2) \text{ (cases 2, 7)}, \\ \min\{\tau_1, \tau_2\}, & \text{if } v \in [m_2, m_1) \text{ (cases 3,5)}, \\ \sigma_1, & \text{if } v \in [\max\{m_1, m_2\}, +\infty) \text{ (cases 4,6,8,9)}. \end{cases}$$

An example of equilibrium for $\delta > \bar{\delta}$.

Since, according to the industry numbers, the probability that the sniping bid does not go through is less than 1%, in the main model we only consider the case in which δ is relatively small. However, it is theoretically interesting to know what happens for larger values of δ. Depending on the value of δ and other parameters in the model, the full analysis leads to too many cases the discussion of which is beyond the scope of this paper. However, to gain some intuition, in the following we discuss one example of the equilibrium structure for $\delta > \bar{\delta}$. Interestingly, we see that the platform's revenue could be non-monotone in δ.

Consider the following parameter values, $c = 1$, $q = 0.1$, $r = 0.5$, $p = 0.5$, $v = 0.7$, and $\delta = 0.44$. We have $v > M_2 = 0.1$ and $v < m_1 = 0.75$; therefore, we are in case 5 of Lemma 1.2. However, it holds that $\tau_1 = 0.348355$ and $\tau_2 = 0.257284$, i.e., δ exceeds the necessary and sufficient threshold given in Lemma 1.4 for $(s_1^{HE}, s^{LE}, x^{HNE}, t^{LNE})$ to be an equilibrium.

In particular, the low expert does not want to snipe since δ is larger than τ_2, so he moves his bid to the first stage. In other words, he prefers to follow the strategy t^{LE} instead of the strategy s^{LE}. This will cause a change in the strategies of the other bidders. The new equilibrium will be $(s_2^{HE}, t^{LE}, o^{HNE}, t^{LNE})$, which happens to be the same as case 6 of Lemma 1.2.

This is because if we cosnider the proof of case 6, the requirements for the equilibrium are satisfied, even though $v < m_1$. First, it holds that $\delta < \sigma_1 = 0.44949$, so the high expert's best response is s_2^{HE}. Moreover, it is the case that $v > \frac{c(1-p)r(1-q)}{2prq\delta+2p(1-r)q+(1-p)r} = 0.698758$, which makes o^{HNE} the best response for the high nonexpert. The value of v for the low expert and the low nonexpert doesn't matter as long as the others follow the aforementioned strategies. Therefore, $(s_2^{HE}, t^{LE}, o^{HNE}, t^{LNE})$ is an equilibrium.

This causes the following interesting phenomenon. Even though in the interval $[0, 0.257284]$, platform's revenue is a decreasing function of δ, as depicted in Figure 1.B.1; as δ increases more, outside this interval, platform's revenue increases. In this example, for $\delta = 0.257284$, platform's revenue is 0.251106, while for $\delta = 0.44$, it is 0.264497.

This is mainly because the low expert has changed his strategy by moving his bid of c from the second stage to the first stage, i.e., from the point where his bid was not going through with probability $\delta = 0.257284$ to the point where his bid always goes through (when $\delta = 0.44$); this change has a positive effect on the expected price of the item.

As δ increases even more (above σ_1), we see similar patterns: intervals in which platform's revenue is a decreasing function of δ, and some 'jumps' of the revenue in between due to the change of strategies by bidders. In the limit, when $\delta \approx 1$, no-one will bid in the second stage and the auction will be like a sealed-bid second price auction where everyone bids in the first stage.

1.B.3 Choice of Tie-breaking Rule

In this section, first we elaborate on the choice of our tie-breaking rule. We argue that this rule always favors the bidder who is willing to bid slightly higher than the current bid (i.e. his payoff continues to remain positive if he slightly raises his bid) which the other bidder is not able to match. Then, we show that our results are robust to the choice of the tie-breaking rule. In particular, we show that the equilibrium strategies remain almost unchanged, and our main results continue to hold, under a very different tie-breaking rule.

Recall that we use the following tie-breaking rule: If there is a tie between a low-type bidder and a high-type bidder, then the item goes to the high-type. If the two bidders are of the same type but of different expertise levels, then the item goes to the nonexpert. Finally, if the two bidders are of the same type and of the same expertise level, then the winner is determined by a fair coin toss.

There are two interesting cases for which the tie-breaking rule has an effect in the equilibria described in the main lemma. The first is when a high nonexpert faces a low expert who knows that $C = c$, and they both bid c. In this case, the high nonexpert is willing to bid above c to win the tie and take the item, because his valuation is higher than c, but the low expert cannot do the same since his valuation is c. Therefore, the tie-breaking rule favors the bidder who would be willing to pay a slightly higher price.

The second case is when a high expert who knows that $C = 0$ faces a low nonexpert, and they both bid v. In this case, the low nonexpert does not want to win the item, because his valuation is below v. Therefore, he has an incentive to bid a bit below v, whereas the high expert does not want to do the same since his valuation is v. Thus, again the tie-breaking rule favors the bidder who has higher willingness to pay.

Intuitively, if we break the tie in favor of the other bidder in any of the above cases, one of the bidders would want to increase or decrease his bid by the smallest possible amount $\epsilon > 0$. Since in our model the strategy space is continuous and not discrete, such ϵ does not exist. We use this tie-breaking rule to avoid such complications. However, to further demonstrate the robustness of our results, in the following, we show that equilibrium strategies, and therefore all of our main results, continue to hold if we change the rule in the opposite direction and favor the experts over the nonexperts.

Changing the tie-breaking rule.

Consider the following alternate tie-breaking rule: If there is a tie between two bidders of different expertise levels, then the item goes to the expert. Otherwise, the winner is determined by a fair coin toss. To reduce the number of cases in the analysis, we assume that $\delta = 0$, and only focus on the bids of the second stage.

As explained above, this game does not have a pure strategy Nash equilibrium unless we discretize the bidding space. We show that as the size of the discretization step converges to zero (i.e., the bidding space converges to continuous), the equilibrium outcome of the new tie-breaking rule converges to that of the old tie-breaking rule.

To discretize the strategy space, we assume that bidders can bid c or $c + \epsilon$, for some very small $\epsilon > 0$, but they cannot bid anything in between. This assumption will come into play when there is a tie between a low expert and a high nonexpert who both bid c (the first case discussed earlier in this section). Note that without this discretization, high nonexperts sometimes want to bid the smallest number strictly larger than c; this is because high nonexperts want to win against experts but lose against other high nonexperts. We assume that the rest of the strategy space remains unchanged.

We start by defining the strategies for the different types of bidders.

- For a high expert, consider the following strategy:
  - t^{HE}: If $C = 0$, he bids v. If $C = c$, he bids $c + v$.
- For a low expert, consider the following strategy:
  - t^{LE}: If $C = 0$, he does nothing. If $C = c$, he bids c.

- For a high nonexpert, consider the following strategies:
  - x^{HNE}: He bids $cq + v$ with probability a and $c + \epsilon$ with probability $1 - a$, where $a := a(\epsilon) = 1 - \frac{2p(1-r)qv}{(1-p)r(c+\epsilon-(cq+v))}$.
  - o^{HNE}: He bids $c + \epsilon$.
  - t^{HNE}: He bids $cq + v$.
- For a low nonexpert, consider the following strategies:
  - x^{LNE}: He bids v with probability g and cq with probability $1 - g$, where $g = \frac{2pr(1-q)v}{(1-p)(1-r)(cq-v)}$.
  - u^{LNE}: He bids v.
  - t^{LNE}: He bids cq.

We define also a new threshold for v, the analogous of the old m_1, that now depends also on ϵ. We have that $m_1 := m_1(\epsilon) = \frac{(c+\epsilon-cq)(1-p)r}{2p(1-r)q+(1-p)r}$.

Intuitively, high nonexperts who want to over-bid now bid $c + \epsilon$ instead of c. This allows them to win against experts, even though ties are broken in favor of experts. We now describe the equilibrium bidding strategies for buyers in nine cases in the following lemma.

Lemma 1.5. *For the auction model described above with the alternative tie-breaking rule, the buyers' equilibrium bidding strategies are given below.*

1. *If $v \in [0, \min\{m_1(\epsilon), m_2\})$, the set of strategies $(t^{HE}, t^{LE}, x^{HNE}, x^{LNE})$ forms an equilibrium.*
2. *If $v \in [m_1, \min\{m_2, M_1\})$, the set of strategies $(t^{HE}, t^{LE}, o^{HNE}, x^{LNE})$ forms an equilibrium.*
3. *If $v \in [m_2, \min\{m_1(\epsilon), M_2\})$, the set of strategies $(t^{HE}, t^{LE}, x^{HNE}, u^{LNE})$ forms an equilibrium.*
4. *If $v \in [\max\{m_1(\epsilon), m_2\}, \min\{M_1, M_2\})$, the set of strategies $(t^{HE}, t^{LE}, o^{HNE}, u^{LNE})$ forms an equilibrium.*
5. *If $v \in [M_2, m_1(\epsilon))$, the set of strategies $(t^{HE}, t^{LE}, x^{HNE}, t^{LNE})$ forms an equilibrium.*
6. *If $v \in [\max\{m_1(\epsilon), M_2\}, M_1)$, the set of strategies $(t^{HE}, t^{LE}, o^{HNE}, t^{LNE})$ forms an equilibrium.*
7. *If $v \in [M_1, m_2)$, the set of strategies $(t^{HE}, t^{LE}, t^{HNE}, x^{LNE})$ forms an equilibrium.*
8. *If $v \in [\max\{m_2, M_1\}, M_2)$, the set of strategies $(t^{HE}, t^{LE}, t^{HNE}, u^{LNE})$ forms an equilibrium.*
9. *If $v \in [\max\{M_1, M_2\}, +\infty)$, the set of strategies $(t^{HE}, t^{LE}, t^{HNE}, t^{LNE})$ forms an equilibrium.*

Proof. We will prove the first four equilibria, i.e. when $v < M_1$ and $v < M_2$, which are the most general. The rest of the cases are similar to the proof of the Lemma 1.2.

We have that $v < \min\{M_1, M_2\}$. This means that $cq + v < c$ and $v < cq$. Consider the following general set of strategies:

- High Expert: If $C = 0$, he bids v. If $C = c$, he bids $c + v$.
- Low Expert: If $C = 0$, he does nothing. If $C = c$, he bids c.
- High nonexpert: He bids $cq + v$ with probability a and $c + \epsilon$ with probability $1 - a$.
- Low nonexpert: He bids v with probability g and cq with probability $1 - g$.

The probabilities a, g are as yet undetermined. We will examine if anyone has incentive to change strategy and at the same time try to determine the probabilities and the conditions for which the above is an equilibrium. These conditions will give us the proof that equilibria 1 and 3 are correct.

Both the high and the low experts bid truthfully and this is optimal for them. This is because they bid in the second stage, therefore they don't have any fear to reveal the common value to nonexperts. They also know their true valuation, and since we have a second-price auction, it is optimal for them to bid their true values.

Now, we consider a high nonexpert. His expected valuation is $cq + v$. Their payoff by bidding $cq + v$ is

$$\begin{aligned}
A_2 = &\underbrace{prq(0)}_{\text{opponent is high expert and } C=c} \\
&+ \underbrace{pr(1-q)(0)}_{\text{opponent is high expert and } C=0} \\
&+ \underbrace{p(1-r)q(0)}_{\text{opponent is low expert and } C=c} \\
&+ \underbrace{p(1-r)(1-q)(v)}_{\text{opponent is low expert and } C=0} \\
&+ \underbrace{(1-p)r(0)}_{\text{opponent is high nonexpert}} \\
&+ \underbrace{(1-p)(1-r)(g(cq) + (1-g)(v))}_{\text{opponent is low nonexpert}},
\end{aligned}$$

while their payoff by bidding $c + \epsilon$ is

$$\begin{aligned}
B_2 = &\underbrace{prq(0)}_{\text{opponent is high expert and } C=c} \\
&+ \underbrace{pr(1-q)(0)}_{\text{opponent is high expert and } C=0} \\
&+ \underbrace{p(1-r)q(v)}_{\text{opponent is low expert and } C=c} \\
&+ \underbrace{p(1-r)(1-q)(v)}_{\text{opponent is low expert and } C=0} \\
&+ \underbrace{(1-p)r\left((1-a)\frac{cq + v - (c+\epsilon)}{2}\right)}_{\text{opponent is high nonexpert}}
\end{aligned}$$

$$+ \underbrace{(1-p)(1-r)(g(cq)+(1-g)(v))}_{\text{opponent is low nonexpert}}.$$

By bidding $c+\epsilon+\zeta$ for some small $\zeta > 0$, his payoff can only decrease. By bidding c, the payoff is the same as with the bid of $cq+v$. It holds that

$$B_2 - A_2 = p(1-r)q(v) + (1-p)r\left((1-a)\frac{cq+v-(c+\epsilon)}{2}\right),$$

and we want this to be equal to 0 to permit mixing these strategies, which will give us an expression for the mixing probability a. This is

$$a = 1 - \frac{2p(1-r)qv}{(1-p)r(c+\epsilon-(cq+v))}.$$

This is always ≤ 1. The inequality $a > 0$ is equivalent to $v < \frac{(c+\epsilon-cq)(1-p)r}{2p(1-r)q+(1-p)r} = m_1$. So, we need this condition for equilibria 1 and 3. If $1 - \frac{2p(1-r)qv}{(1-p)r(c+\epsilon-(cq+v))} \leq 0$, then it is always better for the high nonexpert to bid $c+\epsilon$, so we set $a = 0$ (equilibria 2 and 4).

Next, we consider a low nonexpert. His expected valuation is cq. His payoff if he bids cq is

$$\begin{aligned} A_3 = &\underbrace{pr(q0+(1-q)(-v))}_{\text{opponent is high expert}} \\ &+ \underbrace{p(1-r)(q0+(1-q)0)}_{\text{opponent is low expert}} \\ &+ \underbrace{(1-p)r(0)}_{\text{opponent is high nonexpert}} \\ &+ \underbrace{(1-p)(1-r)(g(cq-v)+(1-g)0)}_{\text{opponent is low nonexpert}}. \end{aligned}$$

His payoff if he bids v is

$$\begin{aligned} B_3 = &\underbrace{pr(q0+(1-q)0)}_{\text{opponent is high expert}} \\ &+ \underbrace{p(1-r)(q0+(1-q)0)}_{\text{opponent is low expert}} \\ &+ \underbrace{(1-p)r(0)}_{\text{opponent is high nonexpert}} \\ &+ \underbrace{(1-p)(1-r)\left(g\frac{cq-v}{2}\right)}_{\text{opponent is low nonexpert}}. \end{aligned}$$

It must hold that $A_3 = B_3$ to permit mixing these strategies, from which we get an expression for the mixing probability g which is

$$g = \frac{2pr(1-q)v}{(1-p)(1-r)(cq-v)}.$$

This expression is always non-negative, and it is < 1 iff

$$v < \frac{c(1-p)(1-r)q}{2pr(1-q)+(1-p)(1-r)} = m_2.$$

If $v \geq m_2$, then we set $g = 1$ (which corresponds to strategy u^{LNE}).

This ends the proof. □

Notice that as ϵ goes to 0, the bidding strategies of Lemma 1.5 approach the strategies of the main Lemma 1.2. This means that the analogues of Proposition 1.1 and Proposition 1.2 will continue to hold with the alternative tie-breaking rule.

1.B.4 Distribution of Bidders' Private Value

In the main model, we assumed that V has binary distribution with support $\{0, v\}$. In this section we relax that assumption and show that our main result, that nonexperts sometimes bid more than their expected value, still holds. More specifically, for $k \geq 2$, we assume that the private value of each bidder is $V = \frac{i \cdot v}{k-1}$ with probability r_i, where $i \in \{0, 1, \ldots, k-1\}$ and $\sum_{i=0}^{k-1} r_i = 1$. In the main model, we had $k = 2$, $r_2 = r$, and $r_1 = 1 - r$.

The tie-breaking rule is a generalization of what we had in the main model. We assume that in case of a tie the bidder with the highest private value wins. If the two bidders have the same private value, then the nonexpert wins. If both bidders have the same private value and the same expertise level, then the winner is determined with a fair coin toss.

To simplify the analysis, we assume that $\delta = 0$ and only focus on the bids in the second stage. It is easy to see that it is weakly dominant for all the experts to bid their true valuation. In other words, if an expert has private value $\frac{i \cdot v}{k-1}$, he will bid $\frac{i \cdot v}{k-1}$ when the common value is low ($C = 0$), and $c + \frac{i \cdot v}{k-1}$, when the common value is high ($C = c$). We also assume that $v \leq c \cdot \min\{q, 1-q\}$ (which corresponds to the condition $v \leq \min\{M_1, M_2\}$ of our main model).

We show that, in equilibrium, nonexperts will mix between at most three different bids. More specifically, if a nonexpert is of type i, meaning that his private value is $\frac{i \cdot v}{k-1}$, he will mix between $\frac{(i+1) \cdot v}{k-1}$, $c \cdot q + \frac{i \cdot v}{k-1}$, and $c + \frac{(i-1) \cdot v}{k-1}$. The bid $\frac{(i-1) \cdot v}{k-1}$ is employed because he wants to lose against the experts of higher type when $C = 0$. The bid $c \cdot q + \frac{i \cdot v}{k-1}$ is used because this is his expected valuation and this is the bid he wants to have against a nonexpert. The bid $c + \frac{(i-1) \cdot v}{k-1}$ is used because he wants to win against an expert of lower type when $C = c$.

The reason that bidders use only these three bids in equilibrium is that, assuming that the opponent also uses the same strategy in equilibrium, every other bid is dominated by at least one of these three. The intuition is as follows. A bid $y \in [0, cq)$ will lose against all the bids in $[cq, +\infty)$ and will win only against some bids in $[0, v]$. But a bid in $[0, v]$ means that either the opponent is an expert and $C = 0$, in which case we want to win against bids in $[0, v_i]$ and lose against bids in $[v_{i+1}, v]$, or the opponent is a nonexpert who happened to underbid, in which case we want to win all the time, something achieved by the bid of $cq + v_i$. So, depending on the parameters of the model, y is dominated by either v_{i+1} or $cq + v_i$. When these two give the same payoff, i.e. when the nonexpert is mixing between the two, y is dominated by both.

Similarly, a bid $y \in [cq, c)$ will lose against all the bids in $[c, +\infty)$, will win against all bids in $[0, v]$, and will win against some bids in $[cq, cq + v]$. But a bid in $[cq, cq + v]$ means that the opponent is a nonexpert, in which case we want to win against all bids in $[cq, cq + v_{i-1}]$ and lose against bids in $[cq + v_{i+1}, cq + v]$. So, y is dominated by the bid $cq + v_i$.

Finally, a bid $y \in [c, +\infty)$ will win against all bids in $[0, cq+v]$, and will win against some bids in $[c, +\infty)$. But a bid in $[c, +\infty)$ means that either the opponent is an expert and $C = c$, in which case we want to win against bids in $[c, c+v_{i-1}]$ and lose against bids in $[c+v_i, v]$, or the opponent is a nonexpert who happened to overbid, in which case we want to lose all the time, something achieved by the bid of $cq+v_i$. So, depending on the parameters of the model, y is dominated by either $c+v_{i-1}$ or $cq+v_i$. When these two give the same payoff, i.e. when the nonexpert is mixing between the two, y is dominated by both.

Suppose that the nonexperts of type i will bid $\frac{(i+1)\cdot v}{k-1}$ with probability $\theta_{i,1}$, $c \cdot q + \frac{i \cdot v}{k-1}$ with probability $\theta_{i,2}$, and $c + \frac{(i-1)\cdot v}{k-1}$ with probability $\theta_{i,3}$, where $\theta_{i,1} + \theta_{i,2} + \theta_{i,3} = 1$. It holds that $\theta_{0,3} = 0$ and $\theta_{k-1,1} = 0$.

The expected payoff of a nonexpert of type $i < k-1$ when he bids $\frac{(i+1)\cdot v}{k-1}$ is

$$\begin{aligned}\Phi_i =&(1-p)\left(\sum_{j=0}^{i-1} r_j\theta_{j,1}\left(cq+\frac{iv}{k-1}-\frac{(j+1)v}{k-1}\right)+\frac{1}{2}r_i\theta_{i,1}\left(cq+\frac{iv}{k-1}-\frac{(i+1)v}{k-1}\right)\right)\\&+p(1-q)\sum_{j=0}^{i}\frac{v(i-j)r_j}{k-1}.\end{aligned}$$

The expected payoff of a nonexpert of type i when he bids $c \cdot q + \frac{i\cdot v}{k-1}$ is

$$\begin{aligned}\Psi_i =&(1-p)\left(\sum_{j=0}^{k-1} r_j\theta_{j,1}\left(cq+\frac{iv}{k-1}-\frac{(j+1)v}{k-1}\right)+\sum_{j=0}^{i-1} r_j\theta_{j,2}\left(\frac{iv}{k-1}-\frac{jv}{k-1}\right)\right)\\&+p(1-q)\sum_{j=0}^{k-1}\frac{v(i-j)r_j}{k-1}.\end{aligned}$$

The expected payoff of a nonexpert of type $i > 0$ when he bids $c + \frac{(i-1)\cdot v}{k-1}$ is

$$\begin{aligned}\Omega_i =&(1-p)\left(\sum_{j=0}^{k-1} r_j\theta_{j,1}\left(cq+\frac{iv}{k-1}-\frac{(j+1)v}{k-1}\right)+\sum_{j=0}^{i-1} r_j\theta_{j,3}\left(cq-c+\frac{iv}{k-1}-\frac{(j-1)v}{k-1}\right)\right.\\&\left.+\sum_{j=0}^{k-1} r_j\theta_{j,2}\left(\frac{iv}{k-1}-\frac{jv}{k-1}\right)+\frac{1}{2}r_i\theta_{i,3}\left(cq-c+\frac{iv}{k-1}-\frac{(i-1)v}{k-1}\right)\right)\\&+p\left((1-q)\sum_{j=0}^{k-1}\frac{v(i-j)r_j}{k-1}+q\sum_{j=0}^{i-1} r_j\left(\frac{iv}{k-1}-\frac{jv}{k-1}\right)\right).\end{aligned}$$

Consider the case where $\theta_{i,1} > 0$ for every $i < k-1$, $\theta_{i,2} > 0$ for every i, and $\theta_{i,3} > 0$ for every $i > 0$. In other words, all types of nonexperts are mixing between all their potential bids. This case corresponds to the equilibrium in case 1 of Lemma 1.2. To find all the probabilities t, we need to solve the system

$$\begin{aligned}
&\Phi_i = \Psi_i, \quad \text{for } i \in \{0, 1, \ldots, k-2\} \\
&\Psi_i = \Omega_i, \quad \text{for } i \in \{1, 2, \ldots, k-1\} \\
&\theta_{0,3} = 0 \\
&\theta_{k-1,1} = 0 \\
&\sum_{m=1}^{3} \theta_{i,m} = 1, \quad \text{for } i \in \{0, 1, \ldots, k-1\}.
\end{aligned}$$

The solution to this system for $k = 2$ is

$$\begin{aligned}
\theta_{0,1} &= \frac{2p(q-1)r_1 v}{(p-1)r_0(cq-v)} \\
\theta_{0,2} &= 1 - \frac{2p(q-1)r_1 v}{(p-1)r_0(cq-v)} \\
\theta_{0,3} &= 0 \\
\theta_{1,1} &= 0 \\
\theta_{1,2} &= 1 - \frac{2pqr_0 v}{(p-1)r_1(c(q-1)+v)} \\
\theta_{1,3} &= \frac{2pqr_0 v}{(p-1)r_1(c(q-1)+v)},
\end{aligned}$$

which is the same as our solution in the main body of the paper ($\theta_{1,2} = a$ and $\theta_{0,1} = g$).

The solution to the system for $k = 3$ is

$$\theta_{0,1} = \frac{2v(2cq(p(q-1)r_1 + 2(p-1)r_0) + v(-p(q-1)(r_1 - 2r_2) - 4(p-1)r_0))}{(p-1)r_0\,(4c^2q^2 + 4cqv - 7v^2)}$$

$$\theta_{0,2} = \frac{(2cq-v)((p-1)r_0(2cq-v) - 2p(q-1)r_1 v) - 4p(q-1)r_2 v^2}{(p-1)r_0\,(4c^2q^2 + 4cqv - 7v^2)}$$

$$\theta_{0,3} = 0$$

$$\theta_{1,1} = \frac{2v(2cq(p(q-1)r_2 - (p-1)r_0) + v(p(q-1)(2r_1 + 3r_2) + (p-1)r_0))}{(p-1)r_1\,(4c^2q^2 + 4cqv - 7v^2)}$$

$$\begin{aligned}
\theta_{1,2} = \Big(&16c^4(p-1)(q-1)^2q^2r_1 - 16c^3(q-1)qv\Big(r_2(pq^2 - (p+1)q + p) \\
&+ r_0(p((q-1)q+1) + q - 1) + (p-1)r_1\Big) \\
&+ 4c^2v^2\Big(2r_0(p(q((q-1)q+4) - 1) + 3q^2 - 4q + 1) \\
&+ r_1(p(-6(q-1)q - 3) + 18(q-1)q + 7) - 2r_2(p(q((q-2)q+5) - 3) + q(2-3q))\Big) \\
&+ 4cv^3\Big(r_0(p(q(13q-12) - 2) + 5q + 2) + r_1(p(8(q-1)q - 3) + 7) \\
&+ r_2(p(q(13q-14) - 1) - 5q + 7)\Big) + 7v^4\Big(2r_0(-3pq + p - 1) \\
&+ 2r_2(3pq - 2p - 1) + (3p-7)r_1\Big)\Big)
\end{aligned}$$

$$\bigg/ \left((p-1)r_1 \left(4c^2(q-1)^2 - 4c(q-1)v - 7v^2\right) \left(4c^2q^2 + 4cqv - 7v^2\right) \right)$$

$$\theta_{1,3} = \frac{2(p-1)r_2v(2c(q-1)+v) - 2pqv((3r_0+2r_1)v - 2c(q-1)r_0)}{(p-1)r_1\left(4c^2(q-1)^2 - 4c(q-1)v - 7v^2\right)}$$

$$\theta_{2,1} = 0$$

$$\theta_{2,2} = \frac{(p-1)r_2(2c(q-1)+v)^2 + 2pqv(2r_0v - r_1(2c(q-1)+v))}{(p-1)r_2\left(4c^2(q-1)^2 - 4c(q-1)v - 7v^2\right)}$$

$$\theta_{2,3} = \frac{2pqv(r_1(2c(q-1)+v) - 2r_0v) - 8(p-1)r_2v(c(q-1)+v)}{(p-1)r_2\left(4c^2(q-1)^2 - 4c(q-1)v - 7v^2\right)}.$$

Even though we can analytically solve the system for larger values of k, the closed-form solution does not have any meaningful pattern. In the following example, we numerically solve the system for the case of $k = 20$, which would be an approximation of uniform continuous distribution of v.

Figure 1.B.2 shows a plot of the three mixing probabilities as functions of the private value of a nonexpert for $k = 20$ and small v. We can see that as the private value increases, the probability of underbidding decreases and the probability of overbidding increases.

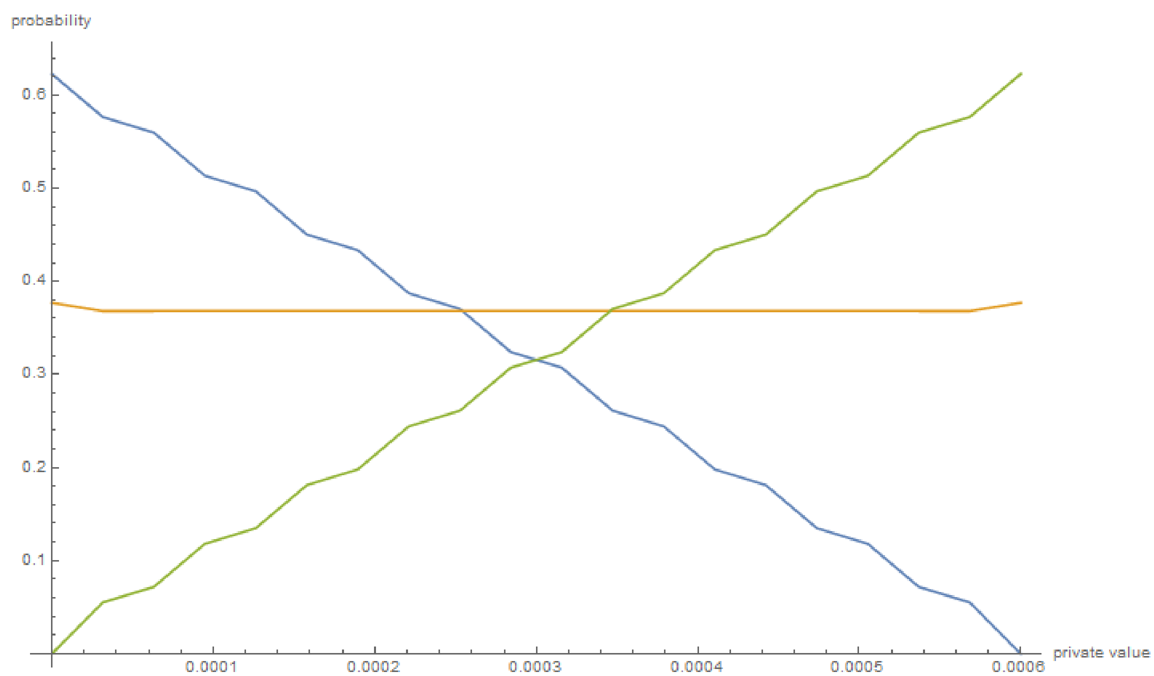


Figure 1.B.2: Mixing probabilities as a function of the private value. Blue is for underbidding, orange for bidding the expected valuation, and green for overbidding. The plot is for $k = 20$, $r_i = 1/k$ for $i \in \{0, \dots, k-1\}$, $c = 1$, $q = 0.5$, $p = 0.999$, and $v = 0.0006$.

As v increases, we will get different equilibria where some types of nonexperts don't mix between all their potential bids, i.e., some $\theta_{i,1}$'s become 0. A complete analysis of the equilibrium is beyond the scope of this paper as the number of cases in equilibrium analysis grows exponentially in k as v increases; however, the general pattern is that as v increases, nonexperts bid more aggressively. This is consistent with our findings in the main body of the paper.

1.B.5 Signaling Using Closing Format: Hard vs. Soft Close

In this section, we consider a situation in which the platform lets the sellers decide whether to sell in an auction with hard-close or soft-close format. We call the seller of an item with high common value a high-type seller, and the seller of an item with low common value a low-type seller. A seller is high-type with probability q where q is common knowledge. A seller naturally knows his own type; experts also know the seller's type (since they know the common value of items being offered). But nonexperts do not know the seller's type. We investigate whether a seller can signal his type using the closing format (soft versus hard). In particular, we derive conditions for existence of a separating equilibrium.

A seller sets his closing format F (soft or hard). For a format F, we assume that all nonexperts have the same belief about a seller who uses F. In general, nonexperts' belief about a format is the probability that they think a seller using that format is high-type. However, since we only consider pure strategy Nash equilibria of the game, nonexperts' belief about a format is limited to three possibilities: Low (L), High (H), and Unknown (X). In belief L, nonexperts believe that a seller using format F is always a low-type seller. In belief H, nonexperts believe that a seller using format F is always a high-type seller. Finally, in belief X, nonexperts cannot infer anything about the seller's type and believe that the seller is high-type with probability q.

Nonexperts have beliefs about each format F. In equilibrium, the beliefs must be consistent with sellers' strategies. In particular, if both types of sellers use the same format in (a pooling) equilibrium, nonexperts' belief for that format must be X. If the two types of sellers use different formats in (a separating) equilibrium, nonexperts' belief for the format used by the low-type seller must be L and for the format used by the high-type seller must be H. Furthermore, in an equilibrium, given the nonexperts' beliefs, sellers should not be able to benefit from changing their strategies.

We use the following notation to explain the results of this section: Let $\pi_T^B(F)$, where $T \in \{L, H\}$ and $B \in \{L, H, X\}$ denote the expected profit of a seller who uses mechanism $F \in \{soft, hard\}$ and has type T, and nonexperts believe has type B.

Lemma 1.6. *For any $B \in \{L, H\}$ and any $T \in \{L, H\}$ we have $\pi_T^B(soft) = \pi_T^B(hard)$. In other words, if nonexperts have no uncertainty about the type of the seller ($B \neq X$), soft-close and hard-close formats both lead to the same revenue for the seller (no matter what type the seller is).*

Proof. Note that when nonexperts have no uncertainty about the type of the seller, they do not infer anything from other bidders' bids, and do not update their expected value. The auction reduces to a full-information second price auction in this case. □

The seller's revenue, in each case, is given by

$$\pi_H^H(soft) = \pi_H^H(hard) = c + r^2 v;$$

$$\pi_L^L(soft) = \pi_L^L(hard) = r^2 v;$$

$$\pi_H^L(soft) = \pi_H^L(hard) = \begin{cases} cp^2 + rv(2(1-p)p(1-r) + r) & \text{if } v \leq c, \\ cp\left(2(1-p)r + p - 2(1-p)r^2\right) + r^2 v & \text{if } v > c; \end{cases}$$

$$\pi_L^H(soft) = \pi_L^H(hard) = \begin{cases} c(1-p)^2 + rv(2(1-p)p(1-r) + r) & \text{if } v \leq c, \\ c(1-p)(p(2(1-r)r - 1) + 1) + r^2 v & \text{if } v > c. \end{cases}$$

Lemma 1.7. *No separating equilibrium exists.*

Proof. Assume for sake of contradiction that there is a separating equilibrium in which the low-type uses format F and the high-type uses format F'. Note that, using the above expressions, we have $\pi_L^L(F) < \pi_L^H(F) = \pi_L^H(F')$ for any F and F'. Therefore, the low-type benefits from mimicking, contradicting the equilibrium condition. □

Depending on out-of-equilibrium beliefs, the game could have multiple pooling equilibria (both soft- and hard-close). We can show that soft-close is always a pooling equilibrium. Furthermore, for regions in which hard-close provides higher expected revenue for the high-type seller, as shown in Figure 1.A.1, hard-close is also a pooling equilibrium. The intuitive criterion is not sufficient for refining the equilibrium set to a unique equilibrium. However, we can show that only soft-close pooling equilibrium can survive the D1 criterion refinement. Intuitively, the high-type always gains more (loses less) than low-type by deviating to soft-close format in a hard-close pooling equilibrium. Therefore, out-of-equilibrium beliefs on soft-close auction, subject to D1 requirement, is high. This makes the deviation to soft-close always profitable (in a hypothetical hard-close equilibrium). Therefore, hard-close pooling equilibrium cannot survive D1 criterion refinement. This is formally proved in the following lemma.

Lemma 1.8. *A hard-close pooling equilibrium cannot survive D1 criterion refinement.*

Proof. Assume for sake of contradiction that there is a hard-close pooling equilibrium. Let z be the buyer's belief, the probability that the seller is high-type, on observing the soft-close format. We show that for any $z < 1$, if a low-type seller weakly benefits from deviating to soft-close, a high-type seller strictly benefits from deviating. Then, according to D1 criterion, this implies that out-of-equilibrium belief on soft-close has to be high. Therefore, hard-close cannot be an equilibrium.

Assume for sake of contradiction that there is a z for which a low-type seller weakly benefits from deviating to soft-close, but a high-type seller does not strictly benefit from deviating to soft-close. First, note that if both buyers are nonexperts, the equilibrium outcome is not affected by the type of the seller. In other words, both types of sellers would have the same revenue in each closing format. Similarly, if both buyers are experts, the equilibrium outcome is not affected by the closing format. Therefore, to compare the benefit of deviation (for sellers), we can assume that the buyers have different levels of expertise: an expert and a nonexpert.

If the expert is low-value, then the revenue is zero in both hard-close and soft-close formats for the low-type seller. The revenue for the high-type seller is always greater than or equal in soft-close (always c) than in hard-close (at most c, depending on the value of z and whether the nonexpert is using an aggressive strategy or not).

Finally, consider the case that the expert is high-value. First, assume that the nonexpert is also high-value. In this case, the revenue of soft-close for high-type seller is always c and for low-type seller is always v. The revenue of hard-close for the low-type seller is always v. Therefore, the low-type cannot prefer hard-close to soft-close in this sub-case. Next, assume that the nonexpert is low-value. In this case, the revenue of the auction in both soft-close and hard-close cases is determined by the bid of the nonexpert. A low-value nonexpert has the exact same strategy in soft-close and hard-close formats. Furthermore, this strategy does not depend on whether the common value is high or low. Therefore, low-type and high-type sellers have the same revenue in the hard-close format and in the soft-close format. As shown, there is no case (for any z) in which a low-type seller benefits from deviating to soft-close while a high-type seller does not. Furthermore, it is easy to see that there are cases in which the high-type seller strictly benefits from this deviation. Therefore, according to D1 criterion, buyers' out-of-equilibrium belief on soft-close auction has to be high.

Given that, in a hypothetical hard-close equilibrium, buyers' belief on soft-close is high, sellers always benefit from deviating to soft-close. Therefore, hard-close cannot be a pooling equilibrium. □

Finally, note that a soft-close pooling equilibrium always survives D1 criterion refinement. This is because whenever the high-type seller benefits from deviating to hard-close, the low-type seller also benefits from deviating to hard-close (the proof is very similar to the proof of Lemma 1.8). Therefore, a soft-close pooling equilibrium in which out-of-equilibrium belief on hard-close is low survives D1 criterion refinement.

Given that the only pure strategy equilibrium that survives D1 criterion refinement is a soft-close pooling equilibrium, we show that, compared to the case where the platform decides closing format, a low-type seller and the platform are both (weakly) worse off if the closing format decision is left to the sellers. A high-type seller may be worse off or better off, depending on other parameters, as shown in Figure 1.A.1.

Chapter 2

Multi-Channel Attribution: The Blind Spot of Online Advertising[1]

In this chapter, we study the problem of attributing credit for customer acquisition to different components of a digital marketing campaign using an analytical model. We investigate attribution contracts through which an advertiser tries to incentivize two publishers that affect customer acquisition. We situate such contracts in a two-stage marketing funnel, where the publishers should coordinate their efforts to drive conversions.

First, we analyze the popular class of multi-touch contracts where the principal splits the attribution among publishers using fixed weights depending on their position. Our first result shows the following counterintuitive property of optimal multi-touch contracts: higher credit is given to the portion of the funnel where the existing baseline conversion rate is higher. Next, we show that social welfare maximizing contracts can sometimes have even higher conversion rate than optimal multi-touch contracts, highlighting a prisoners' dilemma effect in the equilibrium for the multi-touch contract. While multi-touch attribution is not globally optimal, there are linear contracts that "coordinate the funnel" to achieve optimal revenue. However, such optimal-revenue contracts require knowledge of the baseline conversion rates by the principal. When this information is not available, we propose a new class of 'reinforcement' contracts and show that for a large range of model parameters these contracts yield better revenue than multi-touch.

2.1 Introduction

The last decade has seen a large shift of advertising effort from offline to online channels. Internet advertising revenue is projected to overtake TV advertising for the first time in 2017.[2] The key benefit of the online medium is the accountability provided by the user clicks and the cookie trails left by their visits to various advertising venues. This fine-grained view of the user's journey through the decision funnel along with the specific advertising actions

[1]Based on joint work with Vibhanshu Abhishek and R. Ravi

[2]http://www.pwc.com/us/outlook (accessed March 2017).

they are exposed to (banner and display ads, video ads, text ads after search, emails) provides a unique opportunity to solve the traditional marketing mix problem in a very user-specific way. The key to arriving at optimal resource allocations across the channels is to determine the response model of how each of the interventions affects the decision-making journey of the customer. This is commonly phrased as an **attribution** problem of how the credit for a digital conversion should be split among the various advertising actions. Attribution allows an advertiser to determine the impact of each ad-type so that the effectiveness of different types of ad activities can be taken into consideration while deciding how to split the advertising budget.

The problem of attribution is not new. It arises in traditional advertising channels like television and print as well, where advertisers have resorted to marketing mix models using aggregate data (Naik et al., 2005; Ansari et al., 1995; Ramaswamy et al., 1993). However, online advertising offer a unique opportunity to address the attribution problem as advertisers have disaggregate individual-level data which were not previously available (Goel, 2014). Disaggregate data offer the possibility of determining the effectiveness of an ad on an individual customer at a specific time. Although better data has improved the accountability and performance in online advertising, several advertisers still use simplistic approaches like last/first touch attribution that might be suboptimal under a variety of conditions (PWC, 2014; Abhishek et al., 2016). The rapidly growing size of the industry and concerns raised by the advertisers have lead to tremendous recent focus on attribution. Companies like Google, Marketo, and Datalogix have designed and offered several new algorithmic attribution techniques in the last few years. Google (2017a) alone offers a variety of such models including last-touch and first-touch (where the last and first publisher gets the full credit respectively), as well as other weighted models including options for weighting based on fixed weights, time decay or position. Google (2017b) also offers an alternate data-driven attribution method that is based on the Shapley value of each publisher.

At the same time, many researchers have proposed empirical models of attribution. At its core, the attribution rule determines the payment received by publishers for showing an advertiser's ad.[3] Most of the existing literature has assumed that the publishers are not strategic (with the exception of Berman (2015)) and that their actions are not affected by the attribution methodology used. However, publishers (e.g. Facebook or Google) have to exert a considerable amount of effort (such as investments in technology) to match an ad impression with the right customer, and this effort level is affected by the incentives. Unfortunately, a typical advertiser cannot observe the effort exerted by a single publisher in the funnel while the final conversion is based on the total effort by all publishers in it. This might lead to free-riding where a publisher aims to benefit from the effort exerted by another publisher in the pipeline, and creates an opportunity for moral hazard (Holmstrom, 1982). In addition, the publishers are much better informed about the consumers as they observe them in many different contexts as compared to the advertiser. This information asymmetry prevents the advertiser from estimating the effectiveness of a publisher's marketing action and can lead to adverse selection. As an example of such an inefficiency, Abhishek et al. (2016) show that

[3]In this paper, we denote publisher as an entity that is responsible for matching ads to consumers and the delivery of ads. In most cases, this role would be fulfilled by an ad network such ad Double Click of Yahoo Display Network, a search engine like Google or a large publishers like Facebook or Snap. We use the term publisher to be consistent with the prior literature (e.g. Berman (2015))

once display publishers have determined which consumers are most likely to convert, they flood them with ads. This crowds out more effective publishers and drives revenues from them to less effective publishers. Both the moral hazard and adverse selection problems create a misalignment of incentives between the advertiser and the publishers resulting in a loss of efficiency in online advertising.

The misalignment in incentives is manifested in many ways. Ad fraud has become a major concern and publishers charge advertisers for ads that consumers never see. In fact, methodologies like last-touch give credit to publishers even if the consumer would have converted without seeing an ad. There is a substantial increase in low-quality ad inventory, that do not lead to any meaningful outcomes such as conversions (Scott, 2016). A recent Economist (2016) article shows that ad fraud will cost advertisers US$ 7 billion this year and is growing rapidly. Even though advertisers are aware of these issues, they are not able to address them due to the lack of appropriate data and the complexity of the online advertising industry. This complication is exacerbated even further due to the multiple touch-points spanning several publishers that jointly affect the consumer decision-making funnel. The multiplicity of publishers create a lack of accountability (Economist, 2016), which affects the advertisers and ultimately the entire advertising industry adversely. Although advancement in attribution methodologies have made advertising more efficient,[4] they have not eliminated the moral hazard or adverse selection issue completely. One way to eliminate or reduce the moral hazard issue is using newer attribution methodologies, e.g. Berman (2015) presents a Shapley value based attribution scheme that performs better than last-touch attribution when the publishers' ads are strategic complements and the uncertainty in consumer behavior is low.

In this paper, we propose a simplified two-stage model of the purchase funnel to determine the most appropriate attribution methodology. Prior literature (Mulpuru, 2011; Court et al., 2009; Bettman et al., 1998) shows that consumers move through different stages before they purchase a product. Consequently, in our analysis, the two stages considered are awareness and consideration. Incorporating the temporal dynamics is not only a more natural approach to addressing the attribution problem but also leads to interesting new results as opposed to more static models considered in prior literature. We consider two distinct publishers that are responsible for generating ad impressions on behalf of the advertiser and jointly drive consumers towards conversion. In the basic model, we consider that one publisher can create awareness and the other one can drive conversions.[5]

After reviewing related work and background in Sections 2.2 and 2.3, we define our model in Section 2.4. In Section 2.5, we first analyze a linear attribution rule that splits a fraction of the marketing dollar as f and $1-f$ between the two publishers (for a conversion). This rule resembles the commonly used multi-touch attribution rules, such as first- and last-touch. We term the resulting contracts as f-contracts, which in turn determine the efforts exerted by the publishers in the co-production process. One might expect the optimal f to give more credit to the stage where the baseline conversion rate is lower. We show the

[4] For example, "According to Forrester Research, B2B companies are seeing an average of 15 to 18 percent lift in revenue as a result of implementing a closed-loop attribution system." (`https://www.clickz.com/lift-off-a-new-study-on-attribution-and-revenue/41687/`, accessed March 2017).

[5] In an extension of the basic model, we assume that both publishers compete for both awareness and conversions.

counter-intuitive result that, all else being equal, an advertiser using an optimal f-contract should give *more* credit to a publisher with higher baseline probability of advancing (down the purchase funnel) in its stage. We arrive at this counter-intuitive result because of the way multi-touch contracts compensate publishers. There is a complementarity between the baseline rate of one stage and the effort exerted by the publisher in the other stage since the final conversion is a co-production process. In the presence of this effect, to provide an incentive for the publisher in this stage to exert optimal effort, the optimal f-contracts gives this publisher more credit if its baseline rate is high.

Next, we show that social welfare maximizing contracts can lead to have even higher conversion rates than optimal multi-touch contracts, highlighting a prisoners' dilemma effect in the equilibrium for the multi-touch contract. Our result shows that the optimal f-contract gives an inefficient equilibrium due to lack of coordination between the publishers. This result is unique in the attribution literature as we show that the general class of f-contracts, which are commonly used in practice, can result in inefficiencies in the effort exerted by the different publishers.[6]

In Section 2.6, we explore the broader design space of linear contracts. We show that there are optimal linear contracts that can "coordinate the marketing channels" to achieve optimal revenue. However, these optimal contracts suffer from the same problem as optimal multi-touch contracts in that they require full knowledge of the baseline conversion rates by the principal (advertiser). For this reason we propose a new class of 'reinforcement' contracts. These contracts perform significantly better over a wide range of parameters than other multi-touch contracts while not relying on the knowledge of the baselines. Finally, in Section 2.7, we examine several extensions of our underlying models and show that the main findings still hold under a variety of circumstances.

This paper addresses an important gap in the attribution literature, namely the strategic decision by the publishers in a dynamic purchase funnel. As discussed earlier, incorporation of the dynamic nature of the conversion process leads to new and interesting results. This paper has also several managerial implications about designing of multi-channel advertising contracts under information asymmetry and uncertainty. Advertisers can use the contracts outlined in the paper to increase the effectiveness of multi-channel advertising.

2.2 Literature Review

Recent years have seen a tremendous amount of academic interest in the attribution problem given the importance to the industry. Here, we discuss the different streams of literature that are relevant to our research.

We start off with some of the empirical work on attribution from the marketing literature. Shao and Li (2011) propose two multi-touch attribution models, a bagged logistic regression model and a probabilistic model, and they apply these two approaches to a real-world dataset. Jordan et al. (2011) explain why current attribution methods are inefficient and they find an optimal ad allocation and payment scheme in a model they developed. Dalessandro

[6] Given the inefficiency, a related issue is how bad an f-contract can be in terms of social welfare. Using a concept called the Price of Anarchy (Koutsoupias and Papadimitriou, 1999; Roughgarden and Tardos, 2002), we show that this multiplicative ratio is bounded by $\frac{4}{3}$.

et al. (2012) propose an attribution methodology based on a casual estimation problem that uses the concept of Shapley Value. Anderl et al. (2013) introduce a graph-based framework for attribution using Markov models and test it in real-world data-sets. Li and Kannan (2014) propose a measurement model for attributing conversions to different channels and find that the relative contributions of these channels are different from those estimated by traditional metrics like last-touch. Xu et al. (2014) develop a multivariate point process that captures the dynamic interactions among ad clicks and find that even though display ads may have low direct effect on conversions, they have also an indirect effect by stimulating subsequent visits through other ad formats. Abhishek et al. (2016) use a hidden Markov model for consumer behavior to find that different channels and types of ads affect consumers in different stages in the purchase funnel.

One of the first analytical papers on attribution was by Berman (2016). Berman uses an analytical model with two publishers and two advertisers that involves externalities between the publishers and uncertainty about consumer visit order. Using his model, he shows that bidding truthfully in ad auctions is not an equilibrium for the advertisers. He also shows that last-touch attribution results in lower profits for advertisers compared to not using attribution, while an attribution based on Shapley value can result in higher profits when conversion rates are low.

Our work differs from Berman's work in that we explicitly model the marketing funnel and we take into account customer's microtrails. We believe that the nature of some websites make them more or less likely to be at a certain point in a customer's trail. For example, when a customer searches for a product he is already interested in, we can assume that he is in the final stage of consideration, and therefore an ad in the search engine is more likely to be the last-touch point before purchase. This affects the behavior of publishers who publish ads for different stages in the funnel, and as a result it is important especially for multi-touch contracts. Another difference is that we model a wider range of attribution rules, where last-touch and Shapley value are special cases, and we examine the advertiser's problem to determine the optimal attribution. However, we assume a functional form for the relation between the effort of the publisher and the increase in conversion rate, while Berman models the process of price setting at each stage more realistically via a second-price auction.

The attribution problem is also related to the team production literature in contract theory. There is a big literature on the topic, but to mention just a few, Holmstrom (1982) studies the problem of moral hazard in team compensation and how it can be dealt with by breaking the budget-balancing constraint. Eswaran and Kotwal (1984) followed by explaining how not balancing the budget can result in a new source of moral hazard. Holmstrom and Milgrom (1987) study compensation schemes for incentivizing agents. Dearden and Lilien (1990) consider a problem in which the firm learns over time. McAfee and McMillan (1991) study the interaction between moral hazard and adverse selection in a team model.

Besides the context, another key difference to our work is that we explore the multiplicative effect of agents' efforts, which is a result of the marketing funnel. In other words, instead of all agents putting efforts together that adds to a total effort, they put efforts in stages in a way that the result of the effort an agent puts in a stage is affected by what happened in other stages. Incorporating this difference, which is a key to the marketing funnel approach, leads to some interesting results.

2.3 Background

We present some background on the purchase funnel, attribution models, and fairness considerations.

2.3.1 Purchase Funnel and Common Attribution Rules

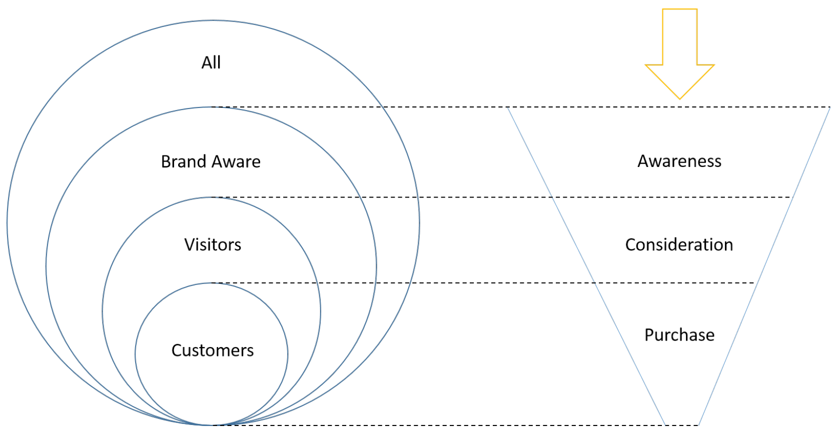


Figure 2.1: Purchase funnel.

In the online world, consumers are exposed daily to a number of different types of ads, e.g. display ads, search ads, affiliate ads and sponsored content. Before a consumer purchases a product, he is influenced by these ads as he moves through his decision making process, which has been commonly captured using the purchase funnel as shown in Figure 2.1. The purchase funnel captures the progression of individuals from being unaware about the firm to purchasing products and becoming the firm's customers. A fraction of the total population becomes aware of the firm and moves into the state of awareness. Some of these brand aware individuals might be further interested in purchasing products from the firm and move into the next consideration state. Finally, a small fraction of individuals that consider the product will eventually purchase it. Since each of these stages contains fewer number of consumers than the previous one, the progression is typically illustrated as a funnel. Consumers enter through the top of the funnel, pass through the different stages, and some of them convert. Each type of ad that a consumer is exposed to helps his move to a further stage on this path, and some ads are more effective in some stages than others (e.g. display ads in the awareness stage, and search ads in the consideration stage).

For advertisers using both stages of the funnel, it is important to have an attribution rule for eventual conversions. This rule will determine who gets credit every time there is a conversion, conditional on all the ads the user has previously seen. As an example, last-touch attribution is a widely used attribution model, where all the credit is assigned to the ad responsible for the last ad exposure before conversion. Last touch is very popular for its convenience and because it is easy to implement. However it might not be optimal, since it fails to take into account ads used to build awareness and interest to the consumer.

To address this issue, firms have started offering alternative models like first-touch (Google, 2017a), where the credit goes to the first ad the user was exposed to because it was the ad that made him enter the funnel. In even more general multi-touch models, the advertiser determines how to split the credit between all the ads in a consumer's trail. For example, he could give equal weight to all the ads in the trail or give more weight to the first and the last touch points and less in the middle.

2.3.2 Fairness and Shapley Value

One approach that has been gaining popularity in the advertising literature for deciding payment rules is fairness (Berman, 2015; Dalessandro et al., 2012; Abhishek et al., 2016). Fairness has also been gaining attention in the economics literature since its introduction by Rabin (1993) to examine game-theoretic problems. Defining fairness has been a challenging problem and researchers have used a set of axioms to delineate what is fair (van den Brink, 2002; Lan et al., 2010). Shapley (1953) proposed four such natural axioms and proved that there is a unique rule that satisfies them. We call this rule the *Shapley Value*, and the payoff of each player according to this rule is a weighted sum of his marginal contributions to every subset of players.[7] This approach, also referred to as the incremental approach, has been commonly adopted in the attribution literature (Abhishek et al., 2016; Berman, 2015) and practice (Google, 2017b). Shapley value is a great attribution rule if our goal is to achieve a fair result and if the attribution rule we use does not affect the performance or actions of the players. Now suppose that the effort of each player depends on the attribution rule we use, and our goal is to maximize the total output of the game. Even though Shapley value based attribution meets the fairness axioms, it might not be the most optimal for an advertiser. For example, if a publisher doesn't get enough credit with the Shapley value, he might put less effort in showing ads and the advertiser will end up with fewer conversions overall. It is possible that an alternative attribution scheme might give the publisher higher payouts so that he puts in more effort with a better overall result. How should we then choose an attribution rule that maximizes the total value of the game? Our paper tries to answer this question.

2.4 Model

2.4.1 Overview

Before purchasing a product, a consumer moves through two stages: awareness and consideration.[8] This model is based on the idea of a conversion funnel which is frequently used in marketing (Mulpuru, 2011; Court et al., 2009; Bettman et al., 1998). Every period, a new consumer arrives in the system and moves to the first stage (awareness).[9] A consumer in the first stage either moves to the second stage (consideration stage) with probability $f(a)$,

[7] For a formal definition, see Appendix 2.A.3.

[8] We analyze two stages for clarity of exposition, but out results continue to hold with more than two stages.

[9] The model can be generalized by including some stochastic waiting time in each stage, but this does not influence the results.

or leaves the system. The function $f(a)$ has as argument the advertising level a for the first stage. The ads in the first stage create awareness, e.g. display ads or sponsored content. Similarly, a consumer in the second stage either purchases the product with probability $g(c)$, or leaves the system. c is the advertising level for the second stage. Ads in the second stage are more transaction oriented and lead directly to conversion, e.g. search ads. Note that it is not necessary for consumers to see ads in either stage before they purchase as $f(0)$, $g(0)$, the baseline rates at each stage, can be strictly positive.

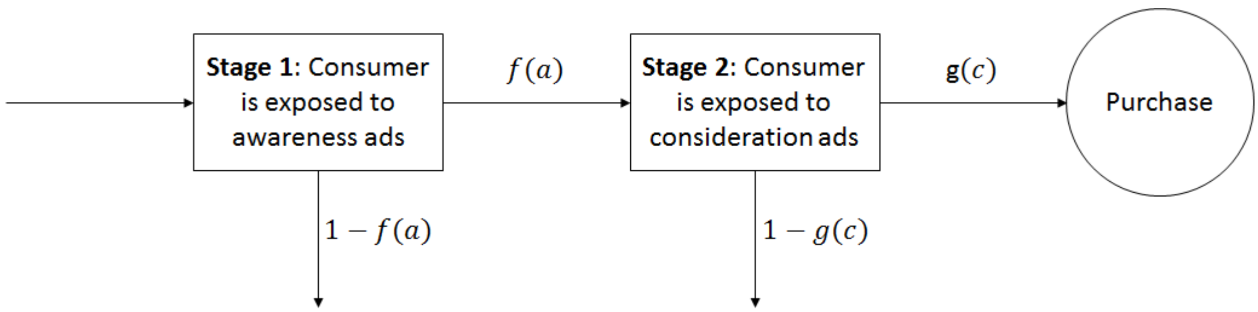


Figure 2.1: General representation of the model.

2.4.2 Consumer Model

Awareness Stage

For the customer to purchase a product, it should belong in the customer's consideration set. If the consumer is not aware of the product, then it is unlikely that he will eventually purchase the product. We assume that a fraction of consumers (≥ 0) might be aware of the product even in the absence of advertising. We represent the **baseline rate** of these consumers by $q_0 \in \left[0, \frac{1}{2}\right]$. Some consumers learn about the product because they have seen an ad from Publisher 1. The rate of these consumers is $a \in \left[0, \frac{1}{2}\right]$, where a is a decision variable for Publisher 1. In other words, we assume the functional form $f(a) = a + q_0$ for the function f (see also Figure 2.2). When a consumer learns about the product, they move to the consideration stage. If a consumer does not learn about the product in this first stage, which happens with probability $1 - q_0 - a$, they leave the system.[10]

We assume the convex functional form of $w \cdot a^2$ for the effort that Publisher 1 has to exert to result in additional conversion a in the first stage. The probability a captures how effective the publisher is in showing the ads to the relevant audience, while the convexity of the functional form models diminishing returns to effort by the publisher.[11]

[10]In the Appendix 2.A.2 we also consider a more general model of the funnel and show how it can be reduced to the simple model analyzed here.

[11]Note that the probabilities a and q_0 are at most $\frac{1}{2}$, to make sure that $f(a) \leq 1$. In reality, these probabilities are very small, so we don't lose anything by assuming the upper bound of $\frac{1}{2}$. Moreover, to make sure that a will not exceed this bound when we find the equilibria, we assume that w is sufficiently high. We will use the bound $w \geq 1$, which is sufficient for our model.

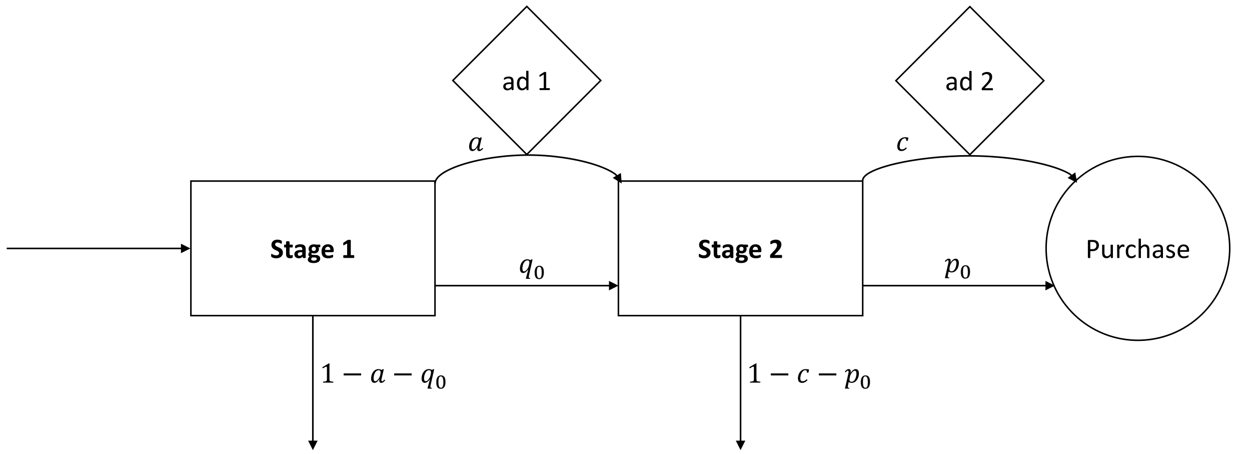


Figure 2.2: A more detailed representation of the model.

Consideration Stage

After consumers reach the consideration stage and are interested in the product, then we have the second stage of our model. Some consumers in that stage will decide to buy the product after seeing an ad from publisher 2. This will happen with probability $c \in \left[0, \frac{1}{2}\right]$, which is a decision variable for Publisher 2. Some other aware consumers will buy the product without seeing any ad and this will happen with probability $p_0 \in \left[0, \frac{1}{2}\right]$. In other words, we assume that the function g has the form $g(c) = c + p_0$. The probability that a consumer in the second stage will not buy the product in the end is $1 - p_0 - c$.

Publisher 2 can be considered as a website with content related to the product, where aware consumers go to look for more information before they decide if they will buy (e.g. a search engine, a site with reviews or comparisons of similar products). As before, we assume the convex functional form of $v \cdot c^2$ for the effort that Publisher 2 has to exert to result in additional conversion c in the second stage.[12]

2.4.3 Firm's Problem

In the absence of advertising, a fraction $q_0 \cdot p_0$ of consumers convert. With online advertising, the firm gains the ability to convert more consumers ($(q_0 + a) \cdot (p_0 + c)$) by showing them ads. The firm cannot advertise directly, but can use the advertising real estate provided by the publishers to reach its customers.

The firm can observe the consumers who decided to buy the product and all the ads they've seen prior to that (e.g., through the trail captured in browser cookies). It may not know the efforts the publishers put (that resulted in the additional rates a and c in the two stages), or the baseline probabilities q_0, p_0.[13] What the firm can infer by observing a large

[12]Similarly to the first stage, we assume that c and p_0 are at most $1/2$ and to enforce this upper bound of c in the equilibria, we assume that $v \geq 1$.

[13]The advertiser can potentially observe consumers who do not convert. However, as long as the advertiser does not observe *all* the consumers in the market, it cannot determine p_0, q_0 with certainty.

number of consumers are the percentages of people who followed each of the four possible paths before conversion. For example, it will know that an $a \cdot c$ fraction of consumers have seen an ad from the first publisher in the first stage and an ad from the second publisher in the second stage before they convert. Similarly, it will know that an $a \cdot p_0$ fraction of consumers have seen an ad from the first publisher in the first stage and no ad in the second stage before they convert, and so on (Figure 2.3).

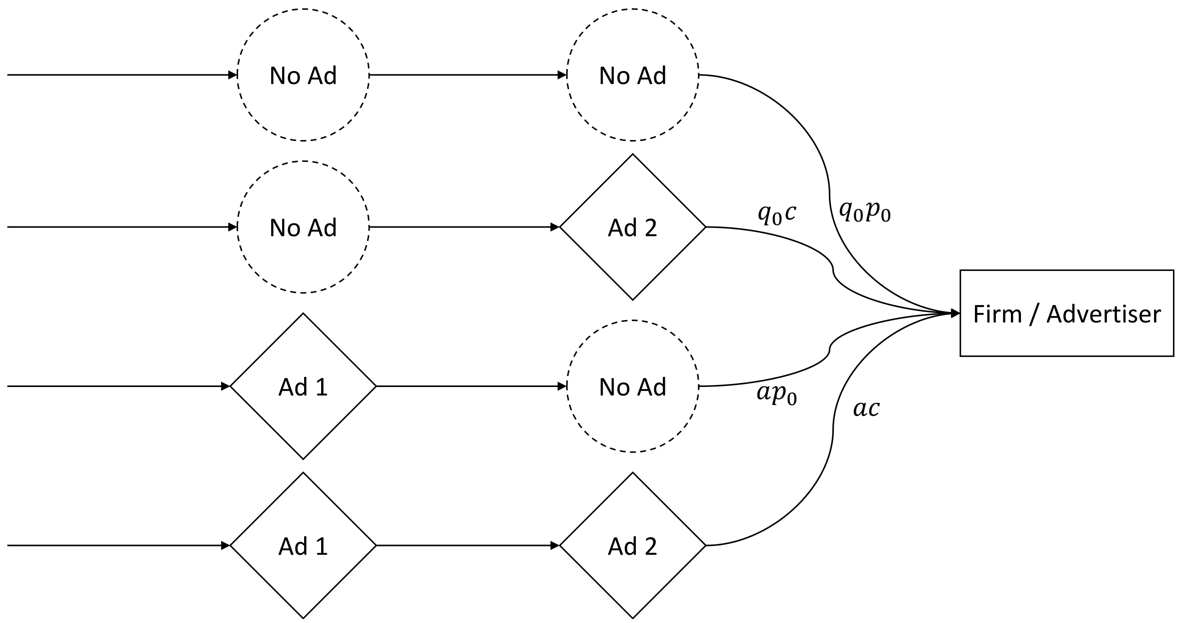


Figure 2.3: The four types of click-streams the firm can observe.

For every conversion, the firm wants to spend some fixed amount in advertising, which we normalize to \$1.[14] The question that we address is the optimal way to split this dollar between the two publishers in order to maximize the conversion rate $(a + q_0)(c + p_0)$.

2.4.4 Valid Contracts and Publishers' Problem

Any attribution rule the firm uses will lead to a contract with the publishers. The publishers try to maximize their profits by putting in the optimal amount of effort based on the contract they have with the firm. A contract is defined by two payment functions g_1, g_2 that satisfy the equality

$$g_1(a, q_0, c, p_0) + g_2(a, q_0, c, p_0) = (a + q_0)(c + p_0),$$

where g_1 is the payment to the first publisher and g_2 is the payment to the second publisher. Since $(a + q_0)(c + p_0)$ is the conversion rate and we assume that the firm wants to spend \$1 per conversion, the functions g_1 and g_2 specify how the firm should split their advertising budget between the two publishers.

[14]This limit on the spend can arise from the cost of the competing outside advertising options, e.g. print or television advertising.

Valid Contracts. The contracts that the firm can offer to the publishers must be based on the information that the firm knows. In this paper we will focus on the class of linear payment functions with respect to the four products q_0p_0, q_0c, ap_0, ac. This type of contracts have two advantages. First, they don't require the knowledge of the baseline rates p_0, q_0 or the publishers' efforts a, c to be implemented, but only knowledge of the percentages of people who followed each of the four possible paths before conversion (Figure 2.3).[15] Second, even if the baseline rates p_0, q_0 are known to the advertiser, attribution rules that depend linearly only on the products q_0p_0, q_0c, ap_0, ac can be applied in an online fashion. Every time there is a conversion, the advertiser can attribute one unit of credit among publishers based on the conversion path the customer followed. Note also that the widely used class of multi-touch attribution contracts is a special case of linear contracts.

The profit of the first publisher is given by

$$\pi_1 = g_1(a, q_0, c, p_0) - w \cdot a^2,$$

where $w \cdot a^2$ is the advertising cost in the first stage. Publisher 1 decides a to maximize his profits. Similarly, the profit of the second publisher is given by

$$\pi_2 = g_2(a, q_0, c, p_0) - v \cdot c^2,$$

where $v \cdot c^2$ is the advertising cost in the second stage. Publisher 2 decides c to maximize his profits.

2.4.5 Benchmarks

Next, we consider two benchmark models to compare with our main model. In both models, we assume that the firm and the publishers are integrated, i.e. the firm controls the advertising efforts. In the first model, the goal for the firm is to maximize the social welfare, i.e. the sum of profits of the two publishers. In the second model, the goal is to maximize the number of conversions.

Maximizing Social Welfare (Publishers' Optimal) In the first benchmark model, the firm will make all the decisions about advertising, and its goal is to maximize the social welfare. The social welfare is equal to the total revenue minus the total cost for the publishers. Therefore, the optimization problem for the firm is as follows.

$$\begin{aligned} \max_{a,c} \quad & (a + q_0)(c + p_0) - (wa^2 + vc^2) \\ & a, c \geq 0. \end{aligned}$$

Maximizing Conversion Rate (Firm's Optimal) There are two ways we could potentially model this benchmark case. One is to model it as an optimization problem with the conversion rate as the objective and no constraints. In other words, we don't care about the

[15] Note that if the firm can only observe the four products q_0p_0, q_0c, ap_0, ac, it cannot infer the individual efforts. This is because for any solution (q_0, a, p_0, c) that satisfy these products, there are other solutions of the form $\left(q_0x, ax, \frac{p_0}{x}, \frac{c}{x}\right)$, for $x > 0$ that satisfy them too.

cost of advertising effort, but we want to find the optimal effort level that maximizes the number of conversions. However, this would be unrealistic, simply because we can always achieve a conversion rate of 1 in this optimization problem.

The more appropriate benchmark is achieved by including a constraint on the cost of advertising effort. Note that in the main model, the firm can spend \$1 for every conversion, which goes to the two publishers. Therefore, the appropriate benchmark is to determine the optimal conversion rate given that the cost of advertising effort is exactly the number of conversions (multiplied by unit revenue per conversion). In other words, the optimization problem for the firm is as follows.

$$\begin{aligned} \max_{a,c} \quad & (a+q_0)(c+p_0) \\ & (a+q_0)(c+p_0) = wa^2 + vc^2 \\ & a, c \geq 0. \end{aligned}$$

We denote by a^*, c^* the efforts in the optimal solution of this problem.

2.5 Multi-Touch Attribution

2.5.1 Definition of f-contract

We start by considering simple contracts that split conversion credit among the touch points of a consumer's trail. The canonical form of such a contract, that we term an f-contract, is summarized in the following table.

Ad in the awareness stage	Ad in the conversion stage	Credit to Publisher 1	Credit to Publisher 2
no	no	$\frac{1}{2}$	$\frac{1}{2}$
no	yes	0	1
yes	no	1	0
yes	yes	f	$1-f$

The parameter $f \in [0,1]$ is some value determined by the firm or externally. Note that in the table above, in the case that the consumer sees no ad, we split the dollar equally to the two publishers. This is an arbitrary choice made to be consistent with the assumption that the firm always pays out a dollar for every conversion. However, since this is just the same constant amount each publisher gets, its value does not affect the equilibrium behavior of the publishers. In other words, we could also assume that in case of no ad, both publishers get 0, without any change in our results.

In an f-contract, the profit of the publishers are as follows,

$$\begin{aligned} \pi_1 &= \frac{1}{2}q_0p_0 + ap_0 + fac - wa^2, \\ \pi_2 &= \frac{1}{2}q_0p_0 + q_0c + (1-f)ac - vc^2. \end{aligned} \tag{2.1}$$

Some examples of contracts of this form are the last-touch ($f = 0$), where all the credit goes to the last ad a consumer had seen prior to the conversion, and the first-touch ($f = 1$)

where all the credit goes to the first ad. It is also interesting to note that for $f = \frac{1}{2}$, we get the Shapley value attribution of this model. In other words, the f-contract captures a wide range of attribution models commonly used in practice (Google, 2017a).

If we solve for the equilibrium under an f-contract, the equilibrium efforts exerted by the publishers are as follows,

$$a(f) = \frac{fq_0 + 2vp_0}{4vw - f(1-f)},$$
$$c(f) = \frac{(1-f)p_0 + 2wq_0}{4vw - f(1-f)},$$

while the conversion rate as a function of f is given by

$$r(f) = \frac{(f^2 q_0 + 2v(p_0 + 2wq_0))((1-f)^2 p_0 + 2w(q_0 + 2vp_0))}{(4vw - f(1-f))^2}.$$

2.5.2 Optimal f-contract

We can now investigate the properties of the optimal contract and show how different parameters affect the split of the advertising dollar between Publishers 1 and 2. We study the effect of the baseline conversion rates here and consider the effect of the publisher costs later in Section 2.7.1. We then investigate the social welfare properties of the equilibrium of the f-contract.

Effect of baseline conversion rates

Our first proposition shows how the optimal contract f^* varies with baseline rates q_0 and p_0. To motivate it, note that if the value of the baseline rate q_0 is high, there is already a high rate of intrinsic conversion in the first stage. This suggests that there is less value in increasing the incentive to the first publisher in the co-production process, and hence could lead to lowering f as q_0 increases. However, we find the opposite to be the case. (All proofs are provided in the appendix.)

Proposition 2.1. *Let $f^* = f^*(q_0, p_0)$ be the value of f in the optimal f-contract for the firm as a function of q_0 and p_0. Then, f^* is increasing in q_0 and decreasing in p_0.*

This proposition shows that as the baseline rate of the first stage (q_0) increases, the amount we give to the first publisher in the optimal f-contract increases. We arrive at this counter-intuitive result because of the way multi-touch contracts compensate publishers. If the baseline rate in the first stage increases, the number of consumers who don't see an ad in the first stage but see an ad from the second publisher in the second stage increases. This implies that the payoff of the second publisher increases (as shown in Equation 2.1). Since there is complementarity between q_0 and the effort exerted by Publisher 2, this increases his incentive to put more effort. However, the incentive for the first publisher does not increase due to an increase in q_0. To balance things out and give incentive to both publishers to put more effort to increase the overall conversion rate, we should increase the value of f, i.e. the amount we give to the first publisher. Increasing f gives Publisher 1 more incentive to increase a, leading to an overall increase in the total conversion rate. The argument for the decrease in f^* with an increase in p_0 is similar as the advertiser wants to incentivize Publisher 2 to put more effort.

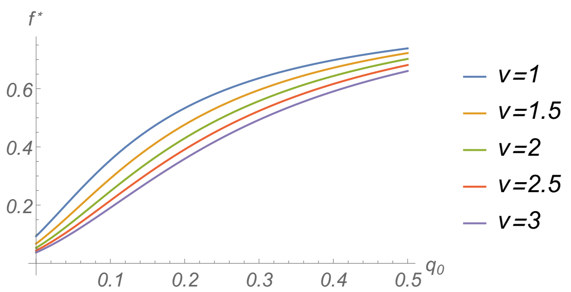


Figure 2.1: The value of f^* as a function of q_0 for various values of v, $p_0 = 0.25$, and $w = 2$.

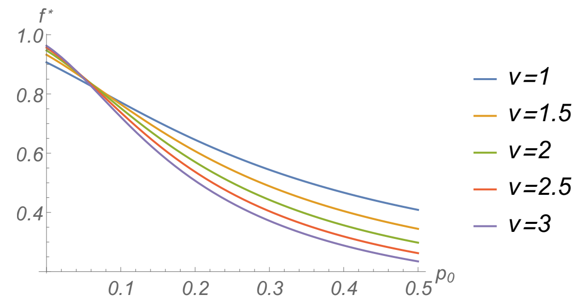


Figure 2.2: The value of f^* as a function of p_0 for various values of v, $q_0 = 0.25$, and $w = 2$.

Our result indicates that if the level of awareness of a product or brand is high, the family of f-contracts needs to incentivize the publisher creating awareness even more, as the second publisher that leads to conversions is automatically receiving a relatively higher payoff. This suggests the counterintuitive recommendation that well known advertisers should be allocating relatively more resources to informational advertising. On the other hand, products that deliver high intrinsic value to consumers so that consumers are likely to buy them if they are aware of the products represent products with high p_0: advertisers for such products should increase their spending on persuasive forms of advertising (in the second stage). As explained in the previous paragraph, these will best balance the automatic incentive for increase in effort in the other stage in optimally splitting the budget. Figures 2.1 and 2.2 illustrate the proposition and show the optimal contract as a function of q_0 and p_0 for some fixed values of the parameters.

In the preceding discussion, we presented the properties of the optimal contract. In reality however, the advertiser may not know the baseline rate of awareness and consideration. However, the following corollary offer insights on how the advertiser should split the advertising dollar if he knows the relationship between the two baseline rates in the first and second stage. If the advertiser has reasons to believe that a larger fraction of consumers are aware of the product but a relatively smaller fraction are likely to convert on their own then he should compensate Publisher 1 more if the cost of advertising is the same across the two publishers.

Corollary 2.1. *Assume that $w = v$. It holds that*

- *If $q_0 > p_0$, then $f^* > \frac{1}{2}$.*
- *If $q_0 = p_0$, then $f^* = \frac{1}{2}$.*
- *If $q_0 < p_0$, then $f^* < \frac{1}{2}$.*

Social Welfare and Price of Anarchy

How does the optimal f-contract perform compared to a centralized solution that maximizes the social welfare? It is clear that the f-contract will be worse than the first-best or socially optimal solution in terms of the social welfare, but it is not clear if the resulting conversion rate would be higher or lower. The following proposition answers this question. It shows that the first benchmark model, where a central planner maximizes the social welfare, can yield a result where everyone (including the firm) is better off compared to the optimal f-contract.

Proposition 2.2. *Consider the first benchmark (where we maximize social welfare of the publishers) and the equilibrium in the optimal f-contract. In the first benchmark, the conversion rate is higher. Moreover, sometimes the publishers' payments are higher too.*

Proposition 2.2 shows that the f-contract gives an inefficient equilibrium, since there can be an alternative solution where everyone is better off. This is a result of the lack of coordination between the publishers. In the centralized solution, both publishers put more effort which turns out to be good for both of them. However, in the f-contract equilibrium we observe a version of the prisoner's dilemma. In particular, when one of the publishers puts a lot of effort, the other one prefers to lower his effort and free-ride. We believe that this result is quite unique in the attribution literature as we show that the general class of f-contracts, which are commonly used in practice, can result in a prisoner's dilemma that adversely affects the efforts exerted by the different publishers.

This result illustrates how multi-channel attribution can be a blind spot for advertisers. Not only they might have limited information that obstructs their view of what the optimal attribution is, but there is also some inherent inefficiency in the widely used multi-touch rules. Thus, even if an advertising firm obtains good estimates of the conversion propensities in the market and determines the best multi-touch, the outcome can still be inefficient. In Section 2.6, we explore how we can use alternative attribution rules to resolve this problem.

Given the inefficiency, a related issue is how bad an f-contract can be in terms of social welfare. This can be measured by a concept called the Price of Anarchy (Koutsoupias and Papadimitriou, 1999), defined as the maximum value of the ratio between the social welfare in the optimal centralized (first-best) solution and the worst possible social welfare in any equilibrium.[16] The higher the price of anarchy, the more inefficient the equilibrium.[17]

As an example in our setup, if we consider the last-touch contract ($f = 0$), with the publishing costs set as low as possible, i.e. $w = v = 1$, and the baseline rate of the second stage set as low as possible, i.e. $p_0 = 0$, then the social welfare in the first benchmark is $\frac{q_0^2}{3}$, while the social welfare in the equilibrium is $\frac{q_0^2}{4}$. This gives a ratio of $\frac{4}{3}$. In the following proposition, we show that this is actually the worst possible case, i.e. the price of anarchy is equal to $\frac{4}{3}$. In other words, the social welfare in an f-contract is never worse that $\frac{3}{4}$ of the optimal.

[16] In our case the equilibrium is unique, so we don't have to worry about determining the worst.

[17] For a prominent example where this concept is used in computer science, see Roughgarden and Tardos (2002).

Proposition 2.3. *Let SW_{OPT} be the social welfare in the first benchmark and SW_f be the social welfare in an f-contract equilibrium. It holds that*

$$1 \le \frac{SW_{OPT}}{SW_f} \le \frac{4}{3}.$$

Moreover, as $w \to +\infty$ or $v \to +\infty$, the ratio $\frac{SW_{OPT}}{SW_f}$ tends to 1.

The intuition behind why an increase in imbalance of the publishing costs w and v results in the social welfare approaching the optimal (first-best) social welfare, is that the high costs resolve the lack of coordination between the publishers in the prisoner's dilemma described above. As the cost of advertising in a particular stage increases, the publisher of this stage puts lesser and lesser effort. At some point he will put no effort at all in both the centralized solution and the equilibrium. At that point only one publisher exerts effort, which means that there is perfect coordination resulting in the first-best outcome.

Figure 2.3 shows the values of the ratio $\frac{SW_{OPT}}{SW_{\text{Last-Touch}}}$ for various parameters. We see that last-touch performs the worst in terms of social welfare when the cost v in the second stage is low and when the baseline p_0 in the second stage is low. Figure 2.4 shows the comparison of the conversion rate in the first benchmark case to the conversion rate under the optimal f-contract. Note that the conversion rate of the optimal f-contract, similarly to the social welfare, performs the worst when the costs (v and w) are low.

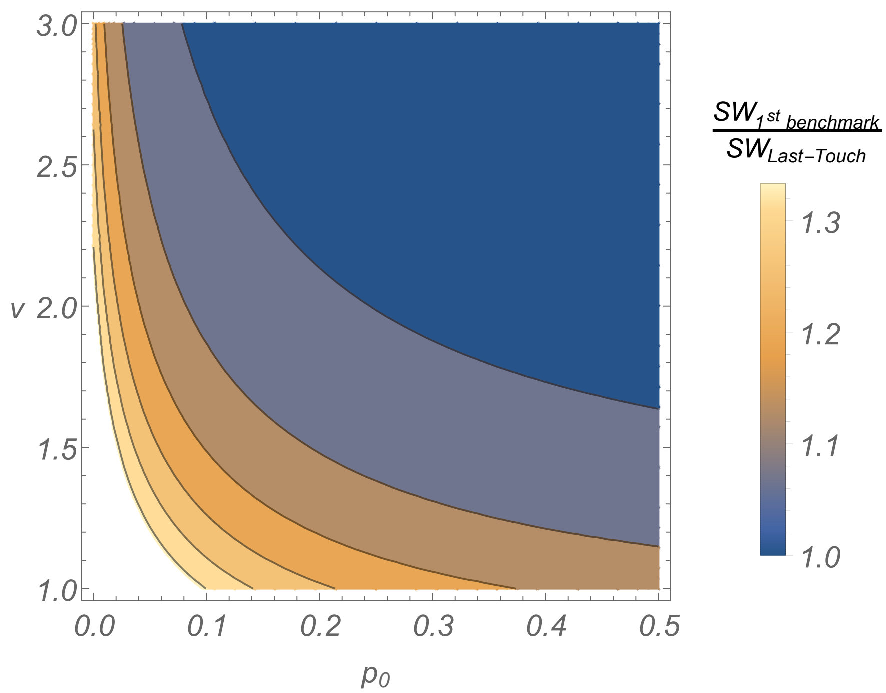


Figure 2.3: Social welfare in the first benchmark (publishers' optimal) over the social welfare in the last-touch contract as a function of p_0 and v, for $q_0 = 0.25$, and $w = 2$.

In summary, for both the firm and the publishers, multi-touch contracts (even the optimal one) are sub-optimal, while they perform better when there is a high degree of heterogeneity

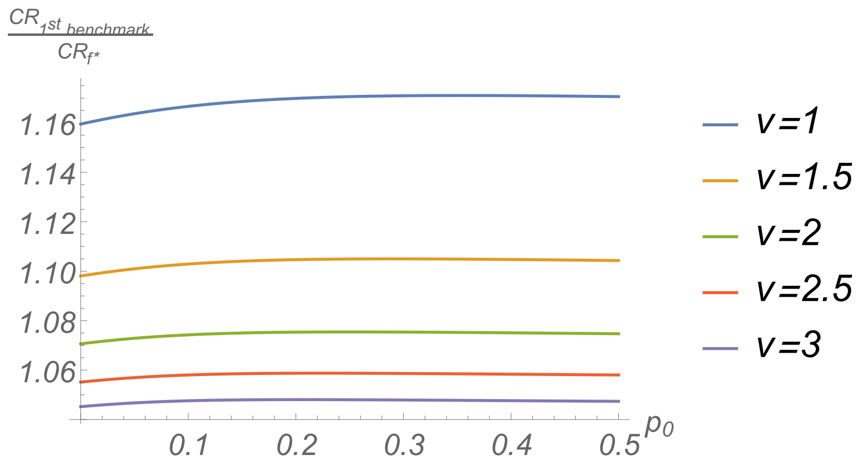


Figure 2.4: Conversion rate in the second benchmark (firm's optimal) over the conversion rate in the optimal f-contract as a function of p_0 for various values of v, $q_0 = 0.25$, and $w = 2$.

in the advertising costs across the two channels (i.e. $w \gg v$ or $w \ll v$). This indicates that advertising will be more efficient (and profitable) in industries where there is a wide discrepancy in the advertising costs.

2.6 Beyond Multi-touch Contracts

In the previous section we showed that the general class of f-contracts are generally sub-optimal. We now consider alternative payment functions and compare them with each other in terms of the conversion rate in the equilibrium.

2.6.1 Optimal Contract

One may wonder if there is a valid contract (in the sense of Section 2.4.4) that achieves the optimal conversion rate (the one achieved in the second benchmark by a central planner). The following proposition shows that this is actually possible with a contract that is linear w.r.t. the observed products q_0p_0, q_0c, ap_0, ac.

Proposition 2.4. *There is a contract that achieve the optimal conversion rate, given by the following payment functions.*

$$g_1(a, q_0, c, p_0) = \frac{1}{2}q_0p_0 + sap_0 + tq_0c + fac$$
$$g_2(a, q_0, c, p_0) = \frac{1}{2}q_0p_0 + (1-s)ap_0 + (1-t)q_0c + (1-f)ac$$

The values of s, t, f are defined below.

$$s = 1 + \frac{q_0 p_0 + 2v(c^*)^2}{2a^* p_0}$$
$$t = -\frac{q_0 p_0 + 2w(a^*)^2}{2q_0 c^*}$$
$$f = \frac{1}{2} + \frac{-a^* p_0 + q_0 c^* + 3w(a^*)^2 - 3v(c^*)^2}{2a^* c^*}$$

Note that a^*, c^* in the above proposition are the optimal values of the optimization problem in the second benchmark (Section 2.4.5).

In order to understand why the contract in Proposition 2.4 achieves the optimal rate, it is useful to compare it to the payments for the publishers in Equation 2.1. As mentioned earlier, any f-contract suffers from a free-riding problem. Both publishers want the *other* publisher to exert effort that leads to a prisoner's dilemma. In the payoff outlined here, the publishers are punished for the effort exerted by the other publisher (both t and $1-s$ are negative). This payment scheme disincentivizes both publishers from free-riding. By appropriately choosing s, t and f, the advertiser is able to achieve perfect coordination between the publishers. In equilibrium the publishers' profits are zero (their outside option).[18]

A drawback of the aforementioned contract is that it is not clear how an advertiser can implement this contract if he has incomplete information about q_0 and p_0, which are required to determine the values of a^* and c^*.

Even though the preceding contract might not be implementable, it provides us the intuition that contracts that penalize publishers for the effort of the other publisher, instead of just compensating them for their efforts, tend to work better. The penalty due to under-performance can increase the efficiency of the system, because it gives incentive to both publishers to put more effort instead of free-riding and depending on the effort of the other publisher. In the following subsection, we define one such contract. We call it a reinforcement contract[19] and show that it performs better than any multi-touch contract in the majority of the cases. Moreover, the firm does not need to know the exact values of the baseline probabilities to implement it, which makes it more practical.

2.6.2 Reinforcement Contracts

Perhaps the simplest form of a valid contract is given by the payment function $g_1 = \frac{1}{2}(a + q_0)(c + p_0)$, which we call *equal-split*. According to this payment, both publishers get the same credit for every conversion. The reinforcement or (r)-contract is a generalization of the equal-split $(r = 0)$ and it is given by the following payment function.

$$g_1 = \left(\frac{1}{2} + \left(\frac{a}{a+q_0} - \frac{c}{c+p_0}\right) r\right)(a+q_0)(c+p_0), \text{ for } r \in \mathbb{R}.$$

[18]We can extend the model such that the publishers have a non-zero outside option, but the qualitative results do not change.

[19]The name is motivated by the similarity to rewarding good performance and penalizing bad performance in reinforcement learning.

It is evident from the preceding equation that an (r)-contract not only compensates publishers for the effort they put, but it also penalizes them for the effort the other publisher puts. The following table shows the canonical form of the corresponding attribution rule.

Ad in the awareness stage	Ad in the conversion stage	Credit to Publisher 1	Credit to Publisher 2
no	no	$\frac{1}{2}$	$\frac{1}{2}$
no	yes	$\frac{1}{2} - r$	$\frac{1}{2} + r$
yes	no	$\frac{1}{2} + r$	$\frac{1}{2} - r$
yes	yes	$\frac{1}{2}$	$\frac{1}{2}$

Note that r can take values above $\frac{1}{2}$, i.e. this attribution rule has the unique property of being able to assign negative credit to channels that do not appear in conversion paths.[20] Thus, instead of focusing only on the touch points like the multi-touch attribution rules do, it pays attention to non-touch points as well. This change of focus can help resolve the commitment problem illustrated in Section 2.5.2. The following proposition will help us determine the optimal (r)-contract.

Lemma 2.1. *The conversion rate under an (r)-contract is a convex function of r.*

Lemma 2.1 tells us that the optimal (r)-contract is achieved for the maximum possible value of r that keeps both publishers into the game, i.e. both publishers have non-negative payoff by putting some effort in the equilibrium. This is useful, because in a practical situation the firm can determine the optimal (r)-contract by increasing the value of r until one of the publishers drops out. In other words, implementing the (r)-contract reduces to a simple search problem in a repeated dynamic environment.

A practical way to implement a type of reinforcement contract like the (r)-contract is to use an attribution method that assigns negative credit to publishers who do not appear in a conversion path. This will incentivize publishers to put more effort into being part of conversion paths, and as a result the conversion rates will increase.

In Figure 2.1, we can see a comparison of performance of the different contracts we considered in this paper. As we can see, the optimal (r)-contract performs significantly better than any other contract that does not require the knowledge of the baselines in the majority of the cases (except for very small values of q_0, where an f-contract is better). The chosen values of the parameters for drawing the plot are arbitrary and the picture looks very similar for different values as well. Note also that in the symmetric case, i.e. when $w = v$ and $q_0 = p_0$, the optimal (r)-contract achieves the first best conversion rate.

2.7 Extensions

In this section, we first discuss how the optimal f-contract behaves as a function of the cost parameters of the publishers, and how a firm can choose the optimal f-contract under limited uncertain information about baseline conversion rates. Then we examine the effect of competition on the advertising game.

[20]In an f-contract, even if we allow f to take negative values, it is never optimal to do so. As we can see in the proof of Proposition 2.1, it is always $0 \leq f^* \leq 1$.

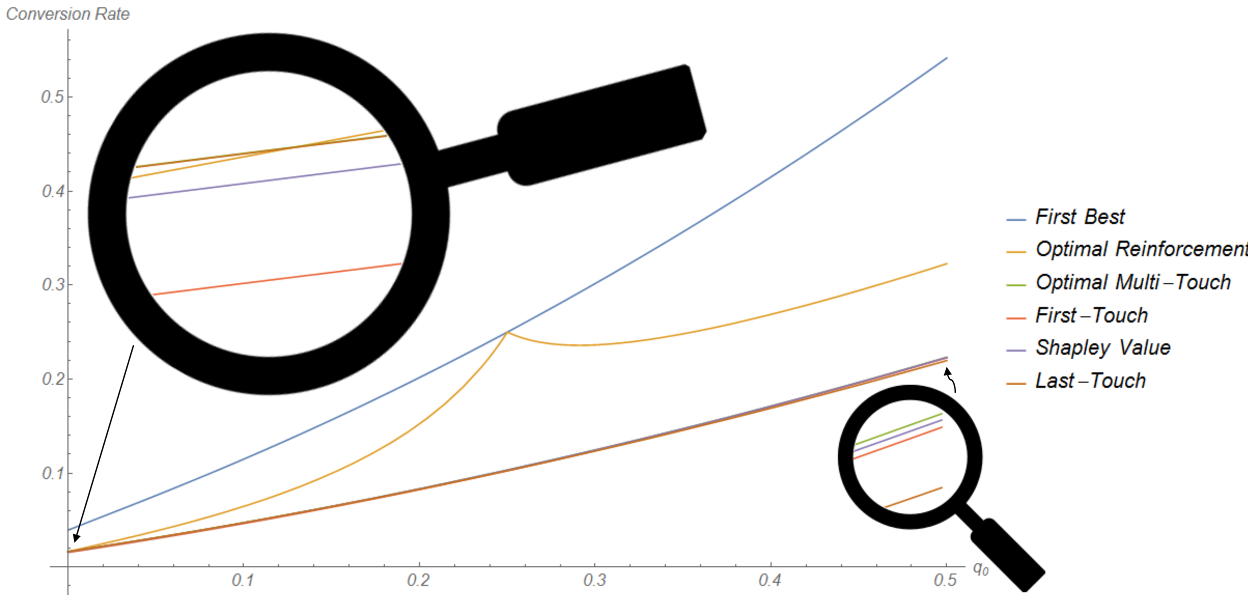


Figure 2.1: Conversion rate for different contracts as a function of q_0 (for $w = v = 2$ and $p_0 = 0.25$). The first best is the one that maximizes conversion rate. The reinforcement contract is second best for a wide range of parameter values.

2.7.1 Effect of Cost Parameters of Publishers in Optimal f-contracts

We now return to the popular f-contracts and explore how the advertising dollar might be split between the two publishers based on the difficulty of advertising in the different stages. The advertising cost can be driven by several factors, e.g. the reach of a publisher, its ability to target the right set of consumers, the intensity of advertising by competitors, or type of the product being sold.

Recall that w represents the cost of advertising effort in the first stage, and v the corresponding value for the second.

Proposition 2.5. *There is a threshold $\overline{w} \in [1, +\infty]$ such that f^* is decreasing in w for $w \in [1, \overline{w}]$ and increasing in w for $w \in [\overline{w}, +\infty)$.*[21] *Similarly, there is a threshold $\overline{v} \in [1, +\infty]$ such that f^* is increasing in v for $v \in [1, \overline{v}]$ and decreasing in v for $v \in [\overline{v}, +\infty)$.*

Interestingly, f^* is not monotone with respect to w. It is easy to see that when the cost of advertising in the first stage increases, Publisher 1 would reduce the amount of effort. As the result, the advertiser compensates Publisher 1 by increasing f^* to induce more effort as w increases. However, we find that when w is sufficiently small, f^* decreases when w increases. In order to understand this counterintuitive finding, it is important to note that even though Publisher 1 is affected directly due to the increase in advertising cost, Publisher 2 is also affected indirectly as w increases. This results in Publisher 2 exerting less effort because his payoff is reduced due to the multiplicative factor of the two efforts in the payoff function presented in Equation 2.1. Consequently, the advertiser needs to incentivize both publishers

[21] $\overline{w}$ is given implicitly by the solution to the equation $\overline{w} = \frac{p_0(1-f^*(\overline{w}))}{2q_0}$.

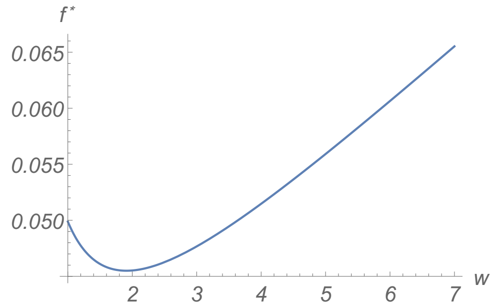


Figure 2.1: The value of f^* as a function of w, for $q_0 = 0.1$, $p_0 = 0.4$, and $v = 10$. The monotonicity changes at $\overline{w} = 1.909$.

to increase the effort. When w is relatively small, it is easier to increase the effort exerted by Publisher 1 as compared to Publisher 2. Under this condition, it is more important to increase the fraction paid to Publisher 2 to increase the overall conversion rate, which leads to a decrease in f^* with w. Figure 2.1 shows the variation in f^* with w.

The following corollary shows how the different advertising costs affect f^*. A firm which uses multi-touch attribution for its campaign should give more credit to the advertising channel with the highest publishing cost, all else being equal.

Corollary 2.2. *Assume that $q_0 = p_0$. It holds that*

- *If $w > v$, then $f^* > \frac{1}{2}$.*
- *If $w = v$, then $f^* = \frac{1}{2}$.*
- *If $w < v$, then $f^* < \frac{1}{2}$.*

2.7.2 Implementing the f-contract Under Uncertainty

To find the optimal f-contract, the firm needs to know the baseline rates q_0, p_0. In this section, we propose an approach that can be used by the firm to determine the optimal f under asymmetric information. The following proposition shows that using very little information the firm can derive the optimal (in expectation) f-contract.

Proposition 2.6. *Assume that the firm does not know the values of q_0, p_0, but it has some information about the distributions from which they are drawn. Then, the firm can calculate the value f^* of the optimal (in expectation) f-contract, by only using the moments $\mathrm{E}[q_0 p_0]$, $\mathrm{E}[q_0^2]$, and $\mathrm{E}[p_0^2]$.*

This proposition shows that the advertiser does not need too much information to derive f^*. As long as he knows the second moments of the baseline rates , he can find the optimal split between the two publishers. Now, suppose that a firm knows that both the publishing costs and the expected baselines rates are the same in the two stages, the following corollary sheds light on how the advertiser should set f^*.

Corollary 2.3. *Let f^* be the value of the optimal (in expectation) f-contract for the firm. If $v = w$ and $\mathrm{E}[q_0] = \mathrm{E}[p_0]$, then*

- *If* $\mathrm{Var}[q_0] = \mathrm{Var}[p_0]$, *then* $f^* = \frac{1}{2}$.

- *If* $\mathrm{Var}[q_0] > \mathrm{Var}[p_0]$, *then* $f^* > \frac{1}{2}$.

- *If* $\mathrm{Var}[q_0] < \mathrm{Var}[p_0]$, *then* $f^* < \frac{1}{2}$.

Corollary 2.3 shows that a firm which uses multi-touch attribution for its campaign should give more credit to the advertising channel with higher uncertainty about the baseline rate, all else being equal.

2.7.3 Competition in Each Stage

In this extension, we assume that both publishers can show ads in both stages of the game. In other words, there is competition not only across stages but also in each stage. In this variation, we want to see how increased competition between the publishers will affect our results.

As in the original model, there is a continuum of consumers of mass 1 moving as in the diagram of Figure 2.1. The variable a now will be a vector (a_1, a_2), where a_1 is a decision variable for the first publisher, with cost $w_1 a_1^2$, and a_2 is a decision variable for the second publisher, with cost $w_2 a_2^2$. Similarly, the variable c will be a vector (c_1, c_2), where c_1 is a decision variable for the first publisher, with cost $v_1 c_1^2$, and c_2 is a decision variable for the second publisher, with cost $v_2 c_2^2$. For the functions f and g, we assume the functional forms $f(a) = a_1 + a_2 + q_0$ and $g(c) = c_1 + c_2 + p_0$, where q_0 and p_0 are the baseline probabilities.

In each stage, we assume that a consumer can see at most one ad. The effort a_1 represents the probability that during the awareness stage he will see an ad from the first publisher, a_2 is the probability that he will see an ad from the second publisher, and q_0 is the probability that he will move to the conversion stage without seeing an ad. Similarly, the effort c_1 represents the probability that during the conversion stage he will see an ad from the first publisher, c_2 is the probability that he will see an ad from the second publisher, and p_0 is the probability that he will decide to buy the product without seeing an ad during the conversion stage.

As before, we assume that the baseline probabilities q_0, p_0 cannot be more than $\frac{1}{2}$ and that the cost parameters v_1, v_2, w_1, w_2 are sufficiently large (so that $f(a), g(c) \in [0, 1]$ in the equilibrium).

For every conversion, the firm wants to spend \$1 in advertising. The question is what is the optimal way to split this dollar between the two publishers in order to maximize the number of conversions, which is $f(a)g(c)$.

The equivalent of an f-contract in this variation is summarized in the following table.

Ad in the awareness stage	Ad in the conversion stage	Credit to Publisher 1	Credit to Publisher 2
0	0	$\frac{1}{2}$	$\frac{1}{2}$
0	1	1	0
0	2	0	1
1	0	1	0
1	1	1	0
1	2	f	$1-f$
2	0	0	1
2	1	$1-f$	f
2	2	0	1

A 0 in the first two columns means no ad, a 1 means ad from the first publisher and a 2 means ad from the second publisher. The parameter $f \in [0,1]$ is some value determined by the firm or externally.

As in the original model, some examples of contracts of this form are the last-touch ($f = 0$), where all the credit goes to the last ad the consumer had seen prior to the conversion, and the first-touch ($f = 1$) where all the credit goes to the first ad, which made the consumer aware of the product. The following lemma will help us determine the optimal f-contract in this extended model.

Lemma 2.2. *Let $r(f)$ be the conversion rate in an f-contract as a function of f. Then, $r(f)$ is concave in $(0,1)$.*

Since $r(f)$ is concave, we know that the optimal value f^* is either the single root of the equation $r'(f) = 0$ in $(0,1)$ (when it exists), or 0, or 1. More specifically, if $r'(f) < 0$ in $(0,1)$, then $f^* = 0$. If $r'(f) > 0$ in $(0,1)$, then $f^* = 1$. If there is a root f such that $r'(f) = 0$, then it is unique and f^* is equal to this root.

To show the robustness of our results in this extended model, we consider the following numerical examples.

Figures 2.2 and 2.3 are the analogs of Proposition 2.1. As we can see, f^* is increasing in q_0 and decreasing in p_0.

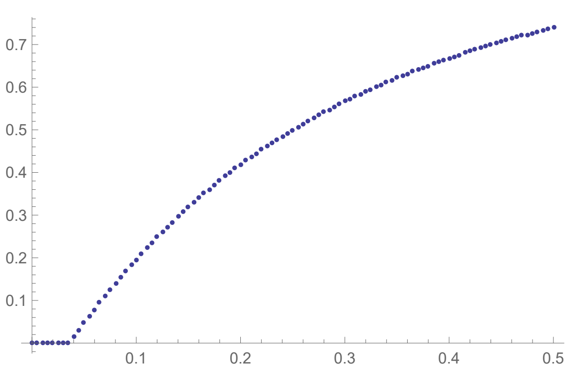


Figure 2.2: Optimal f for different values of q_0, for $v_1 = v_2 = w_1 = w_2 = 2$, $p_0 = \frac{1}{4}$.

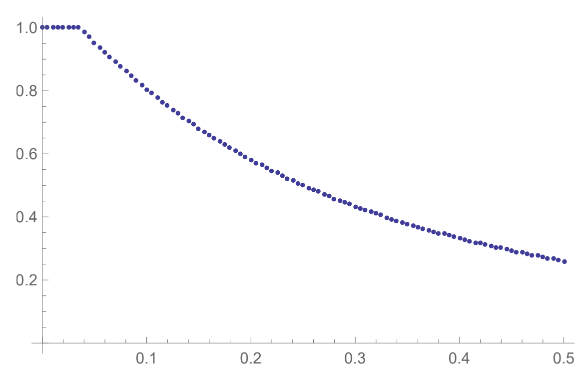


Figure 2.3: Optimal f for different values of p_0, for $v_1 = v_2 = w_1 = w_2 = 2$, $q_0 = \frac{1}{4}$.

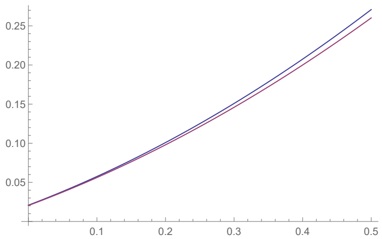


Figure 2.4: Social Welfare in the equilibrium (purple) and in the first best (blue) for different values of q_0, for $v_1 = v_2 = w_1 = w_2 = 2$, $p_0 = \frac{1}{4}$, $f = 0$.

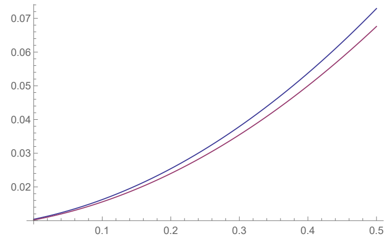


Figure 2.5: First publisher's profit in the equilibrium (purple) and in the first best (blue) for different values of q_0, for $v_1 = v_2 = w_1 = w_2 = 2$, $p_0 = \frac{1}{4}$, $f = 0$.

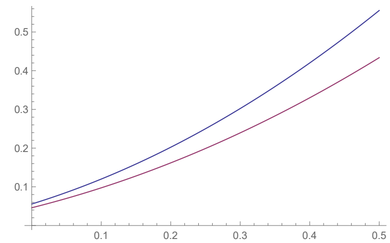


Figure 2.6: Firm's revenue in the equilibrium (purple) and in the first best (blue) for different values of q_0, for $v_1 = v_2 = w_1 = w_2 = 2$, $p_0 = \frac{1}{4}$, $f = 0$.

Figures 2.4, 2.5, and 2.6 are the analogs of Proposition 2.2. As we can see, in the first best the conversion rate is higher, and sometimes publisher's profits are higher too.

Figures 2.7 and 2.8 are the analogs of Proposition 2.3. In this case, the upper bound of the ratio is even tighter at $\frac{512}{507}$.

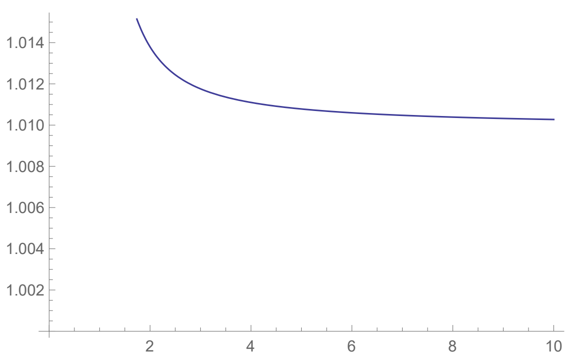


Figure 2.7: Ratio of Social welfare in the first best over the Social welfare in the equilibrium for different values of a, for $q_0 = p_0 = \frac{1}{4}$, $f = 0$, $v_1 = w_2 = lo = 2$, $v_2 = w_1 = hi$, $hi = a \cdot lo$. The limit as $a \to \infty$ is $\frac{512}{507} = 1.00986$.

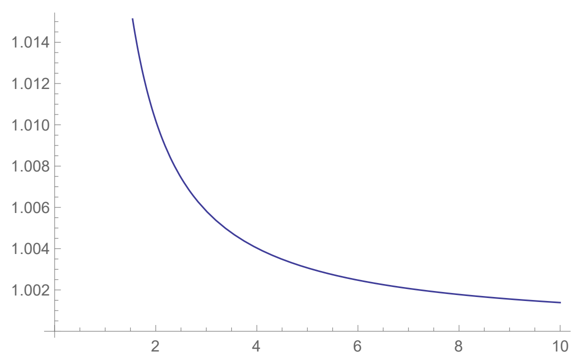


Figure 2.8: Ratio of Social welfare in the first best over the Social welfare in the equilibrium for different values of a, for $q_0 = p_0 = \frac{1}{4}$, $f = 0$, $v_1 = v_2 = lo = 2$, $w_1 = w_2 = hi$, $hi = a \cdot lo$. The limit as $a \to \infty$ is 1.

Finally, Figures 2.9 and 2.10 are the analogs of Proposition 2.6 and Corollary 2.3. The firm can determine f^* by using the moments $\mathrm{E}[q_0 p_0]$, $\mathrm{E}[q_0^2]$, and $\mathrm{E}[p_0^2]$. Moreover, if q_0 and p_0 have the same mean, but q_0 has higher variance, then $f^* > \frac{1}{2}$.

2.8 Conclusion

The online advertising industry has continued to grow rapidly in the last two decades. While it offers several advantages over traditional advertising, the information asymmetry and

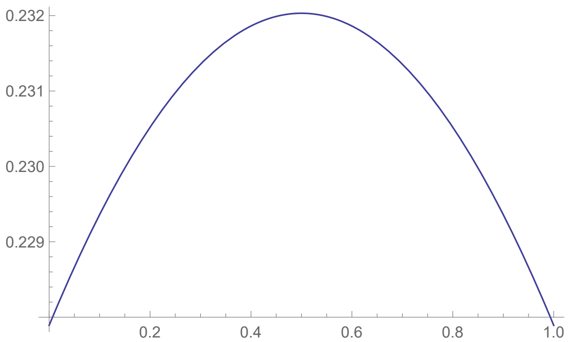


Figure 2.9: Firm's revenue for different values of f when $\mathrm{E}[p_0] = \mathrm{E}[q_0] = \frac{1}{4}$ and $\mathrm{E}[p_0^2] = \mathrm{E}[q_0^2] = \frac{1}{12}$, for $v_1 = w_1 = v_2 = w_2 = 2$. The optimal value is for $f = \frac{1}{2}$.

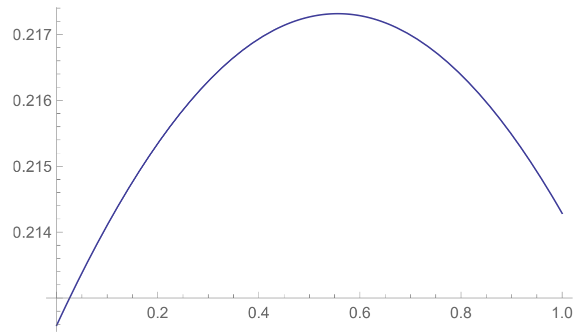


Figure 2.10: Firm's revenue for different values of f when $\mathrm{E}[p_0] = \mathrm{E}[q_0] = \frac{1}{4}$ and $\mathrm{E}[p_0^2] = \frac{1}{16}$, $\mathrm{E}[q_0^2] = \frac{1}{12}$, for $v_1 = w_1 = v_2 = w_2 = 2$. The optimal value is for $f = 0.556179$.

misalignment of incentives between the advertisers and publishers poses a threat to this industry in form of ad fraud, substandard ad inventory, and sub-optimal effort by advertisers. In this paper, we formalize this problem in the context where consumers move through two stages before they purchase and show how standard contracts, also known as multi-touch contracts, might not lead to the most effective outcome for advertisers *and* publishers. Multi-touch contracts can result in the advertiser receiving lower return on investment and the publishers generating smaller revenues in equilibrium. The inefficiency in multi-touch attribution arises due to a prisoner's dilemma where the dominant strategy of publishers in both stages is to exert less effort in showing ads. We also show that our results are in line with real-world observations. Furthermore, motivated by optimal contracts, we introduce reinforcement contracts, that penalize publishers if they exert relatively less effort than the other publisher, and show how they can reduce the prisoner's dilemma and lead to higher profits for the advertiser.

Our research has several managerial implications. One important findings of our research is that advertisers should spend a relatively larger fraction of their advertising budget on the stage of the purchase funnel that they believe has a higher baseline conversion rate. E.g. if an advertiser knows that the level of brand (or product) awareness is higher than the baseline conversion probability in the consideration stage, they should spend more on brand advertising, even if the exact levels are unknown. Secondly, in industries where the cost to create awareness is similar to the cost of showing consumers ads to drive conversion, the advertisers should engage with publishers that can show ads in both stages of the funnel. This will reduce the prisoner's dilemma across the stages and help in better coordination of the incentives between the advertiser and the publisher. Furthermore, the results presented in this paper underscore the need for better transparency in online advertising. The online advertising ecosystem is becoming extremely complicated with many participants, often with misaligned incentives. This has led to an increase in data fragmentation, which increases the information asymmetry. This can lead to online advertising becoming a less effective medium and reduce advertisers' desire to move their advertising budget online. Sharing information across all the participants resolves this problem and can increase the effectiveness of online advertising, making both advertisers and publishers better off. Finally, if the advertiser has enough market power, it should carefully consider reinforcement contracts for attribution to increase the returns from advertising. This will prevent the publishers from getting stuck in the unfavorable prisoner's dilemma equilibrium.

Our research has a few limitations and presents directions for future work. Our model focuses on a single advertiser as he allocates credit between competing publishers. In that sense, our advertiser has market power to dictate the attribution rule. A possible extension is to consider how competition between different advertisers can affect the attribution process, where the market power of one advertiser would be considerably reduced. Our insights can be used by each advertiser in the stage of budget allocation among different advertising channels, but advertisers' competition could give some extra insights on patterns of overall spending.

The reinforcement attribution contracts we suggest as improvement over multi-touch contracts have some good theoretical properties and are more practical than optimal contracts, but there might still be some challenges in their implementation. The key insight from them is that assigning negative credit to non-appearances of advertising channels in conversion paths can be beneficial for the advertiser, because it increases competition between the publishers and it incentivizes them to put more effort. However, an advertiser who wants to implement such an attribution contract should also take into account the quality of the ads. A low-quality ad that appears in a lot of conversion paths should not be incorrectly considered as important for conversions. To do this properly, an advertiser needs to have sufficient data about consumer paths that did not lead to a conversion. A future research direction, therefore, is to dive more into the intricacies regarding making reinforcement contracts more practical.

Finally, one important result of our research is to show that an attribution based on Shapley value is often not optimal for advertisers. Shapley value uses the marginal contribution of each publisher to the conversion rate in order to provide a fair attribution, but it does not take into account the publishers' incentives. As a result, we have shown that there are alternative attribution schemes with better results for the advertiser. Another future research direction is to extend the framework of this paper and provide a general formulation of an optimal scheme for ad attribution.

2.A Appendix

2.A.1 Analyses and Proofs

Proof of Proposition 2.1. The payoff of the first publisher in the f-contract is $\frac{1}{2}q_0p_0 + ap_0 + fac - wa^2$, while the payoff of the second publisher is $\frac{1}{2}q_0p_0 + q_0c + (1-f)ac - vc^2$. This means that the equilibrium efforts are $a = \frac{fq_0+2vp_0}{4vw-f(1-f)}$ and $c = \frac{(1-f)p_0+2wq_0}{4vw-f(1-f)}$. Therefore, the conversion rate in the equilibrium is

$$r(f) = \frac{(f^2q_0 + 2v(p_0 + 2wq_0))((1-f)^2p_0 + 2w(q_0 + 2vp_0))}{(4vw - f(1-f))^2}.$$

It holds that $r'(f) =$

$$\frac{-4f^3(vp_0^2+wq_0^2)+12f^2vp_0^2-4f(4v^2wp_0^2+3vp_0^2+8vwp_0q_0+4vw^2q_0^2)+4(p_0^2v+4vwp_0q_0+4vw^2q_0^2)}{(4vw-f(1-f))^3}.$$

The discriminant of the cubic formula in the numerator is

$$\Delta = -256v^2w^2(2(vp_0^2+wq_0^2)+p_0q_0)^3(2(vp_0^2+wq_0^2)(27+16vw)+9(3+16vw)p_0q_0) \leq 0,$$

which means that the equation $r'(f) = 0$ has a single real root. It's also $r'(0) = \frac{p_0^2v+4vwp_0q_0+4vw^2q_0^2}{16v^3w^3} \geq 0$ and $r'(1) = -\frac{(q_0+2p_0v)^2}{16v^3w^2} \leq 0$. Therefore, the single root is in the interval $[0,1]$. That root is the value of f^*. The result is a long expression, so to write it down we use the following notation. Let

$$\begin{aligned}
&x = wq_0^2, y = vp_0^2, z = q_0p_0, H = 2(x+y)+z\\
&A = H\left(H(16xy+27z^2)+128xyz\right),\\
&B = \sqrt[3]{xyH\left(9(x-y)zH+\sqrt{3}(x+y)\sqrt{A}\right)},\\
&n = \sqrt[3]{12}, m = 2\sqrt[3]{18}.
\end{aligned}$$

Then

$$f^* = \frac{y}{x+y} + \frac{nB^2 - mxyH(H+2z)}{6(x+y)zB}.$$

Now, we need to prove that f^* is increasing in q_0 and decreasing in p_0. We'll start by proving that it is increasing in q_0. Since $r'(f^*) = 0$, it holds that

$$-a(f^*)^3 + b(f^*)^2 - cf^* + d = 0,$$

where $a = 4(vp_0^2+wq_0^2), b = 12vp_0^2, c = 4(4v^2wp_0^2+3vp_0^2+8vwp_0q_0+4vw^2q_0^2)$, and $d = 4(p_0^2v+4vwp_0q_0+4vw^2q_0^2)$. By taking partial derivatives with respect to q_0 and solving for $\frac{\partial f^*}{\partial q_0}$, we get

$$\frac{\partial f^*}{\partial q_0} = \frac{-\frac{\partial a}{\partial q_0}(f^*)^3 + \frac{\partial b}{\partial q_0}(f^*)^2 - \frac{\partial c}{\partial q_0}f^* + \frac{\partial d}{\partial q_0}}{3a(f^*)^2 - 2bf^* + c}$$

The denominator is equal to

$$\begin{aligned}
&12(vp_0^2+wq_0^2)(f^*)^2 - 24vp_0^2f^* + 4(4v^2wp_0^2+3vp_0^2+8vwp_0q_0+4vw^2q_0^2) =\\
&12vp_0^2(f^*-1)^2 + 12wq_0^2(f^*)^2 + 16vw(vp_0^2+2p_0q_0+wq_0^2) \geq 0,
\end{aligned}$$

therefore, it is enough to prove that the numerator is positive. The numerator is equal to

$$8w(-q_0(f^*)^3 - 4vw(p_0+wq_0)f^* + 2vw(p_0+2wq_0^2)),$$

thus we need to prove that $-q_0(f^*)^3 - 4vw(p_0+wq_0)f^* + 2vw(p_0+2wq_0^2) \geq 0$. We know that $1 = \frac{a(f^*)^3-b(f^*)^2+cf^*}{d}$, so it is enough to prove that

$$-q_0(f^*)^3 - 4vw(p_0+wq_0)f^* + 2vw(p_0+2wq_0^2)\frac{a(f^*)^3-b(f^*)^2+cf^*}{d} \geq 0,$$

or equivalently

$$\frac{fp_0\left((2vp_0-q_0)(f^*)^2-6vp_0f^*+4vw(q_0+2vp_0)+2p_0v\right)}{p_0+2wq_0}\geq 0.$$

It is

$$\begin{aligned}
&(2vp_0-q_0)(f^*)^2-6vp_0f^*+4vw(q_0+2vp_0)+2p_0v=\\
&vp_0\left(2(f^*)^2-6f^*+2+8vw\right)+q_0\left(-(f^*)^2+4vw\right)\geq\\
&vp_0\left(2(f^*)^2-6f^*+2+8\right)+q_0\left(-(f^*)^2+4\right)=\\
&2vp_0\left(\left(f^*-\frac{3}{2}\right)^2+\frac{11}{4}\right)+q_0\left(2-f^*\right)\left(2+f^*\right)\geq 0.
\end{aligned}$$

Therefore, it holds that $\frac{\partial f^*}{\partial q_0}\geq 0$, which means that f^* is increasing in q_0. Because of symmetry, $1-f^*$ is increasing in p_0, which means that f^* is decreasing in p_0. □

Proof of Corollary 2.1. Let $w=v$. If $q_0=p_0$, it holds that $wq_0^2=vp_0^2$, therefore using the notation of the proof of Proposition 2.1, $x=y$. This means that

$$A=(4x+z)((4x+z)(16x^2+27z^2)+128x^2z)=(4x+z)(4x+3z)^3$$

and

$$B=\sqrt[3]{2x^3(4x+z)\sqrt{3}\sqrt{A}}=x\sqrt[3]{2\sqrt{3}}\sqrt{(4x+z)(4x+3z)}.$$

Thus,

$$nB^2-mxyH(H+2z)=2x^2\sqrt[3]{18}(4x+z)(4x+3z)-2\sqrt[3]{18}x^2(4x+z)(4x+3z)=0,$$

which means that $f^*=\frac{y}{x+y}=\frac{1}{2}$.

From Proposition 2.1, we know that f^* is increasing in q_0, therefore if $q_0>p_0$, $f^*>\frac{1}{2}$, and if $q_0<p_0$, $f^*<\frac{1}{2}$. □

Proof of Proposition 2.2. The optimal efforts (in terms of social welfare) are $a_{OPT}=\frac{q_0+2vp_0}{4vw-1}$ and $c_{OPT}=\frac{p_0+2wq_0}{4vw-1}$. In an f-contract, the equilibrium efforts are $a_f=\frac{fq_0+2vp_0}{4vw-f(1-f)}$ and $c_f=\frac{(1-f)p_0+2wq_0}{4vw-f(1-f)}$. It holds that $a_{OPT}\geq a_f$ and $c_{OPT}\geq c_f$, which means that the conversion rate is higher in the first benchmark. This is true for every f-contract, therefore for the optimal f^*-contract as well.

The first publisher's payment in the first best solution (under an f-contract) is

$$p_{1,OPT}=\frac{q_0p_0}{2}+\frac{(2p_0v+q_0)(p_0(f+2vw-1)+(2f-1)q_0w)}{(4vw-1)^2},$$

while in the equilibrium of an f-contract is

$$p_{1,f}=\frac{q_0p_0}{2}+\frac{w(fq_0+2p_0v)^2}{(4vw-f(1-f))^2}.$$

We want to prove that sometimes $p_{1,OPT} \geq p_{1,f^*}$. More specifically, we'll show that this is true for the symmetric case where $q_0 = p_0$ and $v = w$. In that case, it is $f^* = \frac{1}{2}$ and the inequality $p_{1,OPT} \geq p_{1,f^*}$ is equivalent to

$$\frac{1}{2(2w-1)} \geq \frac{4w}{(4w-1)^2},$$

which is true, since it is equivalent to $16w^2 - 8w + 1 \geq 16w^2 - 8w$. □

Proof of Proposition 2.3. The inequality $\frac{SW_{OPT}}{SW_f} \leq \frac{4}{3}$ is equivalent to

$$a' q_0 p_0 + b' q_0^2 + c' p_0^2 \leq 0,$$

where

$$\begin{aligned}
a' &= 4(1-f)^2 f^2 + 32(1-f) f v^2 w^2 - 4((1-f)^2 f^2 + 8(1-f)f - 4)vw - 64 v^3 w^3, \\
b' &= w\left(\left(3(1-f)^2 + 4\right) f^2 + 8(1-f)(2-f)vw - 16 v^2 w^2\right), \\
c' &= v\left(\left(3f^2 + 4\right)(1-f)^2 + 8f(1+f)vw - 16 v^2 w^2\right).
\end{aligned}$$

Therefore, it is enough to prove that $a' \leq 0, b' \leq 0$, and $c' \leq 0$.

The expression $(3f^2 + 4)(1-f)^2 + 8f(1+f)vw - 16v^2w^2$ is a second degree polynomial with respect to vw with largest root

$$\frac{f(1+f) + 2\sqrt{1 - (1-f)f\left(2 + f^2\right)}}{4} \leq 1.$$

Since $vw \geq 1$, i.e. larger than the larger root, and the coefficient of v^2w^2 is negative, the value of the polynomial is non-positive. This means that $c' \leq 0$. Similarly, $b' \leq 0$ (we just replace f with $1-f$).

For a', we have that

$$a' + 48 = 4(1 - vw)\left(4(3 + 4vw) + (4vw - (1-f)f)^2\right) \leq 0,$$

which means that $a' < 0$, and we are done.

The upper bound of the ratio is achieved for $(p_0, f, v, w) = (0, 0, 1, 1)$ or $(q_0, f, v, w) = (0, 1, 1, 1)$.

It remains to show that as $w \to +\infty$ or $v \to +\infty$, the ratio $\frac{SW_{OPT}}{SW_f}$ tends to 1. It holds that $\frac{SW_{OPT}}{SW_f} =$

$$\frac{(4vw - f(1-f))^2 \left(vp_0^2 + wq_0(4vp_0 + q_0)\right)}{(4vw-1)\left(w\left(f^2 q_0(8vp_0 + q_0) - 8fvp_0q_0 + 4vp_0(vp_0 + q_0)\right) + (1-f)^2 p_0\left(f^2 q_0 + vp_0\right) + 4vw^2 q_0(4vp_0 + q_0)\right)}.$$

Both the numerator and the denominator are third degree polynomials in w. The coefficients of w^3 in both of these polynomials are equal to $16v^2q_0(4vp_0 + q_0)$. Therefore, as $w \to +\infty$, the ratio goes to $\frac{16v^2q_0(4vp_0+q_0)}{16v^2q_0(4vp_0+q_0)} = 1$. Similarly, as $v \to +\infty$, the ratio goes to 1. □

Proof of Proposition 2.4. Let a^*, c^* be the efforts in the optimal solution (second benchmark). We define the payments functions

$$g_1(a, q_0, c, p_0) = \frac{1}{2}q_0p_0 + sap_0 + tq_0c + fac$$

and

$$g_2(a, q_0, c, p_0) = \frac{1}{2}q_0p_0 + (1-s)ap_0 + (1-t)q_0c + (1-f)ac,$$

where

$$\begin{aligned} s &= 1 + \frac{q_0p_0 + 2v(c^*)^2}{2a^*p_0}, \\ t &= -\frac{q_0p_0 + 2w(a^*)^2}{2q_0c^*}, \\ f &= \frac{1}{2} + \frac{-a^*p_0 + q_0c^* + 3w(a^*)^2 - 3v(c^*)^2}{2a^*c^*}. \end{aligned}$$

Since a^*, c^* are the optimal efforts, we know that they satisfy the equality

$$(a^* + q_0)(c^* + p_0) = w(a^*)^2 + v(c^*)^2.$$

It is

$$\left.\frac{\partial(g_1(a, q_0, c^*, p_0) - wa^2)}{\partial a}\right|_{a=a^*} = sp_0 - 2wa^* + fc^* = \frac{(a^* + q_0)(c^* + p_0) - w(a^*)^2 - v(c^*)^2}{2a^*} = 0.$$

Similarly,

$$\left.\frac{\partial(g_2(a^*, q_0, c, p_0) - vc^2)}{\partial c}\right|_{c=c^*} = (1-t)q_0 - 2vc^* + (1-f)a^* = \frac{(a^* + q_0)(c^* + p_0) - w(a^*)^2 - v(c^*)^2}{2c^*} = 0.$$

This means that (a^*, c^*) is the equilibrium under the contracts g_1, g_2. To complete the proof, we need to verify that the payoffs of the two publishers in the equilibrium are non-negative. For the payoffs, we have that

$$g_1(a^*, q_0, c^*, p_0) - w(a^*)^2 = \frac{(a^* + q_0)(c^* + p_0) - w(a^*)^2 - v(c^*)^2}{2} = 0$$

and

$$g_2(a^*, q_0, c^*, p_0) - v(c^*)^2 = \frac{(a^* + q_0)(c^* + p_0) - w(a^*)^2 - v(c^*)^2}{2} = 0.$$

□

Proof of Lemma 2.1. The conversion rate under an (r)-contract is equal to

$$\left(\frac{(2r+1)(q_0 + 4p_0v)}{16vw - 1} + q_0\right)\left(\frac{(2r+1)(p_0 + 4q_0w)}{16vw - 1} + p_0\right),$$

which is a convex function of r. Furthermore, the conversion rate is maximized when the participation constraint for one of the publisher is binding. □

Proof of Proposition 2.5. Using the same notation as in the proof of Proposition 2.1, we have that

$$\frac{\partial f^*}{\partial w} = \frac{-\frac{\partial a}{\partial w}(f^*)^3 + \frac{\partial b}{\partial w}(f^*)^2 - \frac{\partial c}{\partial w}f^* + \frac{\partial d}{\partial w}}{3a(f^*)^2 - 2bf^* + c}$$

and we know that the denominator is positive, therefore we need to see when the numerator is positive. Let N be the numerator, then it holds that

$$wN = wN - (-a(f^*)^3 + b(f^*)^2 - cf^* + d) = 4(1-f^*)v(4w^2q_0^2 - (1-f^*)^2p_0^2).$$

Therefore, the numerator is positive iff $2wq_0 > (1-f^*)p_0$ and negative if $2wq_0 < (1-f^*)p_0$. This means that when f^* is bellow the line $1-\frac{2wq_0}{p_0}$, f^* is decreasing in w, and when it's above that line, it's increasing in w. From this and the fact that f^* is continuous, we conclude that f^* can cross the line $1-\frac{2wq_0}{p_0}$ in at most one point. If that point exists, then $\overline{w}$ is the root of the equation $f^* = 1-\frac{2wq_0}{p_0}$ with respect to w. If that point doesn't exist, then either f^* is always increasing, which means $\overline{w} = 1$, or always decreasing, which means $\overline{w} = +\infty$.

The result for $\overline{v}$ now follows because of symmetry. □

Proof of Corollary 2.2. Let $q_0 = p_0$. Similarly to the proof of Corollary 2.1, if $w = v$, it is $x = y$ and therefore $f^* = \frac{1}{2}$.

For the other two parts, we need the fact that if $q_0 = p_0$, then f^* is increasing in w. This doesn't come immediately from Proposition 2.5, but we'll show that it is true. Using the notation of the proof of Proposition 2.5, we have that

$$\begin{aligned} wN &= 4(1-f^*)v(4w^2q_0^2 - (1-f^*)^2p_0^2) = \\ &= 4(1-f^*)v(4w^2 - (1-f^*)^2)q_0^2 \geq \\ &\geq 4(1-f^*)v(4 - (1-f^*)^2)q_0^2 = \\ &= 4(1-f^*)v(1+f^*)(3-f^*)q_0^2 \geq 0. \end{aligned}$$

Therefore, $\frac{\partial f^*}{\partial w} \geq 0$ and the result follows. □

Proof of Proposition 2.6. For the expected conversion rate, we have that

$$\mathrm{E}[r(f)] = \frac{\mathrm{E}[p_0^2](2(1-f)^2v + 8v^2w) + \mathrm{E}[q_0^2](2f^2w + 8vw^2) + \mathrm{E}[p_0q_0]((1-f)^2f^2 + 4vw + 4(1-f)^2vw + 4f^2vw + 16v^2w^2)}{(4vw - f(1-f))^2}.$$

In other words, it depends only on $\mathrm{E}[q_0p_0]$, $\mathrm{E}[q_0^2]$, and $\mathrm{E}[p_0^2]$. □

Proof of Corollary 2.3. Let $w = v$ and $\mathrm{E}[q_0] = \mathrm{E}[p_0]$. Using the notation of the proof of Proposition 2.1 with the difference that now $x = w\,\mathrm{E}[q_0^2], y = v\,\mathrm{E}[p_0^2]$, and $z = \mathrm{E}[q_0p_0]$ (instead of wq_0^2, vp_0^2, and q_0p_0), we know that f^* is the single real root of the polynomial $-af^3 + bf^2 - cf + d$, where $a = 4(x+y), b = 12y, c = 4(4vwy + 3y + 8vwz + 4vwx)$, and $d = 4(y + 4vwz + 4vwx)$.

If $\mathrm{Var}[q_0] = \mathrm{Var}[p_0]$, then that implies $\mathrm{E}[q_0^2] = \mathrm{E}[p_0^2]$, which means that $x = y$. For $f = \frac{1}{2}$, we get $-af^3 + bf^2 - cf + d = \frac{1}{2}(x-y)(16vw-1) = 0$. This means that $\frac{1}{2}$ is the root, i.e. $f^* = \frac{1}{2}$.

If $\mathrm{Var}[q_0] > \mathrm{Var}[p_0]$, then $x > y$. The coefficient of f^3 in the polynomial above is negative and since it has a single real root, the polynomial is positive for f smaller than the root and negative for f larger the root. For $f = \frac{1}{2}$, we get $-af^3+bf^2-cf+d = \frac{1}{2}(x-y)(16vw-1) > 0$. This means that for the root f^*, it holds that $f^* > \frac{1}{2}$.

Similarly, if $\mathrm{Var}[q_0] < \mathrm{Var}[p_0]$, then $x < y$. Therefore, for $f = \frac{1}{2}$, we get $-af^3 + bf^2 - cf + d = \frac{1}{2}(x-y)(16vw-1) < 0$, which means that $f^* < \frac{1}{2}$. □

Proof of Lemma 2.2. Here is the sketch of the proof in 11 steps:

- To prove that $r(f)$ is concave, it is enough to prove that $r'(f)$ is decreasing in $(0, 1)$.
- To prove that $r'(f)$ is decreasing, it is enough to prove that $r''(f)$ is negative in $(0, 1)$.
- We can write $r''(f)$ as $h(f)g(f)$.
- We can prove that $h(f)$ is positive in $(0, 1)$. Therefore, it is enough to prove that $g(f)$ is negative in $(0, 1)$.
- To prove that $g(f)$ is negative in $(0, 1)$, it is enough to prove that $g(f)$ is convex in $(0, 1)$ and that $g(0) < 0$ and $g(1) < 0$.
- We can prove that $g(0) < 0$ and $g(1) < 0$, so it remains to prove that $g(f)$ is convex in $(0, 1)$.
- To prove that $g(f)$ is convex in $(0, 1)$, it is enough to prove that $g''(f)$ is positive in $(0, 1)$.
- To prove that $g''(f)$ is positive in $(0, 1)$, we can first prove that $g''(f)$ is convex, and then that $g''(m) > 0$, where m is the point in $(0, 1)$ where g'' attains its minimum.
- To prove that $g''(f)$ is convex, we can prove that $g^{(4)}(f)$ is positive in $(0, 1)$.
- To find m, we will solve the equation $g^{(3)}(f) = 0$, which will give a single root.
- Finally, we can prove that $g''(m) > 0$.

□

2.A.2 A More General Model and Equivalence

In this section, we justify some simplifications we made in our model in order to make it easier to analyze. We start with a general model that captures all the important things that we want to model, and then we see how we can transition from the general model to our model. In the end, we prove an equivalence result, where we show that for any set of parameter values of the general model and any attribution rule, there are parameter values in our model that give the same equilibrium behavior and the same conversion rate. In other words, we don't lose much with the simplifications and at the same time, we make the problem more tractable.

The General Model

Every consumer who enters the funnel is exposed to an awareness ad with probability r_1. If a consumer does not see an awareness ad, then the probability that he will move to the next stage in the funnel is b_1 (baseline). If a consumer is exposed to an awareness ad, but he was not going to consider the product without the ad, then he moves to the next stage with probability e_1. We assume that r_1 and b_1 are fixed, while e_1 is a decision variable for the first publisher (effort).

The reason we assume that r_1 is fixed is because this can be something pre-determined and it is also easily measured. What is not observable by a firm is the effort the publisher puts in showing ads. The effort will affect how effective the ads will be, e.g. it will show how good targeting the publisher does. Higher effort comes with a cost for the publisher given by the convex function $c_1 e_1^2$.

Summarizing the above, there are eight types of consumers as shown in the following table

	Exposed to ad	Not exposed to ad
Not in the baseline	$(1-b_1)r_1e_1$ / $(1-b_1)r_1(1-e_1)$	0 / $(1-b_1)(1-r_1)$
In the baseline	b_1r_1 / 0	$b_1(1-r_1)$ / 0

In each cell the upper part is the probability that the consumer will move to the consideration stage (second stage in the funnel), while the lower part is the probability that he will leave from the system (exit the funnel).

As we can see, the fraction of people who move from the first stage to the second without seeing an ad is

$$b_1(1-r_1),$$

while the fraction of people who move from the first stage to the second after seeing an ad is

$$b_1r_1 + (1-b_1)r_1e_1.$$

Similarly, every consumer who enters the consideration stage in the funnel is exposed to a consideration ad with probability r_2. If a consumer in the second stage does not see a consideration ad, then the probability that he will purchase the product is b_2 (baseline probability). If a consumer is exposed to a consideration ad, but he was not going to purchase the product without the ad, then he purchases with probability e_2. We assume that r_2 and b_2 are fixed, while e_2 is a decision variable for the second publisher (effort). The cost for the effort e_2 is given by the convex function $c_2 e_2^2$

The fraction of people who move from the second stage to the purchase without seeing an ad is

$$b_2(1-r_2),$$

while the fraction of people moving from the second stage to the purchase after seeing an ad is

$$b_2r_2 + (1 - b_2)r_2e_2.$$

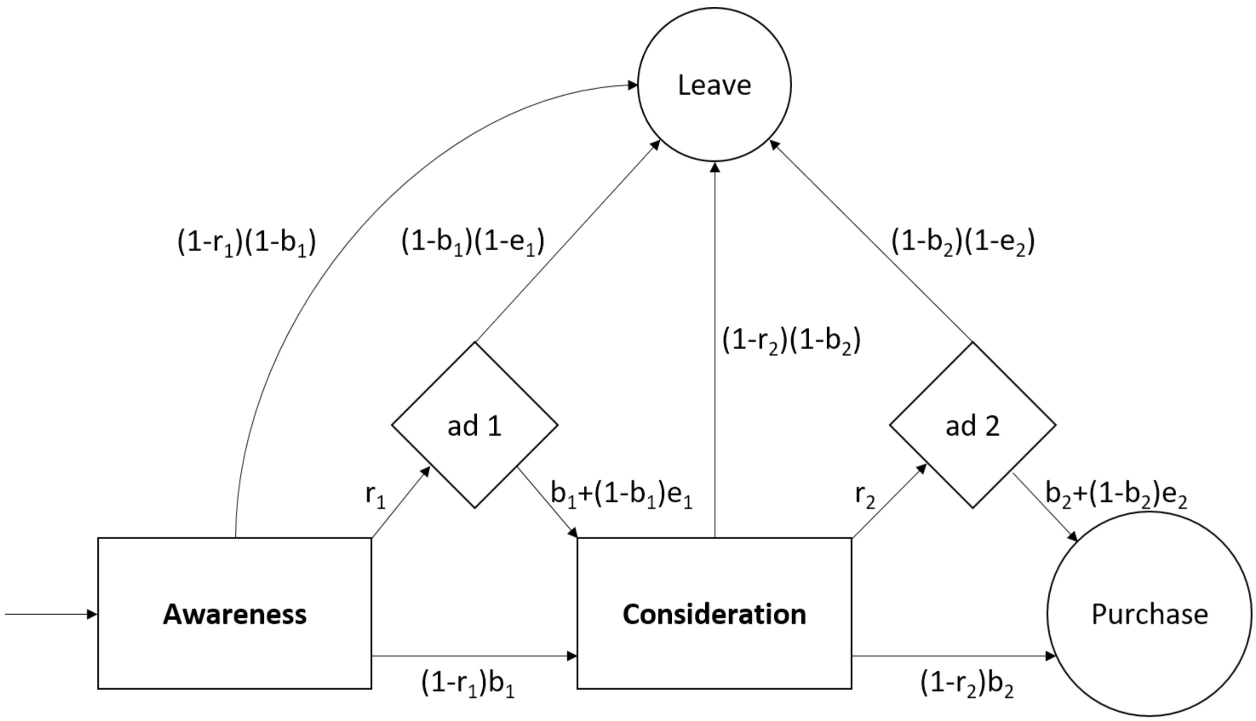


Figure 2.A.1: General model.

The firm can see all the consumers who purchased the product in its website and only them. It can also see, for each consumer who purchased, which ads they were exposed to prior to the purchase.

In other words, the firm can observe that a fraction $b_1(1-r_1)b_2(1-r_2)$ of consumers will purchase the product without seeing any ad, a fraction $(b_1r_1 + (1-b_1)r_1e_1)b_2(1-r_2)$ will purchase after seeing an ad only in the first stage, a fraction $b_1(1-r_1)(b_2r_2+(1-b_2)r_2e_2)$ will purchase after seeing an ad only in the second stage, and a fraction $(b_1r_1+(1-b_1)r_1e_1)(b_2r_2+(1-b_2)r_2e_2)$ will purchase after seeing ads in both stages. So the firm can determine these four quantities even though it doesn't know the individual efforts e_1, e_2 or the baselines b_1, b_2.

Transition from the general model to our model

To simplify the notation and make the analysis a bit easier, we make the following changes to the general model:

- We define the constant $q_0 = b_1(1 - r_1)$, which we will call baseline probability (even though the actual baseline is b_1). This is the probability that a consumer in the first stage will move to the second without seeing any ad.

- We define the decision variable $a = b_1 r_1 + (1 - b_1) r_1 e_1$. This is the probability that a consumer in the first stage will see an ad and then move to the second stage. So now we'll assume that the first publisher will have to decide a instead of deciding e_1. The cost of e_1 was $c_1 e_1^2$, which makes the cost of a something of the form $\xi(a - \psi)^2$, for constants ξ, ψ. This means that the first publisher will choose an a such that $a \geq \psi$. We will simplify this cost further to wa^2 for some constant w. This is without loss of generality because we can always approximate the cost of the original model by adjusting w, and make sure that $a \geq \psi$ in equilibrium. Below we prove this claim formally.

- Similarly, we define the constant $p_0 = b_2(1 - r_2)$; the baseline probability for the transition from the consideration stage to purchase.

- Finally, we define the decision variable $c = b_2 r_2 + (1 - b_2) r_2 e_2$ for the second publisher with an associated cost vc^2.

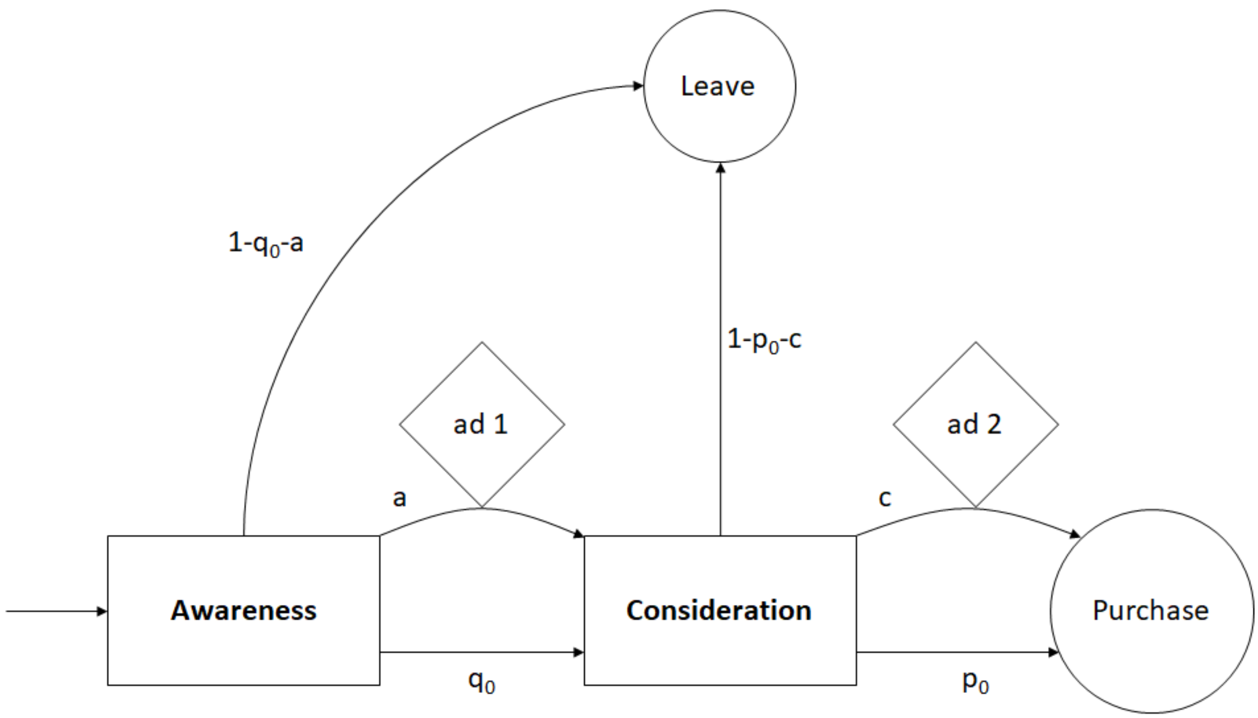


Figure 2.A.2: Simplified model.

In this version of the model, the firm can determine the quantities $q_0 p_0, a p_0, q_0 c$, and ac, even though it doesn't know the individual values of q_0, p_0, a, and c.

Claim 2.1. *The two models are equivalent.*

Proof. We will prove that for any equilibrium (e_1^*, e_2^*) in the general model, there are w, v such that there is an equilibrium (a^*, c^*) in the simplified model that satisfies $a^* = b_1 r_1 + (1 - b_1) r_1 e_1^*$ and $c^* = b_2 r_2 + (1 - b_2) r_2 e_2^*$.

We fix the values of $b_1, b_2, r_1, r_2, c_1, c_2$, and let $p_1(e_1, e_2)$ and $p_2(e_1, e_2)$ be the payment functions in the general model for the first and the second publisher respectively. Let (e_1^*, e_2^*) be an equilibrium of the general model under these payment functions.

We define $q_0 = b_1(1-r_1)$ and $p_0 = b_2(1-r_2)$. We also define the functions $g(e_1) = b_1r_1 + (1-b_1)r_1e_1$ and $h(e_2) = b_2r_2 + (1-b_2)r_2e_2$, and let $a' = g(e_1^*)$ and $c' = h(e_2^*)$. The payment functions in the simplified model will then be $p_1(g^{-1}(a), h^{-1}(c))$ and $p_2(g^{-1}(a), h^{-1}(c))$ for the first and the second publisher respectively. We set

$$w = \left.\frac{\partial p_1(g^{-1}(a), h^{-1}(c'))}{\partial a}\right|_{a=a'} \cdot \frac{1}{2a'}$$

and

$$v = \left.\frac{\partial p_2(g^{-1}(a'), h^{-1}(c))}{\partial c}\right|_{c=c'} \cdot \frac{1}{2c'}.$$

For these w and v, it holds that

$$\left.\frac{\partial(p_1(g^{-1}(a), h^{-1}(c')) - wa^2)}{\partial a}\right|_{a=a'} = 0$$

and

$$\left.\frac{\partial(p_2(g^{-1}(a'), h^{-1}(c)) - vc^2)}{\partial c}\right|_{c=c'} = 0.$$

Notice that $p_1(g^{-1}(a), h^{-1}(c')) - wa^2$ is the utility of the first publisher and $p_2(g^{-1}(a'), h^{-1}(c)) - vc^2$ is the utility of the second publisher in the simplified model. In other words, (a', c') is an equilibrium in the simplified model. □

Example 2.1. Let's consider a $\frac{1}{2}$-contract and let $b_1 = \frac{1}{3}, b_2 = \frac{1}{4}, r_1 = \frac{1}{5}, r_2 = \frac{1}{6}, c_1 = 2, c_2 = 3$. The equilibrium in the general model is $(e_1^*, e_2^*) = (0.00765193, 0.00626063)$. For $w = 1.69573$ and $v = 3.53964$, we get the equilibrium $(a^*, c^*) = (0.0676869, 0.0424492)$ in the simplified model, and it holds that $a^* = b_1r_1 + (1-b_1)r_1e_1^*$ and $c^* = b_2r_2 + (1-b_2)r_2e_2^*$.

2.A.3 Definition of Shapley Value

Let N be a set of players and let $v : 2^N \to \mathbb{R}$ be a function such that for every subset $S \subseteq N$ of players, $v(S)$ gives the total payoff the members of S will get by working together. For a pair (v, N), an attribution rule is a function $\phi_i(v)$ that gives the payoff of player i.

Axiom 2.1. *Symmetry*: If two players are equivalent, then they should have the same payoff. Two players i, j are equivalent if their contribution to every subset of other players is the same, or mathematically if $v(S \cup \{i\}) = v(S \cup \{j\})$ for every $S \subseteq N \setminus \{i, j\}$.

Axiom 2.2. *Null player*: The payoff of a null player should be 0. A player i is called null if he doesn't contribute anything to any subset of other players, or mathematically if $v(S \cup \{i\}) = v(S)$ for every $S \subseteq N \setminus \{i\}$.

Axiom 2.3. *Additivity*: The sum of payoffs that a player gets for two different games should be equal to the payoff he gets if we consider the two games as one big game. Or mathematically, $\phi_i(v + w) = \phi_i(v) + \phi_i(w)$ for every player $i \in N$ and any two functions $v, w : 2^N \to \mathbb{R}$.

Axiom 2.4. *Efficiency*: The total payoff is distributed among all the players. Or mathematically, $v(N) = \sum_{i \in N} \phi_i(v)$.

Shapley (1953) proved that there is a unique rule that satisfies these four axioms. We call this rule the *Shapley Value*, and it is given by the following formula

$$\phi_i(v) = \sum_{S \subseteq N \setminus \{i\}} \frac{|S|!(n - |S| - 1)!}{n!}(v(S \cup \{i\}) - v(S)).$$

Chapter 3

The Beneficial Effects of Ad Blockers[1]

While online advertising is the lifeline of many internet content platforms, the usage of ad blockers has surged in recent years presenting a challenge to platforms dependent on ad revenue. In this chapter, using a simple analytical model with two competing platforms, we show that the presence of ad blockers can actually benefit platforms. In particular, there are conditions under which the optimal equilibrium strategy for the platforms is to allow the use of ad blockers (rather than using an adblock wall, or charging a fee for viewing ad-free content). The key insight is that allowing ad blockers serves to differentiate platform users based on their disutility to viewing ads. This allows platforms to increase their ad intensity on those that do not use the ad blockers and achieve higher returns than in a world without ad blockers. We show robustness of these results when we allow a larger combination of platform strategies, as well as by explaining how ad whitelisting schemes offered by modern ad blockers can add value. Our study provides general guidelines for what strategy a platform should follow based on the heterogeneity in the ad sensitivity of their user base.

3.1 Introduction

3.1.1 Background

Online advertising is like taxes. Nobody likes them, but they exist because people understand that they are necessary. Millions of websites, including some of the largest internet companies, depend on advertising as their main source of revenue. Online advertising revenue in the US in 2015 was \$59.6 billion[2], almost half of it accounted for by Google.[3,4] Google,

[1]Based on joint work with R. Ravi and Kannan Srinivasan

[2]http://www.iab.com/wp-content/uploads/2016/04/IAB-Internet-Advertising-Revenue-Report-FY-2015.pdf (accessed May 2017)

[3]https://www.statista.com/statistics/266249/advertising-revenue-of-google/ (accessed May 2017)

[4]http://www.businessinsider.com/stats-on-googles-revenues-from-youtube-and-google-play-2015-7 (accessed May 2017)

Facebook[5], and Twitter[6] together make up more than 65% of the total revenue.[7] Advertising is the main source of revenue for all these companies. Another example of an industry that depends heavily on advertising is the US news industry, with 69% of its revenue coming from advertising.[8] More generally, advertising is the key reason many content-providing websites are able to offer their services to users for "free" (other than the implicit payment of user attention to the ads). In short, today's internet would not be what it is without advertising.

Of course advertising is not a new phenomenon. Even before the era of the internet, companies advertised products on billboards, newspapers, radio stations, TV channels, and other mass media. However, there is a key difference between advertising on the internet today and the other media. The interactive nature of the internet gives users the easy ability to block ads with ad blockers. An ad blocker is a type of software, usually added conveniently as an extension to an internet browser, that will prevent any ads from appearing on the browsed web pages. When a user with an ad blocker visits a website with ads, the blocker identifies the ad content and blocks it from loading. Consequently, the website does not receive any ad revenue for that user.[9]

Ad blocking is not something new either. After VCRs became popular in the 1980s, there was a trend among viewers for commercial skipping. To combat this, advertisers tried to make ads more entertaining. In 1999, ReplayTV launched the first DVR with a built-in feature to skip commercials.[10] Since then, providers of commercial skipping features have been plagued by lawsuits that claim damages to the copyright of the original content.[11] A difference between these precursors and ad blockers on the internet is that now it is easier than ever before to block ads, since several ad-blocking extensions are just a few clicks away in most browsers.[12]

In Figure 3.1, we can see how adblock usage is changing over the years for desktop and mobile devices. We observe a steady increase in both categories with an average of 44.8% increase for desktop and 63.9% increase for mobile per year. Even though mobile ad blockers were not as popular as their desktop counterparts in 2015, in the beginning of 2017 we see the opposite, since more than 380 million mobile devices have an installed ad blocker versus 236 million desktop devices. PageFair and Adobe (2015) estimated that the cost of ad blockers for publishers in terms of lost revenue in 2015 was $21.8 billion, which was around 14% of the global ad spend. Today with many more devices with ad blockers than in 2015 (Figure 3.1), we expect this number to be much larger.

[5] http://www.businessinsider.com/facebook-us-ad-revenue-2016-1 (accessed May 2017)

[6] https://www.emarketer.com/Article/Twitter-Ad-Revenues-Expected-Continue-Robust-Growth/1013571 (accessed May 2017)

[7] https://www.bloomberg.com/news/articles/2016-04-22/google-and-facebook-lead-digital-ad-industry-to-revenue-record (accessed May 2017)

[8] http://www.journalism.org/2014/03/26/the-revenue-picture-for-american-journalism-and-how-it-is-changing/ (accessed May 2017)

[9] See e.g., http://alternativeto.net/software/adblock-plus/ (accessed May 2017).

[10] https://en.wikipedia.org/wiki/Commercial_skipping (accessed May 2017)

[11] https://www.techdirt.com/articles/20160211/10423633579/dish-agrees-to-cripple-ad-skipping-dvr-to-settle-fox-lawsuit.shtml (accessed May 2017)

[12] https://adblockplus.org/getting_started (accessed May 2017)

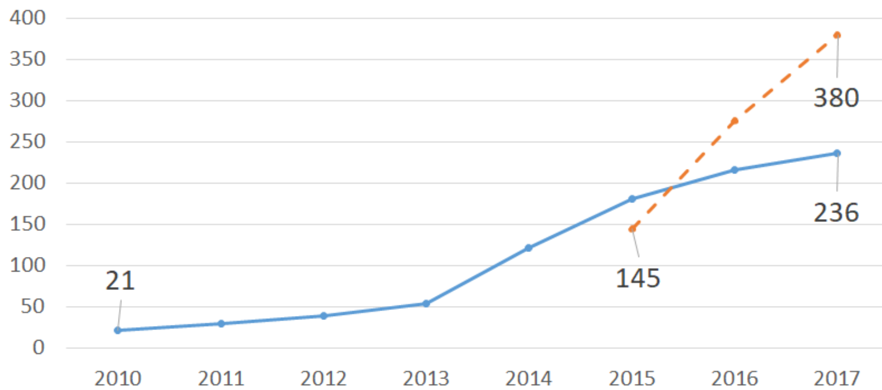


Figure 3.1: Millions of devices with adblock software over the years (— Desktop, - - Mobile). (PageFair, 2017)

3.1.2 How do Platforms Respond?

Websites hosting content and supported by ads act as platforms for gathering viewers and advertisers. Their revenue stream is directly affected by the deployment of ad blockers by the viewers. The response of these platforms to ad blockers have varied considerably.[13,14]

Some platforms disallow the use of ad-blocking software when viewing their sites, by using an **adblock wall**. This is the name for currently available technology that allows websites to detect if a visitor is using an ad blocker and if so, refuse to give access to him. Forbes is an example of a website that uses an adblock wall.[15] City A.M. was the first UK newspaper website to ban the use of ad blockers and prevent adblock users from reading content.[16]

Other platforms offer ad-free or ad-light subscription services for viewing content, by using a **paywall**. Financial Times[17], The Wall Street Journal[18], and The Washington Post[19] are a few examples of news sites with such a paywall. A slightly different but related strategy was adopted by Youtube: it was originally dependent solely on advertising, but in 2015 it launched YouTube Red, a subscription based service that offers ad-free access to all YouTube videos with some additional exclusive content.[20]

Many platforms use a combination of the aforementioned options. They use an adblock wall that offers two choices to the users, either to disable their ad blocker or pay a fee for an

[13] https://www.pubnation.com/blog/publishers-fight-back-how-the-top-50-websites-combat-adblock (accessed May 2017)

[14] http://adage.com/article/media/publishers-weigh-ways-fight-ad-blocking/299116/ (accessed May 2017)

[15] https://digiday.com/media/forbes-ad-blocking/ (accessed May 2017)

[16] https://www.theguardian.com/media/2015/oct/20/city-am-ban-ad-blocker-users (accessed May 2017)

[17] https://www.ft.com/products (accessed May 2017)

[18] https://subscribe.wsj.com (accessed May 2017)

[19] https://subscribe.washingtonpost.com (accessed May 2017)

[20] https://www.theverge.com/2015/10/21/9566973/youtube-red-ad-free-offline-paid-subscription-service (accessed May 2017)

ad-free version of the site. Some examples of websites using this strategy are Wired[21], Bild[22], and Business Insider.[23] The New York Times has also experimented with an adblock wall of this type for some time.[24] This option mirrors ad-free services that have been available in more traditional media, e.g. an alternate to watching a movie or show for free on network TV is to buy or rent an ad-free copy.

Finally, there are some platforms, like The Guardian, that request viewers to disable ad blockers as a gesture of support for the content in the site (without preventing access if they do not).[25] There are also sites that simply ignore the use of ad blockers and **allow** their use. In fact, the majority of the content providing websites in the internet today follow this simple strategy of doing nothing about the existence of ad blockers other than simply allowing their use.

3.1.3 Research Questions

Internet ad blockers motivate some fundamental questions: What is the optimal response of platforms to their presence? Are adblock walls the solution to the ad-blocker problem? Why should platforms ever allow ad blockers, if they can prevent them using a simple adblock wall? When should they erect a paywall and charge a fee for ad-free or ad-light content? Under what conditions should they use these different options? In this paper, we address the central questions above, and explore further the effects ad blockers have on platforms and users. For instance, if we compare a world without ad blockers and the current world with them, are the effects of ad blocking only negative for platforms? How do platforms' ad-revenues change with the availability of ad blockers? How is overall user welfare affected by ad blocking? To answer these questions, there are three important elements we model: competition, ad intensity, and heterogeneity in the ad sensitivity.

Competition. There are several reports that adblock walls do not work.[26] Several websites that implemented some type of adblock wall, like Wired, Bild, and Forbes, have seen their traffic deteriorate right after the introduction of the adblock wall.[27] The main explanation for this is that most adblock users who visit such websites and face a wall prefer to leave the website instead of disabling their ad blocker, even temporarily. In fact, in a survey by PageFair (2017), 74% of adblock users said that they leave websites when faced with an adblock wall, and only 26% disable their ad blockers to read the content.

Competition is the key reason for why adblock walls do not work. Most websites do not

[21] https://www.wired.com/how-wired-is-going-to-handle-ad-blocking/ (accessed May 2017)

[22] http://www.reuters.com/article/us-axelspringer-adblock-idUSKCN0S70S020151013 (accessed May 2017)

[23] http://adage.com/article/media/business-insider-testing-paywall-ad-blocking-response/305951/ (accessed May 2017)

[24] https://www.wsj.com/articles/new-york-times-experiments-with-ways-to-fight-ad-blocking-1457378218 (accessed May 2017)

[25] http://www.economist.com/news/business/21653644-internet-users-are-increasingly-blocking-ads-including-their-mobiles-block-shock (accessed May 2017)

[26] http://www.businessinsider.com/ad-blocking-walls-not-working-2016-2 (accessed May 2017)

[27] https://thestack.com/world/2016/04/21/sites-that-block-adblockers-seem-to-be-suffering/ (accessed May 2017). However, it is unclear whether the loss in traffic after an adblock wall implementation directly translates to loss in revenue.

offer unique content that users cannot find elsewhere. As a result, instead of disabling their ad blocker and facing the inconvenience of ads, users prefer to look for the same or similar content elsewhere.

Ad intensity. Websites can control how many ads they will show, how intrusive or annoying the ads will be, their size, their position, and so on. All these affect the user experience and how much disutility a user will get from the ads. As an example, Forbes.com, when presenting an adblock wall to a user, shows a message promising users that if they disable their ad blocker, they will be presented with an "ad-light" experience in return. In the survey by PageFair (2017), 77% of adblock users said that they were willing to view some ad formats and are not totally against ads. Therefore, ad intensity is a key decision for platforms, since it directly affects how users react.

Heterogeneity in ad sensitivity. The increase in the adoption of ad blockers has other reasons in addition to their ease of installation. Digital ads offering rich-media content such as audio, video, pop-ups and flashing banners have become increasingly intrusive to the content absorption experience. The rise of the mobile internet also puts a premium on the space available for content viewing that is jeopardized by ads that take up too much real-estate, mobile data consumption, and battery life. Finally, re-targeting practices associated with digital ads have increased the perception of privacy intrusion among viewers.

Nevertheless, the adoption of ad blockers among viewers is not likely to be universal since the sensitivity of viewers to ads is sufficiently heterogeneous across sites and devices from which the sites are accessed. Users that access the platforms from public or corporate machines may not have the ability to install new uncertified software such as ad blockers. Casual users who do not spend too much time on sites with annoying ads will not take the effort to employ ad blockers and update/maintain them. Less technical users may not even be aware of the existence and convenience of ad blockers. Many technically-savvy users may also continue to allow ads to support the sites they visit by acknowledging that they indirectly pay for the content they consume. Some users simply continue to view ads so as be kept informed of new products and promotions over time.

In a survey by PageFair and Adobe (2015), when non-adblock users were asked what would cause them to start using an ad blocker, 50% of the respondents stated that misuse of their personal information would be a reason to enable ad blocking. 41% of them responded that an increase in the number of ads from what they typically encounter today would also be a good reason. There was an 11% saying that they would never use an ad blocker, and this proportion increases to 23% for those aged between 35 and 49 years old. In contrast, when adblock users were asked for their main motivation behind adblock usage in PageFair (2017), only 6% of them stated privacy as the main reason. Security and interruption were the two leading reasons at 30% and 29% respectively, while page speed and the fact that there are too many ads came next with 16% and 14% respectively.

This provides evidence that there are fundamentally two classes of users based on whether they use ad blockers or not, and there is a lot of heterogeneity in the ad sensitivity of users in both classes. Furthermore, there is also difference in ad sensitivity between these two classes of adblock and non-adblock users.

3.1.4 Contributions

In this paper we devise a simple analytical model to answer the questions of Section 3.1.3. We model sites as two competing platforms for hosting content and attracting users. We assume two classes of users: one that uses ad blockers and the other that does not; note that the former are typically more ad-sensitive than the latter. Each platform has three options:

- BAN strategy: Continue displaying ads and ban ad blocking (e.g., using an adblock wall). If a viewer uses an ad blocker, he has to disable it to get access to the site.[28]

- ALLOW strategy: Continue to display ads and allow ad-blocking software by any user that installs it.

- FEE strategy: Stop displaying ads and offer only an ad-free site with a subscription fee.

Note that in the second option, the platform will make no revenue from adblock users, but only from those who do not use an ad blocker and can see ads.

Given that banning ad blockers is an option for both platforms, we would expect that this would always emerge as an equilibrium strategy since it would curb the loss of revenues compared to the allow strategy. However the competitive dynamics between even two symmetric platforms results in a surprising equilibrium. Our first result argues that **there are conditions where both platforms arrive at ALLOW as their symmetric optimal strategies** (Proposition 3.1 in Section 3.4). The intuition is that the action of installing ad blockers serves as a filter for more ad-sensitive users that employ ad blockers; with these users gone, each platform can move to a higher intensity of advertising to users and hence increase revenue.

While allowing ad blockers results in increased advertising by both platforms, we may expect the utility of users exposed to this increased advertising to decrease substantially as a result. However, our second result argues that *when platforms* ALLOW *ad blockers, this can increase the overall welfare of users* (Proposition 3.2 in Section 3.4). The result follows from the filtering effect, which can raise the utility of ad-sensitive users substantially by allowing them to filter ad content, overshadowing the potential loss of utility to less ad-sensitive users who might now be subject to more advertising. Perhaps even more surprisingly though, **there are cases where no user is worse off when ad blockers are allowed, while platforms and some users are better off** resulting in a Pareto-improvement in overall welfare as a result of introducing ad blockers.

In Section 3.5.1, we extend the main model by adding the following option for platforms:

- ADS OR FEE strategy: Give the choice to users to either disable their ad blockers and be exposed to ads, or pay a fee for an ad-free version of the site (e.g., using a paywall).

The argument in favor of this new strategy is that it can achieve the filtering effect that the ALLOW strategy had by making ad-sensitive users pay the subscription fee, while users with

[28] Another strategy some platforms use to bypass ad blockers is placing advertising content, e.g. mentions of products, that is organically mixed in with their native content. For the purposes of our model, this strategy can be considered the same as the BAN strategy, since it is just a different way to make adblock users see ads.

lower ad sensitivity see ads. However, we show that **even with the addition of the Ads or Fee strategy, there is still an equilibrium where both platforms Allow ad blockers** (Proposition 3.3). To show this, we further split the class of non-adblock users into two further classes with different ad sensitivities. In this context, Ads or Fee is a better strategy for platforms when there is heterogeneity in the ad sensitivity between the two classes of non-adblock users, because it helps separate very ad-sensitive non-adblock users from the rest, while Allow is a better strategy when the two non-adblock user classes are more homogeneous in their ad sensitivity.

In Section 3.5.2, we extend the main model in a different direction by adding the following option for each platform:

- Whitelist strategy: Allow ad blockers and pay a fee to the ad-blocker company to put the platform in the default whitelist.

This strategy is motivated by the "Acceptable Ads initiative"[29]. This is a program introduced by Adblock Plus, the most popular adblock extension, according to which publishers and advertisers who comply with certain criteria could get whitelisted so that their ads may pass through the filter of the ad blocker. This strategy seems to be inferior to the previous strategies, since this option simply requires platforms to pay for something they had for free before the advent of ad blockers. However, we show that **there is an equilibrium where platforms use the Whitelist strategy and this equilibrium can sometimes increase their revenue even more than when ad blockers did not exist** (Proposition 3.4). For this, we now split the class of adblock users into two further classes with different ad sensitivities. In this context, Whitelist is a better strategy when there is heterogeneity in the ad sensitivity between the two classes of adblock users, since it can help platforms separate ad-sensitive adblock users from the rest, while Allow is a better strategy when the two adblock user classes are more homogeneous in their ad sensitivity.

Finally, in Section 3.6, we extend the main model to include content creators who generate the content of the platforms and share the revenue with them. We show the robustness of our earlier results in this extension; we also show that allowing ad blockers can result in an increased quality of content. This provides an additional benefit for users when ad blockers are allowed.

3.2 Literature Review

Our paper is related to the advertising and marketing avoidance literature (Clancey, 1994; Cho and Cheon, 2004; Speck and Elliott, 1997; Li and Huang, 2016; Seyedghorban et al., 2016). Below we discuss some of the more closely related papers.

Anderson and Gans (2011) consider a model of a content provider who chooses a level of advertising while consumers decide if they will adopt ad-avoidance technology or not. They show that ad-avoidance penetration can increase advertising clutter, but it decreases the content provider's profit. One difference with our setting is that in their model, there

[29] https://adblockplus.org/acceptable-ads

is a price for consumers to adopt ad-avoidance technology.[30] As a result, their setting is more appropriate for more traditional ways of ad avoidance, like DVRs or other physical appliances, where there is a non-zero sunk cost for their adoption. In our setting, where the most popular ad blockers are free of charge and they are just a few clicks away to install in most browsers, this assumption is not as realistic. Another difference is that we consider a model with competition where platforms can actually decide if they will allow ad blocking or not. This is a more appropriate model for a website trying to decide if they will implement an adblock wall or not, instead of just assuming that adblock usage is unavoidable as in their model.

Johnson (2013) examines a model with firms that can target their ads to consumers and consumers who can avoid advertising. He shows that improved targeting can benefit firms but not necessarily consumers. He also shows, that in equilibrium, consumers may under-utilize their ability to block ads. A difference with our model is that there is a direct link between advertising firms and consumers, where firms can target the consumers with the higher probability of buying their product. In other words, there is no intermediate publisher who takes part in the decision process. He also assumes that there is some (positive) cost to consumers for avoiding ads and that the firm has a cost for sending an ad regardless of whether the ad is avoided or not. All these make his setting a better model for more traditional direct advertising campaigns, like direct mail with intentional avoidance by consumers.

Hann et al. (2008) take a different approach by focusing more on the privacy of consumers. In their setting, sellers market their products to consumers through solicitations. Consumers have two ways to avoid solicitations: either by concealment (e.g. registering in a do-not-call list) or by deflection (e.g. with call screening). There are also two types of consumers, consumers with high demand and those with low demand for the products. They show that concealment by low-demand consumers can lead sellers to market more, while the opposite is true when there is concealment by high-demand consumers. They also show that concealment is worse for consumer welfare than deflection.

Wilbur (2008) studies a two-sided empirical model of the television industry with advertisers on one side and viewers on the other. One of his counterfactual findings is that ad avoidance tends to increase advertising quantities and decrease network revenues. Goh et al. (2015) investigate the externalities imposed by consumers who avoid ads on other consumers in the context of the US Do Not Call registry (DNC). They found that the number of subsequent DNC registrations was positively correlated with the number of first wave registrations. This suggests that perhaps telemarketers increased the number of calls to unregistered consumers after the first wave driving even more subsequent registrations.

Aseri et al. (2017) study a similar problem to ours, namely the benefits of ad blockers, in a different setting. They consider a monopolistic platform who might not want to ban ad blockers because of network effects among its users. They also assume that the platform can choose a different ad intensity for each type of user (e.g. show fewer ads to adblock users who whitelisted the site), which is an additional discriminatory tool platforms can use for their

[30] In that sense, ad avoidance from the perspective of consumers is more like the Ads or Fee plan of our model, but the content provider does not receive the fee. This can explain some of their results when viewed in our framework.

benefit. The difference with our paper is that we consider competition between platforms and we show that even when network effects are not present, and also platforms cannot directly discriminate users based on their type, ad blockers can still benefit them (due to a competition-softening effect).

Several papers also study settings where a media provider has to decide between an ad-based and a subscription-based strategy (Prasad et al., 2003; Peitz and Valletti, 2008; Tåg, 2009; Stühmeier and Wenzel, 2011; Vratonjic et al., 2013). Armstrong et al. (2009) study consumer protection policies and their impact on the consumers' incentives to become informed of market conditions. They show that when consumers are able to refuse marketing, price competition can decrease, which can harm consumers. Spam filters can also be considered a form of ad avoidance. Falkinger (2008) studies the equilibria in a model about spam filters with different levels of tolerance. In relation to ad annoyance, Goldstein et al. (2014) study the costs of annoying ads to publishers and users.

To the best of our knowledge, none of the previous work has considered ad avoidance from a perspective of a publisher who has the ability to prevent or limit it in a competitive setting. This is because prior work focused on traditional media providers, like TV stations, or direct marketing actions, like mail, calls, or email. Our setting, on the other hand, is inspired by web ads, ad blockers, and the available anti-adblock technology used by many websites today.[31] We extend the strategy space of publishers to include the most popular responses to ad blockers by websites, like adblock walls, pay walls, combinations of the two, allowing ad blockers, or paying for whitelisting services. This leads to quantitatively different results, with the general surprising conclusion that ad blockers can actually benefit both publishers and users in several different ways.

Outside the ad-avoidance literature, our results are also related (in terms of model mechanics) to price discrimination and price sensitivity (Corts, 1998; Desai, 2001; Desai and Purohit, 2004; Coughlan and Soberman, 2005; Pazgal et al., 2013). As an example of a related result, Jain (2008) uses a model of digital piracy to show that when more price-sensitive consumers are the ones who copy software, then piracy can help firms. Shaffer and Zhang (1995) study the effects of price-discriminating customers by offering promotions based on their past purchase behavior and show that this can reduce the profits of the firms in a competitive environment.

3.3 Model

In this section, we describe the main model that we will use for the results in Section 3.4. In Sections 3.5.1 and 3.5.2, we extend this basic model by adding additional strategies to platforms' strategy space as well as additional segments of users.

3.3.1 Platform Model

There are two platforms, platform 1 and platform 2, competing over a set of users. Each platform can choose one of three different strategies, Ban, Allow, or Fee.

[31] https://www.wsj.com/articles/the-rise-of-the-anti-ad-blockers-1465805039 (accessed May 2017)

In the BAN strategy, the platform bans ad blockers by using an adblock wall. If a user with an ad blocker wants to access the site, she has to disable the ad blocker and see ads. The decision variable for a platform $i \in \{1, 2\}$ with the BAN strategy is the ad intensity $a_i \geq 0$. The revenue of a platform i with the BAN strategy is $r_i =$ (mass of users who pick platform i) $\cdot$ a_i.[32]

In the ALLOW strategy, the platform allows the use of ad blockers. In this case, a user with an ad blocker can access the site without seeing any ads and the platform does not get any ad revenue from them. The decision variable for a platform $i \in \{1, 2\}$ with the ALLOW strategy is again the ad intensity $a_i \geq 0$. The revenue of a platform i with the ALLOW strategy is now $r_i =$ (mass of users who pick platform i and see ads) $\cdot a_i$.

In the FEE strategy, the platform offers content without any ads using a paywall. A user who wants to access the content has to pay a subscription fee for it. The decision variable for a platform $i \in \{1, 2\}$ with the FEE strategy is now the subscription fee $p_i \geq 0$. The revenue of a platform i with the FEE strategy is $r_i =$ (mass of users who pick platform i) $\cdot p_i$.

3.3.2 User Model

We model users using a Hotelling line. We assume that the two platforms are positioned on the two endpoints of the interval $[0, 1]$. Each user draws a value x uniformly at random from $[0, 1]$ that indicates her position in the interval. Users who are closer to a platform prefer that platform more than the other.

Any user's utility consists of three parts. The first part is some intrinsic value they have for accessing the platforms' service, e.g. for reading news, and it is independent of the platform. We use the variable m to indicate this value. The second part is some intrinsic value each user has for the platform. This is where the Hotelling model is used. For a user at position x, this intrinsic value is $1 - x$ if they pick platform 1, and x if they pick platform 2. The third part is the disutility a user gets either from ads they have to see or from the price they have to pay when they choose a platform. We normalize the price sensitivity of users to 1 and we let their ad sensitivity vary. Throughout the paper we will see how the heterogeneity in the ad sensitivity between users can affect how platforms behave.

In the basic model, we assume that there are two segments of users. The first segment consists of users without an ad blocker.[33] This segment is of mass λ and its users have ad sensitivity β. The second segment consists of users with an ad blocker who use it whenever possible. This second segment is of mass μ and its users have ad sensitivity γ.[34] In other

[32] This simple form of revenue proportional to the ad intensity best models display advertising where the platform is compensated proportional to the number of ads shown under a cost-per-mille (thousand impressions) or CPM payment scheme. However, similar arguments as those in the paper can be used for more involved revenue schemes like CPC/cost-per-click and CPA/cost-per-action. The logic then is that removing more ad-sensitive users from the market can increase click-through-rates for platforms.

[33] As mentioned earlier, some possible reasons that these users do not use an ad blocker are either ethical/moral, or because they do not know how to use one, or because they want to support the sites they visit.

[34] Note that λ and μ in this model do not depend on the ad intensities a_1 and a_2, i.e. the decision for users if they will install or not an ad blocker is exogenous. We can endogenize this decision by assuming that there is some cost for installing an ad blocker (e.g. learning cost) and letting users decide if they want to pay this cost based on their ad sensitivity and the ad intensity of the platform they visit. Even though

words, in this basic model, we assume some heterogeneity in the ad sensitivity between non-adblock and adblock users. Later, in extensions of the model, we will explore what happens when there is further heterogeneity in the ad sensitivity inside the segment of non-adblock users and inside the segment of adblock users.

Figure 3.1 summarizes the utility expressions for a user at position x who picked platform 1, based on the strategy the platform chose and the type of the user. To get the user utility for platform 2, we need to change $(1-x)$ to x, a_1 to a_2, and p_1 to p_2. Note the difference in ad sensitivity between the two types of users, and also that the adblock users do not suffer any disutility when the platform chooses the ALLOW strategy.

Platform 1	BAN	ALLOW	FEE
ADS Non-adblock users (mass λ)	$m + (1-x) - \beta \cdot a_1$	$m + (1-x) - \beta \cdot a_1$	$m + (1-x) - p_1$
ADS Adblock users (mass μ)	$m + (1-x) - \gamma \cdot a_1$	$m + (1-x)$	$m + (1-x) - p_1$

Figure 3.1: Utility of the user at position x who picks platform 1 based on the platform's strategy (columns) and the type of the user (rows).

Each user can pick at most one platform. We also assume that m, the intrinsic utility for the service, is large enough so that every user picks at least one of the platforms.[35]

3.3.3 Information Setting and Timeline

For simplicity we assume that all parameters are common knowledge. This is because any information uncertainty in the model would add extra complications without necessarily adding any insights regarding the effects of ad blockers. Moreover, in reality, even if a platform does not know the size of each segment of users or their ad sensitivity, there are ways to estimate these quantities. For example, they can use A/B testing with varying ad intensity to observe how users respond.

The timeline of the game is as follows. First, platforms choose the strategy they want to follow. This is the first step as this is the major decision platforms have to make. For example, using ads or subscription fees as their main business model usually means a different infrastructure for their website. Second, platforms decide the values of their decision variables

the analysis becomes more complicated, we can show that the main results of the paper remain true (for different conditions). There are two reasons to avoid this direction. One is that to truly endogenize user's decision we need to consider other factors that may play a role in that decision, e.g. privacy concerns, but this is beyond the scope of this paper. The second and more important reason is that users usually do not decide to install an ad blocker based on the ad intensity of a single website. They are either installing an ad blocker and use it almost everywhere, no matter the ad intensity of each website, or they do not install one and see ads. Therefore, it is more realistic to assume that the decision is exogenous.

[35] This assumption, which is standard in Hotelling models, reduces the number of cases we need to analyze. When it is not true, i.e. when there are users in the middle of the Hotelling interval who do not pick any platform, there is no interaction between platforms making each platform act as a monopoly in their own part of the market. The fact that there is no competition then leads to less interesting results regarding ad blockers.

(either ad intensity or price, depending on the strategy), as this is an easier decision to adjust. Third, users pick which platform to join based on the plan each platform offers so as to maximize their utility.

3.3.4 Benchmark

In this section, we consider a world without ad blockers as a benchmark to compare the performance of the above model with ad blockers. This way, we can determine the effects of the presence of ad blockers on platforms and users.

When ad blockers do not exist, platforms can choose one of two different strategies: the Ads plan or the Fee plan. A platform with the Ads plan offers its content for free to the users and its revenue comes from showing ads to them. In that case, since there are no ad blockers, every user has to be exposed to ads. A platform with the Fee plan offers its content without ads for a subscription fee.[36]

		Platform 2: Ads	Platform 2: Fee
Platform 1	Ads	$\frac{(\lambda+\mu)^2}{2(\beta\lambda+\gamma\mu)}, \frac{(\lambda+\mu)^2}{2(\beta\lambda+\gamma\mu)}$	$\frac{(\lambda+\mu)^2}{2(\beta\lambda+\gamma\mu)}, \frac{\lambda+\mu}{2}$
	Fee	$\frac{\lambda+\mu}{2}, \frac{(\lambda+\mu)^2}{2(\beta\lambda+\gamma\mu)}$	$\frac{\lambda+\mu}{2}, \frac{\lambda+\mu}{2}$

Table 3.1: Payoff matrix in the benchmark model.

If we analyze the game between the two platforms, we get the payoff matrix in Table 3.1.[37] In this game there are two symmetric equilibria, one where both platforms choose the Ads option and the other where both platforms choose the Fee option.[38] In Figure 3.2, we see the parameter regions of these two equilibria as a function of β (the ad sensitivity of the first segment of non adblock users) and $\frac{\gamma}{\beta}$ (the ratio of the ad sensitivities of the adblock users to that of the non-adblock users). As expected, when the ad sensitivities β and γ of users are low, platforms choose to show ads, while in the opposite case they decide to offer a subscription fee. The curve that separates the two equilibria is the line $\lambda + \mu = \beta\lambda + \gamma\mu$. The (Ads, Ads) strategy profile is an equilibrium iff $\lambda + \mu \geq \beta\lambda + \gamma\mu$, while the (Fee, Fee) strategy profile is an equilibrium iff $\lambda + \mu \leq \beta\lambda + \gamma\mu$.

3.4 Ad Blockers can be Beneficial

In this section we analyze the basic model and show how ad blockers can be beneficial for platforms in Subsection 3.4.1 and for users in Subsection 3.4.2.

[36]The Ads plan is similar to the Ban plan of the main model and the Fee plan is the same as the one in the main model.

[37]For the proof, see the first part of the proof of Proposition 3.1 in Appendix 3.A.1.

[38]Asymmetric equilibria can occur only in the degenerate case where $\lambda + \mu = \beta\lambda + \gamma\mu$. In that case, all possible pair of strategies give the same revenue to the platforms. Since this is a region of measure zero in the parameter space, we ignore it for the remainder of the discussion.

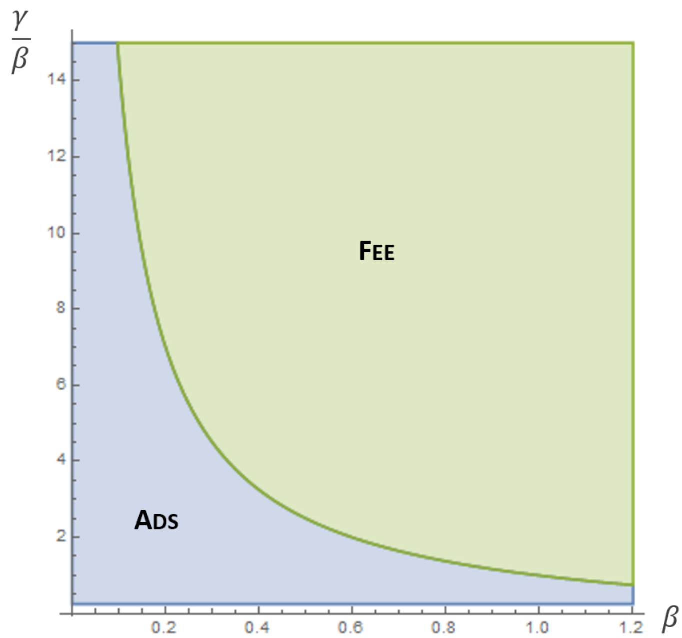


Figure 3.2: Equilibria regions of the benchmark model for $\lambda = 1$ and $\mu = 2$.

3.4.1 Platforms' Welfare

Our first proposition shows that there is an equilibrium where both platforms allow ad blockers. In this equilibrium, even though both platforms get no revenue from adblock users and can ban ad blockers if they want, they still allow adblock users to access the content for free. Moreover, the revenue of the platforms when they allow ad blockers is sometimes higher than their revenue when they ban ad blockers or when they use a fee. As a result, the presence of ad blockers can make platforms better off compared to the benchmark model. (All proofs are in Appendix 3.A.1.)

Proposition 3.1. *There is an equilibrium where both platforms* ALLOW *ad blockers. In this equilibrium, when β is sufficiently low and $\frac{\gamma}{\beta}$ is sufficiently high, platforms are better off than they would be if ad blockers did not exist.*

The intuition for this result is as follows. Let's assume that the ad sensitivity γ of the adblock users is larger than β, the ad sensitivity of non-adblock users (Figure 3.1). When platforms ban ad blockers, they show ads to both segments of users. However, the competition between the two platforms for the ad-sensitive segment will drive the optimal ad intensity of both platforms down. As a result, platforms can end up with low ad-revenue. On the other hand, when platforms allow ad blockers, they do not get any revenue from the segment of adblock users, but they have to compete only for the segment of non-adblock users that are less ad-sensitive. This allows them to increase the advertising intensity, which can result in higher ad-revenue.

The higher the difference in the ad sensitivities of the two segments, the more incentive platforms have to filter ad-sensitive users from the market and focus only on the users with low ad sensitivity. As a result, higher $\frac{\gamma}{\beta}$ makes the ALLOW option more attractive to

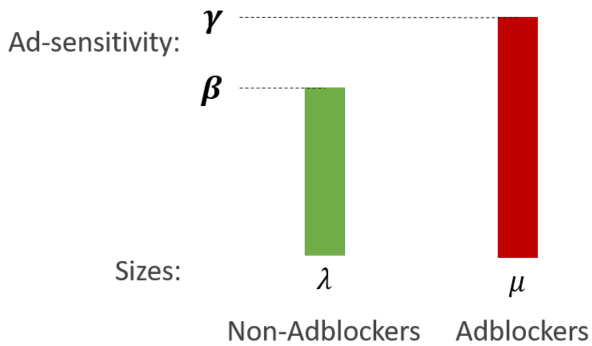


Figure 3.1: Illustration of the ad sensitivities of the two segments of users in the main model.

platforms than the Ban option. Moreover, the lower the ad sensitivity of those who see ads in the Allow strategy (i.e., low β), the more attractive the Allow plan becomes over the Fee plan. This gives the two conditions in Proposition 3.1.

If we analyze all possible pair of strategies for the two platforms, we get the payoff matrix in Table 3.1. As in the benchmark model, there are three symmetric equilibria in this game, one where both platforms ban ad blockers, one where they allow ad blockers, and one where they choose a fee. A difference is that now there are regions in the parameter space with more than one equilibria. In Figure 3.2, we can see the equilibria regions as a function of β (the ad sensitivity of non-adblock users) and $\frac{\gamma}{\beta}$ (the ratio of the ad sensitivity of adblock users over the ad sensitivity of non-adblock users). In regions with more than one equilibria, we list all of them and the first one in the list is the best one for platforms.

	Ban	Allow	Fee
Ban	$\frac{(\lambda+\mu)^2}{2(\beta\lambda+\gamma\mu)}, \frac{(\lambda+\mu)^2}{2(\beta\lambda+\gamma\mu)}$	$\frac{(3\lambda+2\mu)^2(\beta\lambda+\gamma\mu)}{2(3\beta\lambda+4\gamma\mu)^2}, \frac{\lambda(3\beta\lambda+\beta\mu+2\gamma\mu)^2}{2\beta(3\beta\lambda+4\gamma\mu)^2}$	$\frac{(\lambda+\mu)^2}{2(\beta\lambda+\gamma\mu)}, \frac{\lambda+\mu}{2}$
Allow	$\frac{\lambda(3\beta\lambda+\beta\mu+2\gamma\mu)^2}{2\beta(3\beta\lambda+4\gamma\mu)^2}, \frac{(3\lambda+2\mu)^2(\beta\lambda+\gamma\mu)}{2(3\beta\lambda+4\gamma\mu)^2}$	$\frac{\lambda}{2\beta}, \frac{\lambda}{2\beta}$	$\frac{9\lambda(\lambda+\mu)^2}{2\beta(3\lambda+4\mu)^2}, \frac{(\lambda+\mu)(3\lambda+2\mu)^2}{2(3\lambda+4\mu)^2}$
Fee	$\frac{\lambda+\mu}{2}, \frac{(\lambda+\mu)^2}{2(\beta\lambda+\gamma\mu)}$	$\frac{(\lambda+\mu)(3\lambda+2\mu)^2}{2(3\lambda+4\mu)^2}, \frac{9\lambda(\lambda+\mu)^2}{2\beta(3\lambda+4\mu)^2}$	$\frac{\lambda+\mu}{2}, \frac{\lambda+\mu}{2}$

Table 3.1: Payoff matrix in the main model.

If we examine the plot in Figure 3.2 from the bottom towards the top, we see that when $\frac{\gamma}{\beta}$ is low, i.e. when adblock users are less ad-sensitive than non-adblock users or when the ad sensitivity of the two segments is similar, both platforms prefer to ban ad blockers. However, as $\frac{\gamma}{\beta}$ increases, the Allow option becomes more and more attractive for platforms. Thus, first we get a region where both Ban and Allow are equilibria but Ban is the better one for platforms, then a region where both are equilibria but Allow is better, and finally a region where Allow is the unique equilibrium.

If we now review the plot from the right towards the left, we see that as β (the ad sensitivity of non-adblock users, who see ads when platforms allow ad blockers) becomes lower, the Allow option becomes more and more attractive to platforms than the Fee option. This is because the lower the β, the more the platforms can increase the ad intensity, and as a result their ad revenue.

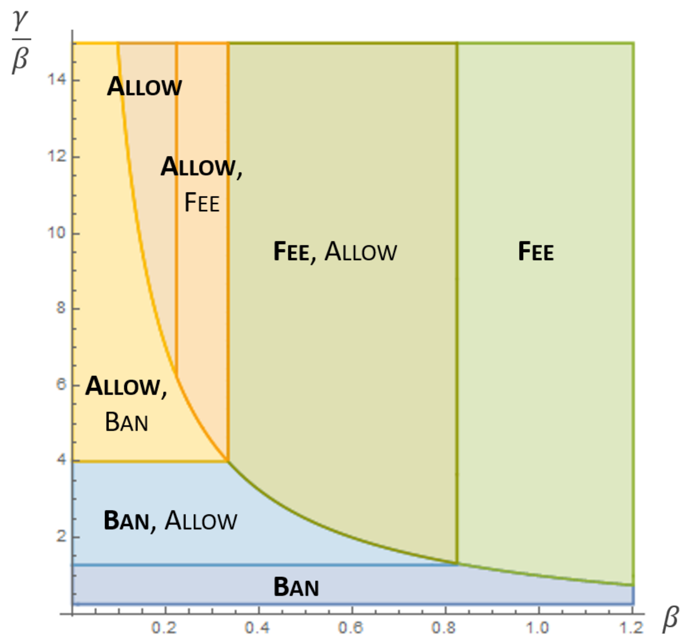


Figure 3.2: Equilibria regions of the main model for $\lambda = 1$ and $\mu = 2$. When there are two equilibrium strategies in the same region, the first one (in bold) is better for platforms.

3.4.2 User Welfare

The next proposition shows that total user welfare goes up when platforms allow ad blockers. In other words, not only do platforms benefit from the presence of ad blockers, but users could benefit too. Moreover, there are cases where total user welfare goes up and no user is worse off compared to the benchmark.

Proposition 3.2. *When both platforms allow ad blockers, total user welfare is higher than in the other two equilibria. Moreover, there are regions in the parameter space where platforms are better off, total user utility goes up, and no user is worse off compared to a world without ad blockers.*

The total user utility in the BAN equilibrium is the same as the one in the FEE equilibrium, while in the ALLOW equilibrium it is higher. The reason is that when both platforms allow ad blockers the segment of adblock users get no disutility from ads since they block them. This improves the overall user utility even if non-adblock users are sometimes worse off because they have to see more ads.

The regions mentioned in the second part of Proposition 3.2 are the regions labeled "ALLOW" and "ALLOW, FEE" in Figure 3.2. In these same two regions in the benchmark model, there is only one equilibrium where both platforms use a subscription fee. The disutility non-adblock users get from the fee in the benchmark model is the same as the disutility they get from ads when platforms allow ad blockers. As a result, their utility in these regions is the same as it was in the benchmark, while every other user and the platforms are better off.

3.5 Additional Plans

In this section, we investigate the effect of adding two different options to the strategy space of the platforms: an Ads or Fee option that lets users choose between watching ads or paying for an ad-free plan, and a Whitelist option to paying a fee to the ad blocker to whitelist their ads among users employing the ad blocker. Even when the Ads or Fee option is available to platforms, we demonstrate that our main counterintuitive finding that the Allow option continues to be an equilibrium still holds in certain parameter regions. When Whitelist is allowed, we show that this is an equilibrium option for both platforms, and perhaps even more surprisingly, despite the payment to the ad blocker, this improves their revenues compared to a world with no ad blockers at all. To show these results, we refine the set of adblock users or non-adblock users into two further subsegments with differing ad sensitivities.

3.5.1 The Ads or Fee Plan

The reason the Allow plan can benefit platforms in the basic model is that it provides a natural way for them to "discriminate" users with different ad sensitivities. Ideally platforms would like to be able to choose a different ad intensity for segments of users with different ad sensitivity. When they cannot do that, ad blockers provide an exogenous mechanism to achieve a similar effect. Ad-sensitive users self select themselves out of the market by using ad blockers, and this way they help not only themselves but also the competing platforms.

However, there is another natural way to "discriminate" users without the help of ad blockers that a lot of web sites currently use.[39] This is by letting users choose between two different options: either get free access to the site with ads (no ad blockers are allowed) or pay a fee for an ad-free version of the site (behind a paywall). We call this new plan the Ads or Fee plan.

The Ads or Fee plan is a combination of the Ban and the Fee plan from the basic model that tries to achieve the best of both worlds. The rationale is that users who are not very ad-sensitive will decide to see ads, while ad-sensitive users will choose the fee option. This solves the problem the Ban strategy had in the basic model of the ad-sensitive segment forcing platforms to decrease ad intensity. With the Ads or Fee strategy platforms can choose a high ad intensity for non-adblock users and also make sensitive adblock users pay a fee to access the content. Thus, this new strategy has the benefits of the Allow strategy plus some potential extra revenue from adblock users that platforms could not get earlier by allowing ad blockers.

The main question we want to answer in this section is whether the Ads or Fee strategy always dominates the benefits of allowing ad blockers. In other words, does the addition of the Ads or Fee plan wipe out the beneficial effects of the Allow strategy and prevents it from ever becoming an equilibrium. To answer this question, we extend the model from Section 3.3 by adding the Ads or Fee strategy to platforms' strategy space. If a platform chooses this strategy they will have two decision variables, an ad intensity a_i and a price p_i. Users who pick a platform with the Ads or Fee plan will choose between being exposed

[39] Some notable examples are Wired, Bild, and Business Insider.

to ads or paying the price, based on which option gives them higher utility.

To get a more refined view of the user ad sensitivities, we also add a third segment of users to the model in this case. This segment has mass ν and is made of non-adblock users with ad sensitivity η with $\beta \leq \eta \leq \gamma$ (Figure 3.1). The reason we add this third segment of users is that the comparison of the ADS OR FEE plan with the ALLOW plan will depend on how heterogeneous the ad sensitivity of non-adblock users is, which we refine by splitting into two subsegments. Then, by allowing $\frac{\eta}{\beta}$ to change, we can compare the two plans.

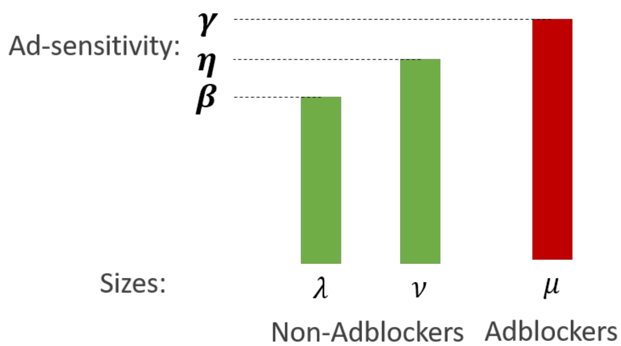


Figure 3.1: Illustration of the ad sensitivities of the three segments of users in the model with the additional ADS OR FEE strategy.

The following proposition shows that even after we add the ADS OR FEE strategy to the game, ad blockers can still be beneficial for platforms. There is still an equilibrium where platforms allow ad blockers, while sometimes this is the unique equilibrium and sometimes it is the best among others, including one where both platforms choose the ADS OR FEE plan.

Proposition 3.3. *When we add the* ADS OR FEE *plan to platforms' strategy space, there is still an equilibrium where both platforms* ALLOW *ad blockers. In this equilibrium, platforms are sometimes better off than in a world without ad blockers. There are regions in the parameter space where this is the unique equilibrium and regions where it is the best equilibrium for platforms among others.*

The reason the ADS OR FEE strategy does not always dominate the ALLOW strategy and there are still cases platforms get higher revenue by allowing ad blockers is the following. When a platform allows ad blockers, it will get revenue only from non-adblock users (first and second segment in Figure 3.1) by showing them ads. When the same platform chooses the ADS OR FEE plan, it gets some revenue from all three segments but the two most ad-sensitive segments will choose to pay the fee instead of seeing ads. Even though the ad intensity in the ADS OR FEE plan is higher than the one in the ALLOW plan (because the platform shows ads to less ad-sensitive users in the first case), when the ad-revenue the platform gets from the ALLOW plan is sufficiently higher than the revenue it gets from the fees in the ADS OR FEE plan, ALLOW is the better option. In other words, there are cases where the platform prefers the second segment of non-adblock users to see ads instead of paying the fee, because they can generate more revenue from ads. In the ADS OR FEE plan though, the platform cannot force these users to see ads instead of paying the fee, because users do what is best for them. This is why in these cases ALLOW is the best option for platforms.

We now describe the conditions under which ALLOW is the best strategy for platforms. For the ALLOW strategy to be better than FEE, we want the users who see ads in ALLOW to have relatively low ad sensitivity, i.e. we want low β and η. To make ALLOW better than BAN, we want the ad sensitivity of those who see ads in BAN but not in ALLOW to be higher than those who see ads in both, i.e. we want high $\frac{\gamma}{\beta}$ and $\frac{\gamma}{\eta}$. Finally, for ALLOW to be better than ADS OR FEE, we want those who see ads in ALLOW but do not see ads in ADS OR FEE to have similar ad sensitivity as those who see ads in both, i.e. we want low $\frac{\eta}{\beta}$. This is because otherwise ADS OR FEE would be the better option to separate non-adblock users of different ad sensitivities. In other words, for ALLOW to be the best strategy, we want the subsegments of non adblock users to be nearly homogeneous in their sensitivities and well separated from the sensitivity of the adblock users.

After we analyze all possible strategy combinations for the two platforms, we obtain the payoff matrix in Table 3.A.1. As before, there are four different symmetric equilibria, one for each strategy that is available to the platforms. In Figures 3.2, 3.3, and 3.4, we can see the equilibria regions for different parameters. To make the pictures a bit simpler, when there are more than one equilibria we list the best one for platforms. Thus, the regions where ALLOW is the unique equilibrium are subregions of the regions labeled ALLOW in the plots and near the borders there are multiple equilibria, one for each region that shares the border.

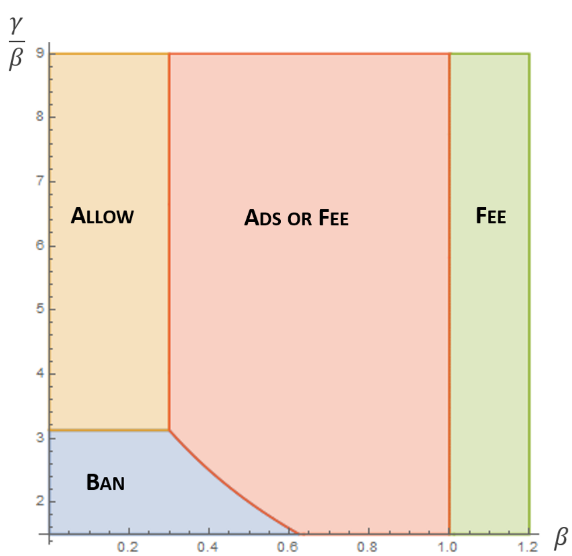


Figure 3.2: Equilibria regions of the model with the ADS OR FEE strategy for $\lambda = \mu = \nu = 1$ and $\frac{\eta}{\beta} = \frac{3}{2}$. When there are more than one equilibria, only the best one for platforms is listed.

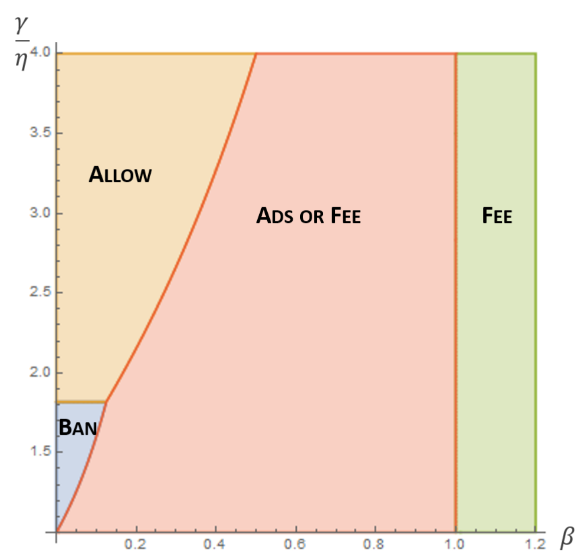


Figure 3.3: Equilibria regions of the model with the ADS OR FEE strategy for $\lambda = \mu = \nu = 1$ and $\frac{\gamma}{\beta} = 4$. When there are more than one equilibria, only the best one for platforms is listed.

In Figure 3.2, we can see that ALLOW is the preferred option for platforms when β is low and $\frac{\gamma}{\beta}$ is high. Moreover, ADS OR FEE is better than ALLOW for relatively higher values of η and because $\frac{\eta}{\beta}$ is fixed in that plot, that means higher values of β make Ads or Fee better than Allow.

Similarly, we can see in Figure 3.3 that Allow is the preferred option when β is low and $\frac{\gamma}{\eta}$ is high. Moreover, when we compare Allow with Ads or Fee, higher η is better for Ads or Fee, which means lower $\frac{\gamma}{\eta}$ is better for it.

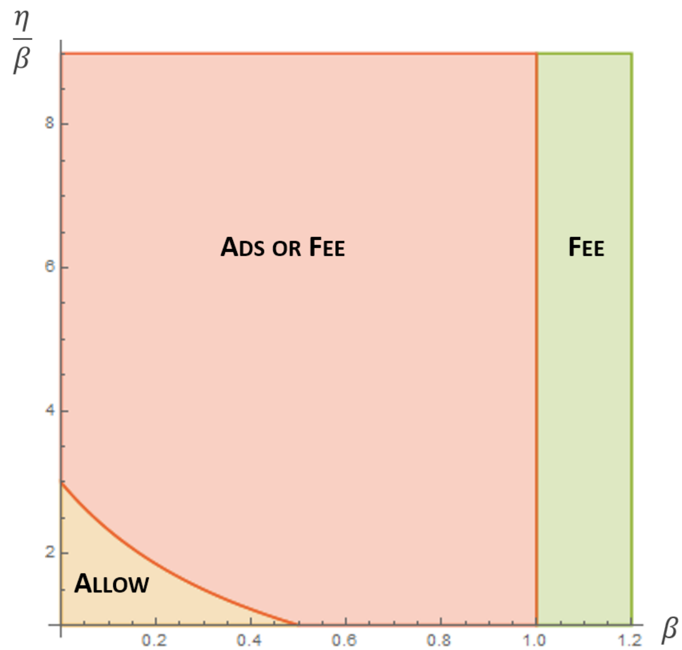


Figure 3.4: Equilibria regions of the model with the Ads or Fee strategy for $\lambda = \mu = \nu = 1$ and $\frac{\gamma}{\eta} = 4$. When there are more than one equilibria, only the best one for platforms is listed.

Finally, in Figure 3.4, we see that lower $\frac{\eta}{\beta}$ makes Allow better than Ads or Fee. This plot is also an example where the Ban region disappears. Even though Ban is an equilibrium when β and $\frac{\eta}{\beta}$ are low, it is never the best for the particular choice of parameter values (γ is too high compared to the other ad sensitivities).

3.5.2 Acceptable Ads and Whitelisting

In 2011, Eyeo, the company that developed Adblock Plus, the most popular adblock extension for browsers, started a program called the "Acceptable Ads initiative".[40] They set a list of criteria of what are considered acceptable ads based on placement, size, etc., and ads that complied with those criteria would be whitelisted by default in their ad blocker. Large companies, like Google, Microsoft, and Amazon were paying monthly fees to Eyeo to participate in this program and let their ads pass through the ad blocker.[41,42] These whitelisting services were also the main source of revenue for Adblock Plus.[43] In 2016, Eyeo extended this

[40] https://adblockplus.org/acceptable-ads (accessed May 2017)
[41] https://www.ft.com/content/80a8ce54-a61d-11e4-9bd3-00144feab7de (accessed May 2017)
[42] http://www.businessinsider.com/google-microsoft-amazon-taboola-pay-adblock-plus-to-stop-blocking-their-ads-2015-2 (accessed May 2017)
[43] https://adblockplus.org/about#monetization (accessed May 2017)

program by launching their own ad marketplace where they also started selling "acceptable" ads to publishers.[44,45]

The controversial nature of these moves by Adblock Plus created a lot of backlash. As an example, the CEO of Interactive Advertising Bureau (IAB) characterized Adblock Plus an "extortion-based business" and their actions "unethical and immoral".[46,47] The main argument is that first the company created an ad blocker that allowed millions of users to block ads on websites and now the same company charges money from advertisers and publishers to unblock their ads. This could be seen as a form of blackmail by those publishers and advertisers who now have to share part of their revenue with the ad-blocker company.

The key question we address in this context is whether this new option where platforms have to pay the ad-blocker company to whitelist their ads could ever be beneficial for them. In the presence of ad blockers, we may expect that sometimes platforms will have to follow this strategy, first if ad blockers hurt their ad revenue a lot, and second to get an advantage over competitors with blocked ads. However, more importantly, when they follow this strategy, how does their revenue change when compared to a world without ad blockers? In other words, can the option of paying the ad blocker to let their ads go through ever benefit them over the benchmark model without ad blockers where their ads were shown "for free"?

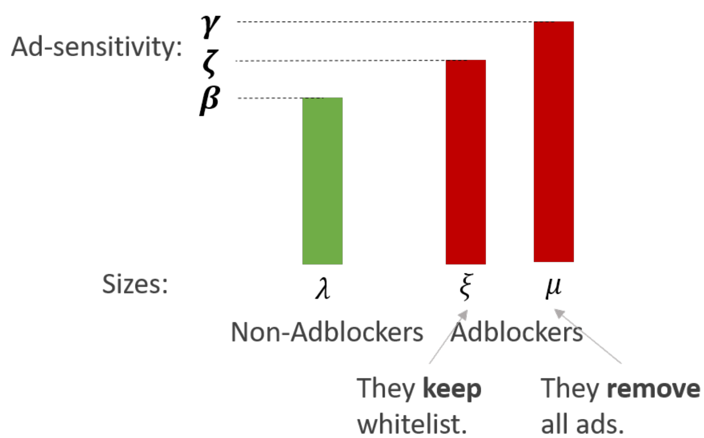


Figure 3.5: Illustration of the ad sensitivities of the three segments of users in the model with the WHITELIST strategy.

To answer this question we use another extension of the basic model. First, we add a new strategy to platforms' strategy space, called WHITELIST. When a platform chooses this WHITELIST option, they allow ad blockers and at the same time they pay a fee $f \geq 0$ [48] to

[44]https://www.theverge.com/2016/9/13/12890050/adblock-plus-now-sells-ads (accessed May 2017)

[45]https://acceptableads.com/en/solutions/ (accessed May 2017)

[46]https://techcrunch.com/2016/05/09/interactive-adverting-bureau-ceo-adblock-plus-is-an-extortion-based-business/ (accessed May 2017)

[47]http://www.businessinsider.com/interactive-advertising-bureau-comments-on-ad-block-plus-2016-1 (accessed May 2017)

[48]In the case of Adblock Plus and their program, this fee is zero for small entities without a lot of ad impressions and strictly positive for larger entities.

the ad-blocker company to whitelist their ads by default. As in the real world, some users who do not like this whitelist feature and do not want to watch even "acceptable" ads, have the option to disable the feature and remove all ads. We again consider three segments of users, but this time refining the segment of adblock users; our model supposes one segment of non-adblock users of mass λ with ad sensitivity β and two segments of adblock users. The first segment of adblock users are those who are fine with the whitelisting program and keep the default whitelist of "acceptable" ads. That segment is of mass ξ with ad sensitivity ζ. The second segment of adblock users are those who are against all ads, remove the default whitelist, and as a result block all ads. That segment is of mass μ with ad sensitivity γ. We do not assume any relationship between β, ζ, and γ, but as we show next, the most interesting results occur when $\beta \leq \gamma$ and $\zeta \leq \gamma$.

The next proposition answers the question above by showing that there is an equilibrium where both platforms choose the WHITELIST plan and that this equilibrium is sometimes better for platforms compared to the benchmark model with no ad blockers.

Proposition 3.4. *When we add the* WHITELIST *plan to platforms' strategy space, there is an equilibrium where both platforms choose the* WHITELIST *option. In this equilibrium, platforms are sometimes better off than they would be if ad blockers did not exist.*

To understand the intuition behind the proposition, consider the case where adblock users are more ad-sensitive than non-adblock users and that those adblock users who remove all ads are even more ad-sensitive than those who keep the ad blocker's whitelist, i.e. $\beta \leq \zeta \leq \gamma$ (Figure 3.5). The main idea is that the WHITELIST option can help platforms separate the two types of adblock users.

A platform with the BAN plan chooses some advertising intensity $a > 0$ and shows ads to all three segments of users (Figure 3.5). With the ALLOW plan, they choose some advertising intensity $a' > a$ and show ads only to the first segment of users. As we have seen in the basic model, sometimes ALLOW is better because of the high a' and sometimes BAN is better because the platforms get ad-revenue from more users. With the WHITELIST plan, the platform chooses some advertising intensity a'' with $a < a'' < a'$ and they show ads to the first two segments of users, the non-adblockers and the adblocker users who keep the whitelist. What happens is that this middle ground between BAN and ALLOW sometimes provides more revenue than the other two, i.e. showing ads to exactly two segments with medium ad intensity is better than showing ads to all three segments with low ad intensity or to just one segment with high ad intensity.

The conditions for which WHITELIST is the best option for platforms are the following. First, we want a sufficiently small fee f, otherwise WHITELIST will become a bad option because it is expensive. Second, we want the ad sensitivity of those who see ads in the WHITELIST plan to be low to make WHITELIST better than FEE, i.e. we want low β and ζ. Third, we want those who see ads in BAN but not in WHITELIST to have comparatively higher ad sensitivity, i.e. we want high $\frac{\gamma}{\beta}$ and $\frac{\gamma}{\zeta}$. Finally, to make WHITELIST better than ALLOW, we want the ad intensity of those who see ads in WHITELIST and not in ALLOW to have similar ad intensity to those who see ads in both, otherwise ALLOW that separates them would be better. Therefore, we also want low $\frac{\zeta}{\beta}$. In other words, for WHITELIST to be the best strategy, we want the subsegments of adblock users to be heterogeneous and well separated in their sensitivities, while the sensitivity of the lower segment among these is comparable to that of non-adblock users.

After we analyze all possible pair of strategies for platforms, we get the payoff matrix in Table 3.A.2. In this game, there are four symmetric equilibria, one for each strategy. In Figures 3.6, 3.7, and 3.8, we see the equilibria regions for different parameter values. As before, to make plots a bit simpler, when there are more than one equilibria we list only the best one for platforms. The regions where each equilibrium is unique are subregions of those labeled in the plots.

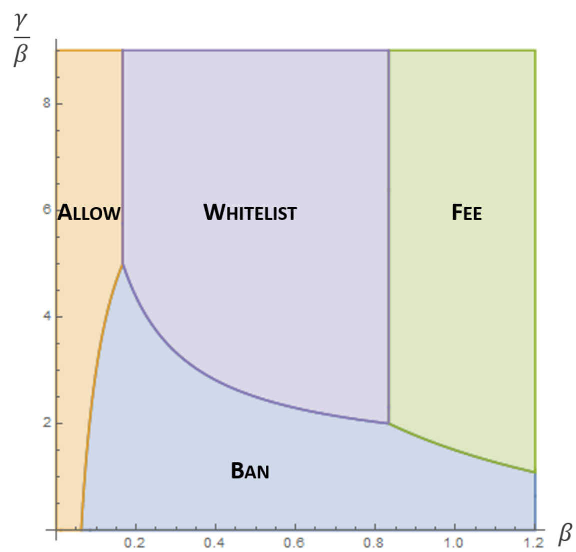


Figure 3.6: Equilibria regions of the model with the Whitelist strategy for $f = 0$, $\lambda = \mu = \xi = 1$, and $\zeta = \frac{1}{2}$. When there are more than one equilibria, only the best one for platforms is listed.

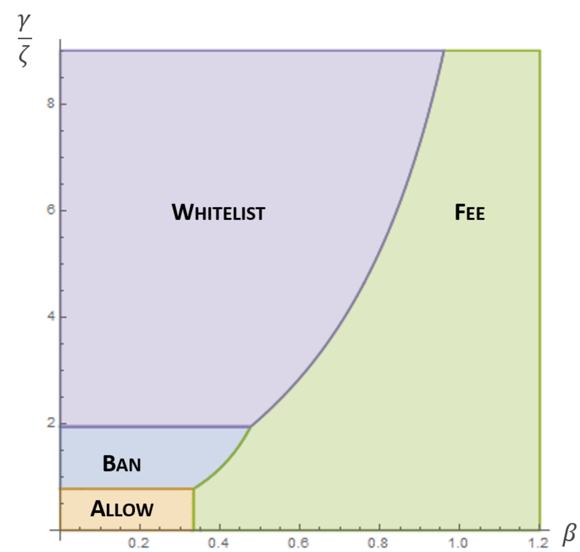


Figure 3.7: Equilibria regions of the model with the Whitelist strategy for $f = 0$, $\lambda = \mu = \xi = 1$, and $\frac{\gamma}{\beta} = \frac{7}{2}$. When there are more than one equilibria, only the best one for platforms is listed.

In Figure 3.6, we see that the Whitelist plan is preferred by platforms when $\frac{\gamma}{\beta}$ is high and β is low. When we compare the Whitelist plan with the Allow plan, Whitelist is better for low $\frac{\zeta}{\beta}$. In this particular plot ζ is fixed, so low $\frac{\zeta}{\beta}$ means high β. Therefore, there is a lower and an upper bound for β to make Whitelist the best option.

In Figure 3.7, we see that the Whitelist plan is preferred when $\frac{\gamma}{\zeta}$ is high and β is low. To understand why the Ban and the Allow regions are in the order they are, let's assume that β is fixed. Since in that plot $\frac{\gamma}{\beta}$ is also fixed, that means γ is fixed. Ban is better than Allow when $\frac{\zeta}{\beta}$ is low, which means when ζ is low, which means when $\frac{\gamma}{\zeta}$ is high.

In Figure 3.8, we see that the Whitelist plan is preferred for medium values of $\frac{\zeta}{\beta}$ and low β. We know that Whitelist is better than Allow for lower values of $\frac{\zeta}{\beta}$. For the comparison between Whitelist and Ban, let's again assume that β is fixed. Higher $\frac{\zeta}{\beta}$ means higher ζ and since $\frac{\gamma}{\zeta}$ for that plot is fixed, that means higher γ. But γ is the ad sensitivity of those who see ads in the Ban plan and do not see ads in the Whitelist plan. Therefore, higher γ makes Whitelist the better option.

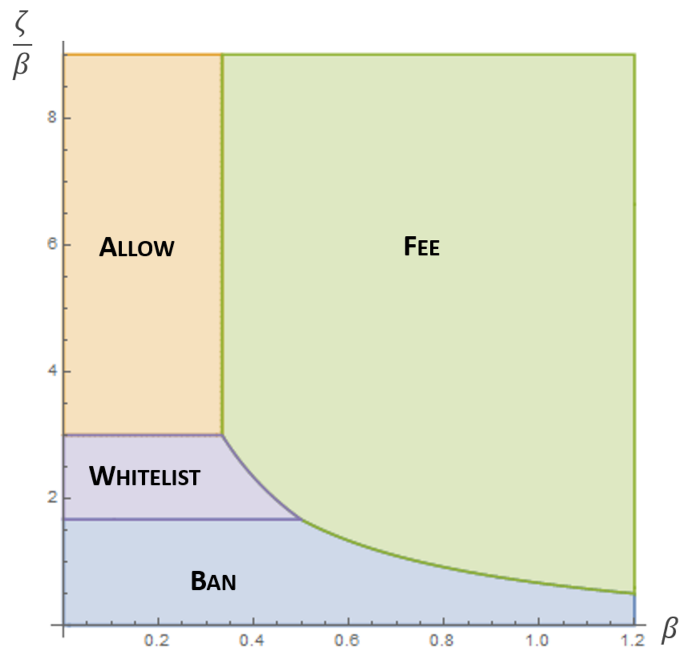


Figure 3.8: Equilibria regions of the model with the WHITELIST strategy for $f = 0$, $\lambda = \mu = \xi = 1$, and $\frac{\gamma}{\zeta} = 2$. When there are more than one equilibria, only the best one for platforms is listed.

In April of 2017, there were surprising reports that Google is planning to create their own built-in ad blocker for the Chrome browser.[49] This ad blocker will remove only "unacceptable" ads from web pages. In other words, Google wants to implement a similar program like the "Acceptable Ads initiative" by Adblock Plus. This is further evidence that Google realizes the benefits of an ad blocker with a whitelisting feature, even though Google itself depends heavily on advertising. Thus, instead of letting third parties implement such a feature and take part of their ad revenue, Google might prefer to do it on its own and thus exercise more control of the ad-blocker market.

3.6 Quality of Content and Content Creators

An important feature of many internet platforms that affects users decision of which platform to join is the quality of content. Some platforms generate their own content, while others depend on third parties to generate content for them.

Youtube is an example of a platform that does not generate its own content. Instead it depends on content creators to create and upload videos on the website that other users watch. For a very long time Youtube was dependent solely on advertising as its main source of revenue. However, recently it started offering subscription plans to its users (named Youtube Red) for an ad-free version of Youtube with some additional exclusive content. Any revenue

[49] https://www.wsj.com/articles/google-plans-ad-blocking-feature-in-popular-chrome-browser-1492643233 and https://www.theverge.com/2017/4/19/15365572/google-ad-blocking-feature-chrome-browser (accessed May 2017)

Youtube gets from advertising and from subscriptions is shared with the content creators. Youtube gets 45% of the revenue, while content creators get the remaining 55%.[50] We examine the question of how the quality of content is affected by the advent of ad blockers under a revenue share model, like the one Youtube implemented. To do that, we extend the basic model by adding content creators.

We assume that each platform has its own content creators. Content creators in platform $i \in \{1, 2\}$ have as decision variable the quality of content q_i and they incur some cost $c_i \cdot q_i^2$ to generate content of this quality.[51] User utility is the same as in the basic model with an additional quality-based utility term of $r \cdot q_i$ for platform i, i.e. higher quality of content in a platform means higher utility for users in that platform. Each platform has also a fixed fraction f_i that determines how they split the revenue with the content creators. The profit for content creators in platform i is $\pi_i = f_i \cdot (\text{total revenue}) - c_i \cdot q_i^2$, while the profit for platform i is $r_i = (1 - f_i) \cdot (\text{total revenue})$. Finally, the timeline of the game is the same as before with the addition that in the second step content creators also decide the value of q_i to maximize their profit, before users choose between the platforms.

We can show that all the results of the basic model are robust under this extension with content creators. More specifically, there is an equilibrium where platforms allow ad blockers and all platforms, content creators, and users are better off compared to a world without ad blockers. The fact that content creators can be better off when ad blockers are allowed causes an increase in the quality of content. This is the main result in the next proposition.

Proposition 3.5. *When ad blockers are allowed by platforms, the quality of content is higher for sufficiently high $\frac{\gamma}{\beta}$ and sufficiently low β. This is an additional benefit for users when ad blockers are allowed, whose welfare can increase even more than the increase with ad blockers in the basic model without content creators.*

Indeed, for the case of symmetric platforms, total user welfare is the same in this extension and in the basic model when platforms Ban ad blockers or when they use a Fee. However, when platforms Allow ad blockers, user welfare is higher in this extension than in the basic model. This is because when platforms Ban ad blockers or use a Fee, all the extra value that is generated by the quality of content goes to the platforms and the content creators in the form of increased ad intensity or price. However, when platforms Allow ad blockers, the higher quality of content allows this extra value to be shared with users who benefit even more with ad blockers.

3.7 Conclusion

3.7.1 Managerial Implications

Our analysis leads to several managerial implications. It can provide websites with some general guidelines regarding the plan they should choose based on how heterogeneous their visitors are in their ad sensitivity. In Figure 3.1, we exhibit a decision diagram summarizing our findings.

[50] https://support.google.com/youtube/answer/6204741 (accessed May 2017)

[51] The convex cost function is a natural choice for increasing quality which has diminishing returns to the

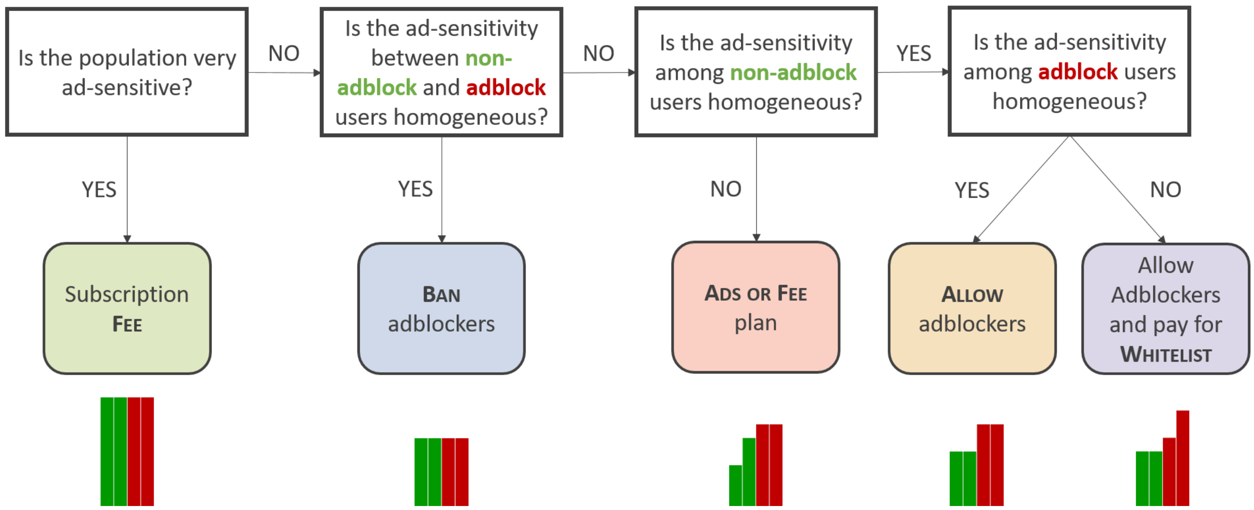


Figure 3.1: Flowchart for platforms to decide which strategy to follow based on the heterogeneity in the ad sensitivity of their user base.

The decision flowchart contains four questions in increasing order of refinement. If the users are generally very ad-sensitive, then the platform cannot expect to receive a lot of ad revenue from them, so it is better to choose a subscription based plan with a fee (FEE). Otherwise, the platform can benefit from serving ads to the users. If the ad sensitivities of non-adblock and adblock users are similar (or if non-adblock users are more ad-sensitive), then ad blockers cannot help the platform and it is better to ban them with an adblock wall (BAN). If, on the other hand, adblock users are more ad-sensitive than non-adblock users, then the platform would be better off with a plan that filters out the very ad-sensitive adblock users. The third question is about the ad sensitivity of non-adblock users. If it is very heterogeneous, then a plan that offers options to the users like the ADS OR FEE plan can help the platform filter out the very ad-sensitive non-adblock users. If non-adblock users have homogeneous ad sensitivity and adblock users are more ad-sensitive than non-adblock users, then allowing ad blockers can be beneficial. The heterogeneity of adblock users plays a role here. If adblock users are homogeneous then just allowing ad blockers can be enough (ALLOW), but if they are very heterogeneous then a whitelist option on top of allowing ad blockers can be the better plan (WHITELIST), since it filters out only the very ad-sensitive part of adblock users and keeps the rest, even if that means the platform has to pay a fee to the ad blocker for whitelisting its ads.

The way these findings can be used is the following. First, a platform can run a few tests with varying types of adblock walls or messages to the users (as many web platforms already do) in order to estimate the ad sensitivities of the various user segments. Then it can use the insights from our analysis to decide the ideal plans to offer.

effort by the creator.

3.7.2 Summary

Ad blockers initiated an existential crisis in the world of online platforms subsisting on advertising. While most speculations point to a grim outlook for advertisers and platforms as a result of ad blockers, our results offer an alternative view that might offer a glimmer of hope for the whole ecosystem, by arguing that ad blockers could actually be beneficial overall.

We suggest several ways in which ad blocking can be beneficial. First, they can make the market more efficient by filtering out ad-sensitive users for more intense or targeted ad serving on the rest. Second, ad-sensitive users can benefit because they can remove ads that annoy them from web sites. Third, ad blockers can also help regulate the ad industry through a whitelisting program of acceptable ads. Finally, a more efficient market can also result in an increase in content quality of web sites, which is an additional benefit for users.

A few years ago, when ad blockers started rising, publishers and advertisers were terrified of their implications for the future of the online ad industry. However, today we see that many of them choose a more friendly approach towards ad blockers. News like the recent plan of Google to create its own ad blocker in its Chrome browser,[52] shows that the industry has started realizing the potential benefits of ad blocking and decided to make it an ally instead of an enemy. As in many existential crises, the result could be rewarding.

3.A Appendix

3.A.1 Analyses and Proofs

Proof of Proposition 3.1. We consider first the scenario where ad blockers do not exist (benchmark model). This is when platforms have only two available strategies: the Ads strategy where everyone who access the websites has to see ads, and the Fee strategy where users have to pay a fee to access the site.

In this scenario, we start by considering four possible cases, one for each combination of plans chosen by the two platforms. For each one of these cases, we will find how the users react and then which of those cases end up in an equilibrium.

1. Both platforms use Ads*.* The indifferent user among those without ad blockers is the one at position x_N that is the solution to the equation $m + 1 - x_N - \beta a_1 = m + x_N - \beta a_2$, i.e. $x_N = \frac{1+\beta(a_2-a_1)}{2}$.

The indifferent user among those with ad blockers (when they are available) is the one at position x_A that is the solution to the equation $m + 1 - x_A - \gamma a_1 = m + x_A - \gamma a_2$, i.e. $x_A = \frac{1+\gamma(a_2-a_1)}{2}$.

Therefore, the expected market share of platform 1 is $z_1 = \lambda x_N + \mu x_A$, while the expected market share of platform 2 is $z_2 = \lambda(1 - x_N) + \mu(1 - x_A)$. The profit for platform 1 is then $z_1 a_1$ and the profit for platform 2 is $z_2 a_2$. Thus, to find the advertising intensities a_1, a_2 in the equilibrium, we need to solve the system $\frac{\partial(z_1 a_1)}{\partial a_1} = \frac{\partial(z_2 a_2)}{\partial a_2} = 0$. The solution is $a_1 = a_2 = \frac{\lambda+\mu}{\beta\lambda+\gamma\mu}$. From this, we get that the profit for both platforms is equal to $\frac{(\lambda+\mu)^2}{2(\beta\lambda+\gamma\mu)}$.

[52] https://www.engadget.com/2017/06/01/google-chrome-ad-blocking-baked-in-2018/ (accessed June 2017)

2. Both platforms have a subscription FEE. In this case, both types of users have similar payoff function, thus the indifferent user for both types is the one at position x that is the solution to the equation $m + 1 - x - p_1 = m + x - p_2$, i.e. $x = \frac{1+p_2-p_1}{2}$.

Therefore, the expected market share of platform 1 is $z_1 = (\lambda + \mu)x$, while the expected market share of platform 2 is $z_2 = (\lambda + \mu)(1 - x)$.

The profit for platform 1 is then z_1p_1 and the profit for platform 2 is z_2p_2. Thus, to find the prices p_1, p_2 in the equilibrium, we need to solve the system $\frac{\partial(z_1p_1)}{\partial p_1} = \frac{\partial(z_2p_2)}{\partial p_2} = 0$. The solution is $p_1 = p_2 = 1$. From this, we get that the profit for both platforms is equal to $\frac{\lambda+\mu}{2}$.

3. First platform uses ADS*, while the second uses a* FEE*.* The indifferent user among non-adblock users is the one at position x_N that is the solution to the equation $m+1-x_N-\beta a_1 = m + x_N - p_2$, i.e. $x_N = \frac{1+p_2-\beta a_1}{2}$.

The indifferent user among adblock users is the one at position x_A that is the solution to the equation $m + 1 - x_A - \gamma a_1 = m + x_A - p_2$, i.e. $x_A = \frac{1+p_2-\gamma a_1}{2}$.

The expected market share of platform 1 is $z_1 = \lambda x_N + \mu x_A$, while the expected market share of platform 2 is $z_2 = \lambda(1 - x_N) + \mu(1 - x_A)$. The profit for platform 1 is then z_1a_1 and the profit for platform 2 is z_2p_2. Thus, to find the advertising intensity a_1 and the price p_2 in the equilibrium, we need to solve the system $\frac{\partial(z_1a_1)}{\partial a_1} = \frac{\partial(z_2p_2)}{\partial p_2} = 0$. The solution is $a_1 = \frac{\lambda+\mu}{\beta\lambda+\gamma\mu}$ and $p_2 = 1$.[53] From this, we get that the profit for platform 1 is $\frac{(\lambda+\mu)^2}{2(\beta\lambda+\gamma\mu)}$, while the profit for platform 2 is $\frac{\lambda+\mu}{2}$.

4. First platform uses a FEE*, while the second uses* ADS*.* This case is similar to the previous case. The profit for platform 1 is $\frac{\lambda+\mu}{2}$, while the profit for platform 2 is $\frac{(\lambda+\mu)^2}{2(\beta\lambda+\gamma\mu)}$.

The following payoff matrix summarizes the four cases above.

		Platform 2: ADS	Platform 2: FEE
Platform 1	ADS	$\frac{(\lambda+\mu)^2}{2(\beta\lambda+\gamma\mu)}, \frac{(\lambda+\mu)^2}{2(\beta\lambda+\gamma\mu)}$	$\frac{(\lambda+\mu)^2}{2(\beta\lambda+\gamma\mu)}, \frac{\lambda+\mu}{2}$
	FEE	$\frac{\lambda+\mu}{2}, \frac{(\lambda+\mu)^2}{2(\beta\lambda+\gamma\mu)}$	$\frac{\lambda+\mu}{2}, \frac{\lambda+\mu}{2}$

We observe that

- (ADS, ADS) is an equilibrium iff $\lambda + \mu \geq \beta\lambda + \gamma\mu$.
- (FEE, FEE) is an equilibrium iff $\lambda + \mu \leq \beta\lambda + \gamma\mu$.
- (ADS, FEE) and (FEE, ADS) are equilibria iff $\lambda + \mu = \beta\lambda + \gamma\mu$.

Now we consider the scenario of the main model, where ad blockers are introduced and the second type of users can use them if they are allowed. The BAN strategy of the main model is similar to the ADS strategy of the benchmark, since every user has to see ads in both of them. Therefore, for the analysis of the main model we can use the four cases we considered in the benchmark and add to them five more cases for the strategy profiles that include the ALLOW strategy.

[53] Note that the indifferent user for those who do not use ad blockers is the one at position $x_N = \frac{\beta\lambda-\beta\mu+2\gamma\mu}{2(\beta\lambda+\gamma\mu)}$, which is always at most 1, but we also need this to be non-negative. Therefore, we need that $\beta\lambda+2\gamma\mu \geq \beta\mu$. Similarly, for the indifferent user for those with ad blockers, we need the inequality $2\beta\lambda + \gamma\mu \geq \gamma\lambda$. Due to symmetry, we get the same conditions from the fourth case of the proof as well.

5. Both platforms Allow *ad blockers.* The indifferent user among those who do not use ad blockers is the one at position x_N that is the solution to the equation $m+1-x_N-\beta a_1 = m + x_N - \beta a_2$, i.e. $x_N = \frac{1+\beta(a_2-a_1)}{2}$.

The indifferent user among those who use ad blockers is the one at position x_A that is the solution to the equation $m+1-x_A = m+x_A$, since now this type of users can use ad blockers to avoid ads. It is $x_A = \frac{1}{2}$.

The expected market share of platform 1 is $z_1 = \lambda x_N$, while the expected market share of platform 2 is $z_2 = \lambda(1-x_N)$. The profit for platform 1 is then $z_1 a_1$ and the profit for platform 2 is $z_2 a_2$. Thus, to find the advertising intensities a_1, a_2 in the equilibrium, we need to solve the system $\frac{\partial(z_1 a_1)}{\partial a_1} = \frac{\partial(z_2 a_2)}{\partial a_2} = 0$. The solution is $a_1 = a_2 = \frac{1}{\beta}$. From this, we get that the profit for both platforms is equal to $\frac{\lambda}{2\beta}$.

6. First platform Allow*s ad blockers, while the second uses* Ban*.* The indifferent user among those who do not use ad blockers is the one at position x_N that is the solution to the equation $m+1-x_N-\beta a_1 = m+x_N-\beta a_2$, i.e. $x_N = \frac{1+\beta(a_2-a_1)}{2}$.

The indifferent user among those with ad blockers is the one at position x_A that is the solution to the equation $m+1-x_A = m+x_A-\gamma a_2$. It is $x_A = \frac{1+\gamma a_2}{2}$.

The expected market share of platform 1 is $z_1 = \lambda x_N$, while the expected market share of platform 2 is $z_2 = \lambda(1-x_N)+\mu(1-x_A)$. The profit for platform 1 is then $z_1 a_1$ and the profit for platform 2 is $z_2 a_2$. Thus, to find the advertising intensities a_1, a_2 in the equilibrium, we need to solve the system $\frac{\partial(z_1 a_1)}{\partial a_1} = \frac{\partial(z_2 a_2)}{\partial a_2} = 0$. The solution is $a_1 = \frac{3\beta\lambda+\beta\mu+2\gamma\mu}{\beta(3\beta\lambda+4\gamma\mu)}$ and $a_2 = \frac{3\lambda+2\mu}{3\beta\lambda+4\gamma\mu}$.[54] From this, we get that the profit for platform 1 is equal to $\frac{\lambda(3\beta\lambda+\beta\mu+2\gamma\mu)^2}{2\beta(3\beta\lambda+4\gamma\mu)^2}$, while the profit for platform 2 is $\frac{(3\lambda+2\mu)^2(\beta\lambda+\gamma\mu)}{2(3\beta\lambda+4\gamma\mu)^2}$.

7. First platform Allow*s ad blockers, while the second uses* Fee*.* The indifferent user among non-adblockers is the one at position x_N that is the solution to the equation $m+1-x_N-\beta a_1 = m+x_N-p_2$, i.e. $x_N = \frac{1+p_2-\beta a_1}{2}$.

The indifferent user among those who use ad blockers is the one at position x_A that is the solution to the equation $m+1-x_A = m+x_A-p_2$. It is $x_A = \frac{1+p_2}{2}$.

The expected market share of platform 1 is $z_1 = \lambda x_N$, while the expected market share of platform 2 is $z_2 = \lambda(1-x_N)+\mu(1-x_A)$. The profit for platform 1 is $z_1 a_1$ and the profit for platform 2 is $z_2 p_2$. Thus, to find the advertising intensity a_1 and the price p_2 in the equilibrium, we need to solve the system $\frac{\partial(z_1 a_1)}{\partial a_1} = \frac{\partial(z_2 p_2)}{\partial p_2} = 0$. The solution is $a_1 = \frac{3(\lambda+\mu)}{\beta(3\lambda+4\mu)}$ and $p_2 = \frac{3\lambda+2\mu}{3\lambda+4\mu}$.[55] From this, we get that the profit for platform 1 is equal to $\frac{9\lambda(\lambda+\mu)^2}{2\beta(3\lambda+4\mu)^2}$, while the profit for platform 2 is $\frac{(\lambda+\mu)(3\lambda+2\mu)^2}{2(3\lambda+4\mu)^2}$.

8. First platform Ban*s ad blockers, while the second* Allow*s them.* This is similar to case 6. The profit for platform 1 is equal to $\frac{(3\lambda+2\mu)^2(\beta\lambda+\gamma\mu)}{2(3\beta\lambda+4\gamma\mu)^2}$, while the profit for platform 2 is

[54] Note that the indifferent user for those who do not use ad blockers is the one at position $x_N = \frac{3\beta\lambda+\beta\mu+2\gamma\mu}{6\beta\lambda+8\gamma\mu}$, which is always non-negative, but we also need this to be at most 1. Therefore, we need that $3\beta\lambda+6\gamma\mu \geq \beta\mu$. Similarly, for the indifferent user for those who use ad blockers, we need the inequality $3\beta\lambda+2\gamma\mu \geq 3\gamma\lambda$. We get the same conditions for the symmetric case 8.

[55] The indifferent user for those who do not use ad blockers is the one at position $x_N = \frac{3(\lambda+\mu)}{\beta(6\lambda+8\mu)}$, which is always non-negative and at most 1. Therefore, we do not need any extra inequality here. Similarly, for the indifferent user for those who use ad blockers, we do not need any extra condition either. The same is true for the symmetric case 9.

$\frac{\lambda(3\beta\lambda+\beta\mu+2\gamma\mu)^2}{2\beta(3\beta\lambda+4\gamma\mu)^2}$.

9. First platform uses a Fee*, while the second* Allow*s ad blockers.* This is similar to case 7. The profit for platform 1 is equal to $\frac{(\lambda+\mu)(3\lambda+2\mu)^2}{2(3\lambda+4\mu)^2}$, while the profit for platform 2 is $\frac{9\lambda(\lambda+\mu)^2}{2\beta(3\lambda+4\mu)^2}$.

Summarizing all of the above, we get the following payoff matrix for the two platforms.

	Ban	Allow	Fee
Ban	$\frac{(\lambda+\mu)^2}{2(\beta\lambda+\gamma\mu)}, \frac{(\lambda+\mu)^2}{2(\beta\lambda+\gamma\mu)}$	$\frac{(3\lambda+2\mu)^2(\beta\lambda+\gamma\mu)}{2(3\beta\lambda+4\gamma\mu)^2}, \frac{\lambda(3\beta\lambda+\beta\mu+2\gamma\mu)^2}{2\beta(3\beta\lambda+4\gamma\mu)^2}$	$\frac{(\lambda+\mu)^2}{2(\beta\lambda+\gamma\mu)}, \frac{\lambda+\mu}{2}$
Allow	$\frac{\lambda(3\beta\lambda+\beta\mu+2\gamma\mu)^2}{2\beta(3\beta\lambda+4\gamma\mu)^2}, \frac{(3\lambda+2\mu)^2(\beta\lambda+\gamma\mu)}{2(3\beta\lambda+4\gamma\mu)^2}$	$\frac{\lambda}{2\beta}, \frac{\lambda}{2\beta}$	$\frac{9\lambda(\lambda+\mu)^2}{2\beta(3\lambda+4\mu)^2}, \frac{(\lambda+\mu)(3\lambda+2\mu)^2}{2(3\lambda+4\mu)^2}$
Fee	$\frac{\lambda+\mu}{2}, \frac{(\lambda+\mu)^2}{2(\beta\lambda+\gamma\mu)}$	$\frac{(\lambda+\mu)(3\lambda+2\mu)^2}{2(3\lambda+4\mu)^2}, \frac{9\lambda(\lambda+\mu)^2}{2\beta(3\lambda+4\mu)^2}$	$\frac{\lambda+\mu}{2}, \frac{\lambda+\mu}{2}$

We observe that (Allow, Allow) is an equilibrium iff the following two conditions hold.

$$\frac{\lambda}{2\beta} \geq \frac{(\lambda+\mu)(3\lambda+2\mu)^2}{2(3\lambda+4\mu)^2} \qquad \text{and} \qquad \frac{\lambda}{2\beta} \geq \frac{(3\lambda+2\mu)^2(\beta\lambda+\gamma\mu)}{2(3\beta\lambda+4\gamma\mu)^2}$$

These two conditions are equivalent to

$$\beta \leq \frac{\lambda(3\lambda+2\mu)^2}{(\lambda+\mu)(3\lambda+2\mu)^2} \qquad \text{and} \qquad \frac{\lambda+\mu\cdot\frac{\gamma}{\beta}}{\left(3\lambda+4\mu\cdot\frac{\gamma}{\beta}\right)^2} \leq \frac{\lambda}{(3\lambda+2\mu)^2}.$$

The function $g(x) = \frac{\lambda+\mu\cdot x}{(3\lambda+4\mu\cdot x)^2}$ is decreasing, therefore (Allow, Allow) is an equilibrium for low enough β and high enough $\frac{\gamma}{\beta}$.

Moreover, if

$$\frac{\lambda}{2\beta} \geq \frac{(\lambda+\mu)^2}{2(\beta\lambda+\gamma\mu)} \qquad \text{and} \qquad \frac{\lambda}{2\beta} \geq \frac{\lambda+\mu}{2},$$

then the profits of the two platforms in the (Allow, Allow) equilibrium are larger than their profits in the (Ban, Ban) and (Fee, Fee) equilibria. These conditions are equivalent to $\frac{\gamma}{\beta} \geq 2+\frac{\mu}{\lambda}$ and $\beta \leq \frac{\lambda}{\lambda+\mu}$, so again when β is low enough and $\frac{\gamma}{\beta}$ is high enough. □

Proof of Proposition 3.2. When both platforms choose Allow, the user utility is

$$\begin{aligned} &\lambda\left(\int_0^{\frac{1}{2}}\left(m+1-x-\beta\cdot\frac{1}{\beta}\right)dx+\int_{\frac{1}{2}}^1\left(m+x-\beta\cdot\frac{1}{\beta}\right)dx\right) \\ &\quad+\mu\left(\int_0^{\frac{1}{2}}(m+1-x)\,dx+\int_{\frac{1}{2}}^1(m+x)\,dx\right) \\ =\;&\lambda\left(m-\frac{1}{4}\right)+\mu\left(m+\frac{3}{4}\right). \end{aligned}$$

When both platforms choose Ban, the user utility is

$$
\begin{aligned}
& \lambda\left(\int_0^{\frac{1}{2}}\left(m+1-x-\beta\cdot\frac{\lambda+\mu}{\beta\lambda+\gamma\mu}\right)dx+\int_{\frac{1}{2}}^1\left(m+x-\beta\cdot\frac{\lambda+\mu}{\beta\lambda+\gamma\mu}\right)dx\right) \\
& \quad +\mu\left(\int_0^{\frac{1}{2}}\left(m+1-x-\gamma\cdot\frac{\lambda+\mu}{\beta\lambda+\gamma\mu}\right)dx+\int_{\frac{1}{2}}^1\left(m+x-\gamma\cdot\frac{\lambda+\mu}{\beta\lambda+\gamma\mu}\right)dx\right) \\
= \; & \lambda\left(m+\frac{3}{4}-\frac{\beta(\lambda+\mu)}{\beta\lambda+\gamma\mu}\right)+\mu\left(m+\frac{3}{4}-\frac{\gamma(\lambda+\mu)}{\beta\lambda+\gamma\mu}\right) \\
= \; & (\lambda+\mu)\left(m-\frac{1}{4}\right).
\end{aligned}
$$

When both platforms choose Fee, the user utility is

$$
\begin{aligned}
& \lambda\left(\int_0^{\frac{1}{2}}(m+1-x-1)\,dx+\int_{\frac{1}{2}}^1(m+x-1)\,dx\right) \\
& \quad +\mu\left(\int_0^{\frac{1}{2}}(m+1-x-1)\,dx+\int_{\frac{1}{2}}^1(m+x-1)\,dx\right) \\
= \; & \lambda\left(m-\frac{1}{4}\right)+\mu\left(m-\frac{1}{4}\right) \\
= \; & (\lambda+\mu)\left(m-\frac{1}{4}\right).
\end{aligned}
$$

Note that the total user utility is the same when both platforms choose Ban or both platforms choose Fee, and it is higher when both platforms choose Allow.

Moreover, the utility of users without ad blockers is the same when both platforms choose Allow or both platforms choose Fee, while the utility of users with ad blockers is always higher when both platforms choose Allow. This implies that in the regions where platforms were using Fee in the benchmark but Allow in the main model, no user is worse off with Allow. □

Proof of Proposition 3.3. This proof requires to consider 16 cases, one for each strategy profile, similar to the cases of the proof of Proposition 3.1 but now with three segments of users. Since the proof is very repetitive, we illustrate here only one of the cases and then we provide the payoff matrix we obtain after the full analysis.

Both platforms use the Ads or Fee *plan.* This case has several subcases based on what option (between ads or fee) each segment of users picks. Since $\beta \leq \eta \leq \gamma$, there are four subcases for platform 1:

- $p_1 < \beta a_1 \leq \eta a_1 \leq \gamma a_1$
- $\beta a_1 \leq p_1 \leq \eta a_1 \leq \gamma a_1$
- $\beta a_1 \leq \eta a_1 < p_1 \leq \gamma a_1$
- $\beta a_1 \leq \eta a_1 \leq \gamma a_1 < p_1$

In the first subcase, every user in platform 1 prefers to pay the fee, therefore this subcase is as if platform 1 had chosen the FEE plan, and we can ignore it here.[56] Similarly, the fourth subcase is as if platform 1 had chosen the BAN plan, and we can ignore it too. This leaves us with two subcases for platform 1 and two similar subcases for platform 2. Therefore, there are four possible strategy profiles we need to analyze. For brevity, we show only the two symmetric ones here.

1. $\beta a_1 \leq p_1 \leq \eta a_1 \leq \gamma a_1$ and $\beta a_2 \leq p_2 \leq \eta a_2 \leq \gamma a_2$. The indifferent user among the first segment of non-adblockers is the one at position x_N that is the solution to the equation $m + 1 - x_N - \beta a_1 = m + x_N - \beta a_2$, i.e. $x_N = \frac{1+\beta a_2 - \beta a_1}{2}$.

The indifferent user among the second segment of non-adblockers is the one at position $x_{N,2}$ that is the solution to the equation $m + 1 - x_{N,2} - p_1 = m + x_{N,2} - p_2$, since those non-adblock users pay the fee. It is $x_{N,2} = \frac{1+p_2-p_1}{2}$.

The indifferent user among adblock users is the one at position x_A that is the solution to the equation $m + 1 - x_A - p_1 = m + x_A - p_2$. It is $x_A = \frac{1+p_2-p_1}{2}$.

The expected mass of users in platform 1 who see ads is $z_{1,a} = \lambda x_N$, while the expected mass of users in platform 1 who pay the fee is $z_{1,p} = \nu x_{N,2} + \mu x_A$. Similarly, the expected mass of users in platform 2 who see ads is $z_{2,a} = \lambda(1 - x_N)$, while the expected mass of users in platform 2 who pay the fee is $z_{2,p} = \nu(1 - x_{N,2}) + \mu(1 - x_A)$.

The profit for platform 1 is $z_{1,a}a_1 + z_{1,p}p_1$ and the profit for platform 2 is $z_{2,a}a_2 + z_{2,p}p_2$. Thus, to find the advertising intensities a_1, a_2 and the prices p_1, p_2 in the equilibrium, we need to solve the system $\frac{\partial(z_{1,a}a_1+z_{1,p}p_1)}{\partial a_1} = \frac{\partial(z_{1,a}a_1+z_{1,p}p_1)}{\partial p_1} = \frac{\partial(z_{2,a}a_2+z_{2,p}p_2)}{\partial a_2} = \frac{\partial(z_{2,a}a_2+z_{2,p}p_2)}{\partial p_2} = 0$. The solution is $a_1 = a_2 = \frac{1}{\beta}$ and $p_1 = p_2 = 1$. From this, we get that the profit for both platforms is equal to $\frac{\lambda+\beta(\mu+\nu)}{2\beta}$.

Note that the solution we found satisfy the inequalities of this subcase. Therefore, this subcase gives us a possible equilibrium.

2. $\beta a_1 \leq \eta a_1 < p_1 \leq \gamma a_1$ and $\beta a_2 \leq \eta a_2 < p_2 \leq \gamma a_2$. The indifferent user among the first segment of non-adblock users is the one at position x_N that is the solution to the equation $m + 1 - x_N - \beta a_1 = m + x_N - \beta a_2$, i.e. $x_N = \frac{1+\beta a_2-\beta a_1}{2}$.

The indifferent user among the second segment of non-adblock users is now the one at position $x_{N,2}$ that is the solution to the equation $m + 1 - x_{N,2} - \eta a_1 = m + x_{N,2} - \eta a_2$, since in these subcase those non-adblock users see ads. It is $x_{N,2} = \frac{1+\eta a_2-\eta a_1}{2}$.

The indifferent user among adblock users is the one at position x_A that is the solution to the equation $m + 1 - x_A - p_1 = m + x_A - p_2$, i.e. $x_A = \frac{1+p_2-p_1}{2}$.

The expected mass of users in platform 1 who see ads is $z_{1,a} = \lambda x_N + \nu x_{N,2}$, while the expected mass of users in platform 1 who pay the fee is $z_{1,p} = \mu x_A$. Similarly, the expected mass of users in platform 2 who see ads is $z_{2,a} = \lambda(1 - x_N) + \nu(1 - x_{N,2})$, while the expected mass of users in platform 2 who pay the fee is $z_{2,p} = \mu(1 - x_A)$.

The profit for platform 1 is $z_{1,a}a_1 + z_{1,p}p_1$ and the profit for platform 2 is $z_{2,a}a_2 + z_{2,p}p_2$. Thus, to find the advertising intensities a_1, a_2 and the prices p_1, p_2 in the equilibrium, we

[56] Note that the ADS OR FEE plan is more general than the FEE and the BAN plan, as the platform can always make ad intensity or price equal to $+\infty$, forcing all of its users to either pay the fee or see ads respectively. To avoid confusion throughout the paper, when we say that a platform uses the ADS OR FEE plan, we actually mean that some of its users see ads and some of its users decide to pay the fee. When we say that a platform uses the FEE plan, we actually mean that it either uses the ADS OR FEE plan with ad intensity equal to $+\infty$ or it uses the actual FEE plan with no option for ads. Similarly for the BAN strategy.

need to solve the system $\frac{\partial(z_{1,a}a_1+z_{1,p}p_1)}{\partial a_1} = \frac{\partial(z_{1,a}a_1+z_{1,p}p_1)}{\partial p_1} = \frac{\partial(z_{2,a}a_2+z_{2,p}p_2)}{\partial a_2} = \frac{\partial(z_{2,a}a_2+z_{2,p}p_2)}{\partial p_2} = 0$. The solution is $a_1 = a_2 = \frac{\lambda+\nu}{\beta\lambda+\eta\nu}$ and $p_1 = p_2 = 1$. From this, we get that the profit for both platforms is equal to $\frac{\lambda^2+\beta\lambda\mu+2\lambda\nu+\eta\mu\nu+\nu^2}{2(\beta\lambda+\eta\nu)}$.

Note however, that the solution we found this time does not satisfy the inequalities of this subcase, because $\eta a_1 = \frac{\eta\lambda+\eta\nu}{\beta\lambda+\eta\nu} \geq \frac{\beta\lambda+\eta\nu}{\beta\lambda+\eta\nu} = 1 = p_1$. Therefore, this subcase does not give an equilibrium.

After we analyze all possible cases and subcases, we obtain the payoff matrix in Table 3.A.1. From that matrix, we can also obtain the conditions under which both platforms choose Allow as their equilibrium strategy.

An example where (Allow, Allow) is the unique equilibrium is for $\lambda = \mu = \nu = 1$, $\beta = \frac{1}{4}$, $\eta = \frac{3}{10}$, and $\gamma = 2$. In that case, the payoff matrix becomes

	Ban	Allow	Fee	Ads or Fee
Ban	$1.8, 1.8$	$0.88, 1.4$	$1.8, 1.5$	$1.8, 2.2$
Allow	$1.4, 0.88$	$3.6, 3.6$	$2.9, 0.96$	$3.1, 2.2$
Fee	$1.5, 1.8$	$0.96, 2.9$	$1.5, 1.5$	$1.5, 3$
Ads or Fee	$2.2, 1.8$	$2.2, 3.1$	$3, 1.5$	$3, 3$

□

Proof of Proposition 3.4. As in the proof of Proposition 3.3, this proof requires to consider all the 16 possible strategy combinations for the platforms. For brevity we show only one in detail here and then we provide the payoff matrix with the result of all cases.

Both platforms use the Whitelist *plan.* The indifferent user among the non-adblock users is the one at position x_N that is the solution to the equation $m + 1 - x_N - \beta a_1 = m + x_N - \beta a_2$, i.e. $x_N = \frac{1+\beta a_2-\beta a_1}{2}$.

The indifferent user among the first segment of adblock users, who keep the whitelist, is the one at position $x_{A,2}$ that is the solution to the equation $m+1-x_{A,2}-\zeta a_1 = m+x_{A,2}-\zeta a_2$, since those adblock users see ads. It is $x_{A,2} = \frac{1+\zeta a_2-\zeta a_1}{2}$.

The indifferent user among the second segment of adblock users, who remove all ads, is the one at position x_A that is the solution to the equation $m+1-x_A = m+x_A$, i.e. $x_A = \frac{1}{2}$.

The expected mass of users in platform 1 who see ads is $z_{1,a} = \lambda x_N + \xi x_{A,2}$. Similarly, the expected mass of users in platform 2 who see ads is $z_{2,a} = \lambda(1 - x_N) + \xi(1 - x_{A,2})$.

The profit for platform 1 is $z_{1,a}a_1 - f$ and the profit for platform 2 is $z_{2,a}a_2 - f$, since they also have to pay the fee f to the ad-blocker company. Thus, to find the advertising intensities a_1, a_2 in the equilibrium, we need to solve the system $\frac{\partial(z_{1,a}a_1-f)}{\partial a_1} = \frac{\partial(z_{2,a}a_2-f)}{\partial a_2} = 0$. The solution is $a_1 = a_2 = \frac{\lambda+\xi}{\beta\lambda+\zeta\xi}$. From this, we get that the profit for both platforms is equal to $\frac{(\lambda+\xi)^2}{2(\beta\lambda+\zeta\xi)} - f$.

In Table 3.A.2, we can see the payoff matrix of the full game. An example where (Whitelist, Whitelist) is the unique equilibrium of the game is $\lambda = \mu = \nu = 1$, $\beta = \zeta = \frac{1}{10}$, and $\gamma = \frac{6}{5}$. In that case, the payoff matrix becomes

	BAN	ALLOW	FEE	WHITELIST
BAN	$3.21, 3.21$	$1.13, 1.59$	$3.21, 1.5$	$1.54, 3.3$
ALLOW	$1.59, 1.13$	$5, 5$	$3.35, 0.61$	$3.67, 5.1$
FEE	$1.5, 3.21$	$0.61, 3.35$	$1.5, 1.5$	$0.96, 8.1$
WHITELIST	$3.3, 1.54$	$5.1, 3.67$	$8.1, 0.96$	$10, 10$

□

Proof of Proposition 3.5. This proof is similar to the proof of Proposition 3.1 with the additional step in every case of the decision about qualities by content creators. To avoid repetition, we analyze one case here and then we provide the payoff matrix of the full game.

Both platforms ALLOW *ad blockers.* The indifferent user among the non-adblock users is the one at position x_N that is the solution to the equation $m + 1 - x_N - \beta a_1 + rq_1 = m + x_N - \beta a_2 + rq_2$, i.e. $x_N = \frac{1+\beta(a_2-a_1)-r(q_2-q_1)}{2}$.

The indifferent user among those who use ad blockers is the one at position x_A that is the solution to the equation $m + 1 - x_A + rq_1 = m + x_A + rq_2$, since now this type of users can use ad blockers to avoid ads. It is $x_A = \frac{1-r(q_2-q_1)}{2}$.

The expected mass of users who see ads in platform 1 is $z_1 = \lambda x_N$, while the expected mass of users who see ads in platform 2 is $z_2 = \lambda(1 - x_N)$.

The profit for the content creators of platform 1 is $f_1 z_1 a_1 - c_1 q_1^2$, while the profit for the content creators of platform 2 is $f_2 z_2 a_2 - c_2 q_2^2$.[57] To find the qualities q_1, q_2, we need to solve the system $\frac{\partial(f_1 z_1 a_1 - c_1 q_1^2)}{\partial q_1} = \frac{\partial(f_2 z_2 a_2 - c_2 q_2^2)}{\partial q_2} = 0$. This gives the solution $q_1 = \frac{a_1 f_1 \lambda r}{4c_1}$ and $q_2 = \frac{a_2 f_2 \lambda r}{4c_2}$ as a function of the ad intensities a_1 and a_2.

The profit for platform 1 is $(1 - f_1)z_1 a_1$ and the profit for platform 2 is $(1 - f_2)z_2 a_2$. Thus, to find the ad intensities a_1, a_2 in the equilibrium, we need to solve the system $\frac{\partial((1-f_1)z_1 a_1)}{\partial a_1} = \frac{\partial((1-f_2)z_2 a_2)}{\partial a_2} = 0$. The solution is $a_1 = \frac{4c_1}{4\beta c_1 - f_1 \lambda r^2}$ and $a_2 = \frac{4c_2}{4\beta c_2 - f_2\lambda r^2}$. From this, we get that the profit for platform 1 is equal to $\frac{2c_1(1-f_1)\lambda}{4\beta c_1 - f_1\lambda r^2}$ and the profit for platform 2 is $\frac{2c_2(1-f_2)\lambda}{4\beta c_2 - f_2 \lambda r^2}$. Moreover, the qualities are $q_1 = \frac{f_1\lambda r}{4\beta c_1 - f_1\lambda r^2}$ and $q_2 = \frac{f_2 \lambda r}{4\beta c_2 - f_2\lambda r^2}$.

The user welfare of non-adblock users is

$$\begin{aligned}
&\lambda\left(\int_0^{\frac{1}{2}}\left(m + 1 - x - \beta\cdot\frac{4c_1}{4\beta c_1 - f_1\lambda r^2} + r\cdot\frac{f_1\lambda r}{4\beta c_1 - f_1\lambda r^2}\right)dx\right. \\
&\qquad \left. + \int_{\frac{1}{2}}^1\left(m + x - \beta\cdot\frac{4c_2}{4\beta c_2 - f_2\lambda r^2} + r\cdot\frac{f_2\lambda r}{4\beta c_2 - f_2\lambda r^2}\right)dx\right) \\
= \ &\lambda\left(m - \frac{1}{4}\right).
\end{aligned}$$

The user welfare of adblock users is

$$\begin{aligned}
&\mu\left(\int_0^{x_A}\left(m + 1 - x + r\cdot\frac{f_1\lambda r}{4\beta c_1 - f_1\lambda r^2}\right)dx\right. \\
&\qquad \left. + \int_{1-x_A}^1\left(m + x + r\cdot\frac{f_2\lambda r}{4\beta c_2 - f_2\lambda r^2}\right)dx\right),
\end{aligned}$$

[57] For simplicity in this analysis and to avoid corner solutions we assume that the cost parameters c_1 and c_2 are sufficiently large.

where $x_A = \frac{1}{2} + \frac{2\beta c_1}{4\beta c_1 - f_1 \lambda r^2} - \frac{2\beta c_2}{4\beta c_2 - f_2 \lambda r^2}$. For symmetric platforms, i.e. when $f_1 = f_2 = \hat{f}$ and $c_1 = c_2 = \hat{c}$, the user welfare of adblock users is

$$\mu \left(m + \frac{3}{4} + \frac{\hat{f} \lambda r^2}{4\beta \hat{c} - \hat{f} \lambda r^2} \right) \geq \mu \left(m + \frac{3}{4} \right),$$

larger than it was in the main model.

Table 3.A.3 contains the payoff matrix of the game. The quality in the (Ban, Ban) equilibrium for platform 1 is $q_{1,\text{Ban}} = \frac{f_1 (\lambda + \mu)^2 r}{4\beta c_1 \lambda + 4\gamma c_1 \mu - f_1 (\lambda + \mu)^2 r^2}$. This is less than or equal to $q_{1,\text{Allow}} = \frac{f_1 \lambda r}{4\beta c_1 - f_1 \lambda r^2}$ when $\frac{\gamma}{\beta} \geq 2 + \frac{\mu}{\lambda}$, i.e. for sufficiently high $\frac{\gamma}{\beta}$. Similarly, $q_{2,\text{Ban}} \leq q_{2,\text{Allow}}$ when $\frac{\gamma}{\beta} \geq 2 + \frac{\mu}{\lambda}$.

The quality in the (Fee, Fee) equilibrium for platform 1 is $q_{1,\text{Fee}} = \frac{f_1 (\lambda + \mu) r}{4c_1 - f_1 (\lambda + \mu) r^2}$. This is less than or equal to $q_{1,\text{Allow}} = \frac{f_1 \lambda r}{4\beta c_1 - f_1 \lambda r^2}$ when $\beta \leq \frac{\lambda}{\lambda + \mu}$, i.e. for sufficiently low β. Similarly, $q_{2,\text{Fee}} \leq q_{2,\text{Allow}}$ when $\beta \leq \frac{\lambda}{\lambda + \mu}$.

For the case of symmetric platforms, the total user welfare in the (Ban, Ban) and in the (Fee, Fee) equilibria is $(\lambda + \mu) \left(m - \frac{1}{4} \right)$, i.e. the same as in the main model. However, in the (Allow, Allow) equilibrium, the total user welfare is $\lambda \left(m - \frac{1}{4} \right) + \mu \left(m + \frac{3}{4} + \frac{\hat{f} \lambda r^2}{4\beta \hat{c} - \hat{f} \lambda r^2} \right) \geq \lambda \left(m - \frac{1}{4} \right) + \mu \left(m + \frac{3}{4} \right)$, larger than in the main model. $\square$

3.A.2 Payoff Matrices

	Ban	Allow	Fee	Ads or Fee
Ban	$\frac{(\lambda+\mu+\nu)^2}{2(\beta\lambda+\gamma\mu+\eta\nu)}, \frac{(\lambda+\mu+\nu)^2}{2(\beta\lambda+\gamma\mu+\eta\nu)}$	$\frac{(3\lambda+2\mu+3\nu)^2(\beta\lambda+\gamma\mu+\eta\nu)}{2(3\beta\lambda+4\gamma\mu+3\eta\nu)^2}, C$	$\frac{(\lambda+\mu+\nu)^2}{2(\beta\lambda+\gamma\mu+\eta\nu)}, \frac{1}{2}(\lambda+\mu+\nu)$	$\frac{(\lambda+\mu+\nu)^2}{2(\beta\lambda+\gamma\mu+\eta\nu)}, A$
Allow	$C, \frac{(3\lambda+2\mu+3\nu)^2(\beta\lambda+\gamma\mu+\eta\nu)}{2(3\beta\lambda+4\gamma\mu+3\eta\nu)^2}$	$\frac{(\lambda+\nu)^2}{2(\beta\lambda+\eta\nu)}, \frac{(\lambda+\nu)^2}{2(\beta\lambda+\eta\nu)}$	$\frac{9(\lambda+\nu)^2(\lambda+\mu+\nu)^2}{2(3\lambda+4\mu+3\nu)^2(\beta\lambda+\eta\nu)}, \frac{(\lambda+\mu+\nu)(3\lambda+2\mu+3\nu)^2}{2(3\lambda+4\mu+3\nu)^2}$	$\frac{9(\lambda+\nu)^2(\mu+\nu)^2(\beta\lambda+\eta\nu)}{2(3\beta\lambda(\mu+\nu)+\eta\nu(4\mu+3\nu))^2}, B$
Fee	$\frac{1}{2}(\lambda+\mu+\nu), \frac{(\lambda+\mu+\nu)^2}{2(\beta\lambda+\gamma\mu+\eta\nu)}$	$\frac{(\lambda+\mu+\nu)(3\lambda+2\mu+3\nu)^2}{2(3\lambda+4\mu+3\nu)^2}, \frac{9(\lambda+\nu)^2(\lambda+\mu+\nu)^2}{2(3\lambda+4\mu+3\nu)^2(\beta\lambda+\eta\nu)}$	$\frac{1}{2}(\lambda+\mu+\nu), \frac{1}{2}(\lambda+\mu+\nu)$	$\frac{1}{2}(\lambda+\mu+\nu), \frac{\lambda+\beta(\mu+\nu)}{2\beta}$
Ads or Fee	$A, \frac{(\lambda+\mu+\nu)^2}{2(\beta\lambda+\gamma\mu+\eta\nu)}$	$B, \frac{9(\lambda+\nu)^2(\mu+\nu)^2(\beta\lambda+\eta\nu)}{2(3\beta\lambda(\mu+\nu)+\eta\nu(4\mu+3\nu))^2}$	$\frac{\lambda+\beta(\mu+\nu)}{2\beta}, \frac{1}{2}(\lambda+\mu+\nu)$	$\frac{\lambda+\beta(\mu+\nu)}{2\beta}, \frac{\lambda+\beta(\mu+\nu)}{2\beta}$

Table 3.A.1: Payoff matrix in the model with the Ads or Fee plan.

$$A = \frac{\lambda^2(\mu+\nu)^2\beta^3+\lambda(\mu+\nu)\Big(4\lambda^2+2((\gamma+2)\mu+(\eta+2)\nu)\lambda+(\mu+\nu)(4\gamma\mu+\mu+4\eta\nu+\nu)\Big)\beta^2+(\gamma\mu+\eta\nu)\Big(((\gamma+4)\mu+(\eta+4)\nu)\lambda^2+2(\mu+\nu)(2\gamma\mu+\mu+2\eta\nu+\nu)\lambda+4(\mu+\nu)^2(\gamma\mu+\eta\nu)\Big)\beta+\lambda(\mu+\nu)(\gamma\mu+\eta\nu)^2}{8\beta(\mu+\nu)(\beta\lambda+\gamma\mu+\eta\nu)^2}$$

$$B = \frac{9\beta^3\lambda^2(\mu+\nu)^3+3\beta^2\lambda\Big(12\lambda^2+6(\eta+2)\nu\lambda+\nu(8\eta\mu+12\eta\nu+3\nu)\Big)(\mu+\nu)^2+\beta\eta\nu\Big((48\mu+9(\eta+4)\nu)\lambda^2+6\nu(4(\eta+1)\mu+6\eta\nu+3\nu)\lambda+4\eta\nu(2\mu+3\nu)^2\Big)(\mu+\nu)+\eta^2\lambda\nu^2(4\mu+3\nu)^2}{8\beta(3\beta\lambda(\mu+\nu)+\eta\nu(4\mu+3\nu))^2}$$

$$C = \frac{(2\gamma\mu(\lambda+\nu)+\beta\lambda(3\lambda+\mu+3\nu)+\eta\nu(3\lambda+\mu+3\nu))^2}{2(\beta\lambda+\eta\nu)(3\beta\lambda+4\gamma\mu+3\eta\nu)^2}$$

	Ban	Allow	Fee	Whitelist
Ban	$\frac{(\lambda+\mu+\xi)^2}{2(\beta\lambda+\gamma\mu+\zeta\xi)}, \frac{(\lambda+\mu+\xi)^2}{2(\beta\lambda+\gamma\mu+\zeta\xi)}$	$\frac{(\beta\lambda+\gamma\mu+\zeta\xi)(3\lambda+2(\mu+\xi))^2}{2(3\beta\lambda+4\gamma\mu+4\zeta\xi)^2}, E$	$\frac{(\lambda+\mu+\xi)^2}{2(\beta\lambda+\gamma\mu+\zeta\xi)}, \frac{1}{2}(\lambda+\mu+\xi)$	$\frac{(3\lambda+2\mu+3\xi)^2(\beta\lambda+\gamma\mu+\zeta\xi)}{2(3\beta\lambda+4\gamma\mu+3\zeta\xi)^2}, D$
Allow	$E, \frac{(\beta\lambda+\gamma\mu+\zeta\xi)(3\lambda+2(\mu+\xi))^2}{2(3\beta\lambda+4\gamma\mu+4\zeta\xi)^2}$	$\frac{\lambda}{2\beta}, \frac{\lambda}{2\beta}$	$\frac{9\lambda(\lambda+\mu+\xi)^2}{2\beta(3\lambda+4(\mu+\xi))^2}, \frac{(\lambda+\mu+\xi)(3\lambda+2(\mu+\xi))^2}{2(3\lambda+4(\mu+\xi))^2}$	$\frac{\lambda(2\zeta\xi+\beta(3\lambda+\xi))^2}{2\beta(3\beta\lambda+4\zeta\xi)^2}, \frac{(3\lambda+2\xi)^2(\beta\lambda+\zeta\xi)}{2(3\beta\lambda+4\zeta\xi)^2} - f$
Fee	$\frac{1}{2}(\lambda+\mu+\xi), \frac{(\lambda+\mu+\xi)^2}{2(\beta\lambda+\gamma\mu+\zeta\xi)}$	$\frac{(\lambda+\mu+\xi)(3\lambda+2(\mu+\xi))^2}{2(3\lambda+4(\mu+\xi))^2}, \frac{9\lambda(\lambda+\mu+\xi)^2}{2\beta(3\lambda+4(\mu+\xi))^2}$	$\frac{1}{2}(\lambda+\mu+\xi), \frac{1}{2}(\lambda+\mu+\xi)$	$\frac{(\lambda+\mu+\xi)(3\lambda+2\mu+3\xi)^2}{2(3\lambda+4\mu+3\xi)^2}, F$
Whitelist	$D, \frac{(3\lambda+2\mu+3\xi)^2(\beta\lambda+\gamma\mu+\zeta\xi)}{2(3\beta\lambda+4\gamma\mu+3\zeta\xi)^2}$	$\frac{(3\lambda+2\xi)^2(\beta\lambda+\zeta\xi)}{2(3\beta\lambda+4\zeta\xi)^2} - f, \frac{\lambda(2\zeta\xi+\beta(3\lambda+\xi))^2}{2\beta(3\beta\lambda+4\zeta\xi)^2}$	$F, \frac{(\lambda+\mu+\xi)(3\lambda+2\mu+3\xi)^2}{2(3\lambda+4\mu+3\xi)^2}$	$\frac{(\lambda+\xi)^2}{2(\beta\lambda+\zeta\xi)} - f, \frac{(\lambda+\xi)^2}{2(\beta\lambda+\zeta\xi)} - f$

Table 3.A.2: Payoff matrix in the model with the Whitelist plan.

$$D = \frac{(2\gamma\mu(\lambda+\xi)+\beta\lambda(3\lambda+\mu+3\xi)+\zeta\xi(3\lambda+\mu+3\xi))^2}{2(\beta\lambda+\zeta\xi)(3\beta\lambda+4\gamma\mu+3\zeta\xi)^2} - f$$

$$E = \frac{\lambda(\beta(3\lambda+\mu+\xi)+2(\gamma\mu+\zeta\xi))^2}{2\beta(3\beta\lambda+4\gamma\mu+4\zeta\xi)^2}$$

$$F = \frac{9(\lambda+\xi)^2(\lambda+\mu+\xi)^2}{2(3\lambda+4\mu+3\xi)^2(\beta\lambda+\zeta\xi)} - f$$

	Ban	Allow	Fee
Ban	$-\frac{2c_1(f_1-1)(\lambda+\mu)^2}{4c_1(\beta\lambda+\gamma\mu)-f_1r^2(\lambda+\mu)^2}, -\frac{2c_2(f_2-1)(\lambda+\mu)^2}{4c_2(\beta\lambda+\gamma\mu)-f_2r^2(\lambda+\mu)^2}$	I, J	$-\frac{2c_1(f_1-1)(\lambda+\mu)^2}{4c_1(\beta\lambda+\gamma\mu)-f_1r^2(\lambda+\mu)^2}, -\frac{2c_2(f_2-1)(\lambda+\mu)}{4c_2-f_2r^2(\lambda+\mu)}$
Allow	G, H	$\frac{2c_1(f_1-1)\lambda}{f_1r^2\lambda-4c_1\beta}, \frac{2c_2(f_2-1)\lambda}{f_2r^2\lambda-4c_2\beta}$	$-\frac{18c_1(f_1-1)\lambda\left(4c_1\beta-f_1r^2\lambda\right)(\lambda+\mu)^2}{(3f_1r^2\lambda(\lambda+\mu)-4c_1\beta(3\lambda+4\mu))^2}, L$
Fee	$-\frac{2c_1(f_1-1)(\lambda+\mu)}{4c_1-f_1r^2(\lambda+\mu)}, -\frac{2c_2(f_2-1)(\lambda+\mu)^2}{4c_2(\beta\lambda+\gamma\mu)-f_2r^2(\lambda+\mu)^2}$	$K, -\frac{18c_2(f_2-1)\lambda\left(4c_2\beta-f_2r^2\lambda\right)(\lambda+\mu)^2}{(3f_2r^2\lambda(\lambda+\mu)-4c_2\beta(3\lambda+4\mu))^2}$	$-\frac{2c_1(f_1-1)(\lambda+\mu)}{4c_1-f_1r^2(\lambda+\mu)}, -\frac{2c_2(f_2-1)(\lambda+\mu)}{4c_2-f_2r^2(\lambda+\mu)}$

Table 3.A.3: Payoff matrix in the model with the content creators.

$$G = -\frac{2c_1(f_1-1)\lambda\left(4c_1\beta-f_1r^2\lambda\right)\left(-3f_2r^2(\lambda+\mu)^2+8c_2\gamma\mu+4c_2\beta(3\lambda+\mu)\right)^2}{(3f_1f_2\lambda(\lambda+\mu)^2r^4-4(c_1f_2\beta(\lambda+\mu)(3\lambda+4\mu)+c_2f_1\lambda(3\beta\lambda-\beta\mu+4\gamma\mu))r^2+16c_1c_2\beta(3\beta\lambda+4\gamma\mu))^2}$$

$$H = -\frac{2c_2(f_2-1)\left(3f_1r^2\lambda(\lambda+\mu)-4c_1\beta(3\lambda+2\mu)\right)^2\left(4c_2(\beta\lambda+\gamma\mu)-f_2r^2(\lambda+\mu)^2\right)}{(f_1\lambda(3f_2r^2(\lambda+\mu)^2+4c_2(-3\beta\lambda+\beta\mu-4\gamma\mu))r^2+4c_1\beta(4c_2(3\beta\lambda+4\gamma\mu)-f_2r^2(\lambda+\mu)(3\lambda+4\mu)))^2}$$

$$I = -\frac{2c_1(f_1-1)\left(3f_2r^2\lambda(\lambda+\mu)-4c_2\beta(3\lambda+2\mu)\right)^2\left(4c_1(\beta\lambda+\gamma\mu)-f_1r^2(\lambda+\mu)^2\right)}{(f_1(\lambda+\mu)(3f_2r^2\lambda(\lambda+\mu)-4c_2\beta(3\lambda+4\mu))r^2+4c_1(f_2\lambda(-3\beta\lambda+\beta\mu-4\gamma\mu)r^2+4c_2\beta(3\beta\lambda+4\gamma\mu)))^2}$$

$$J = -\frac{2c_2(f_2-1)\lambda\left(4c_2\beta-f_2r^2\lambda\right)\left(3f_1r^2(\lambda+\mu)^2-4c_1(2\gamma\mu+\beta(3\lambda+\mu))\right)^2}{(f_1(\lambda+\mu)(3f_2r^2\lambda(\lambda+\mu)-4c_2\beta(3\lambda+4\mu))r^2+4c_1(f_2\lambda(-3\beta\lambda+\beta\mu-4\gamma\mu)r^2+4c_2\beta(3\beta\lambda+4\gamma\mu)))^2}$$

$$K = -\frac{2c_1(f_1-1)(\lambda+\mu)\left(3f_2r^2\lambda(\lambda+\mu)-4c_2\beta(3\lambda+2\mu)\right)^2}{(4c_1-f_1r^2(\lambda+\mu))(3f_2r^2\lambda(\lambda+\mu)-4c_2\beta(3\lambda+4\mu))^2}$$

$$L = -\frac{2c_2(f_2-1)(\lambda+\mu)\left(3f_1r^2\lambda(\lambda+\mu)-4c_1\beta(3\lambda+2\mu)\right)^2}{(4c_2-f_2r^2(\lambda+\mu))(3f_1r^2\lambda(\lambda+\mu)-4c_1\beta(3\lambda+4\mu))^2}$$

Bibliography

V. Abhishek, P. Fader, and K. Hosanagar. Media exposure through the funnel: A model of multi-stage attribution. *SSRN*, 2016. 50, 53, 55

E. Anderl, I. Becker, F. V. Wangenheim, and J. H. Schumann. Putting attribution to work: A graph-based framework for attribution modeling in managerial practice. *SSRN*, 2343077, October 2013. 53

S. P. Anderson and J. S. Gans. Platform siphoning: Ad-avoidance and media content. *American Economic Journal: Microeconomics*, 3(4):1–34, 2011. 93

A. Ansari, K. Bawa, and A. Ghosh. A nested logit model of brand choice incorporating variety-seeking and marketing-mix variables. *Marketing Letters*, 6(3):199–210, 1995. 50

M. Armstrong, J. Vickers, and J. Zhou. Consumer protection and the incentive to become informed. *Journal of the European Economic Association*, 7(2-3):399–410, 2009. 95

M. Aseri, M. Dawande, G. Janakiraman, and V. S. Mookerjee. Ad-blockers: A blessing or a curse? *Working paper*, November 2017. 94

P. Bajari and A. Hortacsu. The winner's curse, reserve prices, and endogenous entry: Empirical insights from ebay auctions. *RAND Journal of Economics*, pages 329–355, 2003. 2, 4

P. Bajari, R. McMillan, and S. Tadelis. Auctions versus negotiations in procurement: an empirical analysis. *The Journal of Law, Economics, & Organization*, 25(2):372–399, 2008. 6

R. Berman. Beyond the last touch: Attribution in online advertising. *Working paper*, March 2015. 50, 51, 55

R. Berman. Beyond the last touch: Attribution in online advertising. August 2016. Available at SSRN: https://ssrn.com/abstract=2384211 or http://dx.doi.org/10.2139/ssrn.2384211. 53

J. R. Bettman, M. F. Luce, and J. W. Payne. Constructive consumer choice processes. *Journal of Consumer Research*, 25(3):187–217, December 1998. 51, 55

P. Bhardwaj and S. Balasubramanian. Managing channel profits: The role of managerial incentives. *Quantitative Marketing and Economics*, 3(3):247–279, 2005. 5

J. Bulow and P. Klemperer. Auctions versus negotiations. *The American Economic Review*, 86(1):180–194, 1996. 6

C. M. Campbell and D. Levin. Can the seller benefit from an insider in common-value auctions? *Journal of Economic Theory*, 1(91):106–120, 2000. 13

C.-H. Cho and H. J. Cheon. Why do people avoid advertising on the internet? *Journal of Advertising*, 33(4):89–97, 2004. 93

I.-K. Cho and D. M. Kreps. Signaling games and stable equilibria. *The Quarterly Journal of Economics*, 102(2):179–221, 1987. 14

M. Clancey. The television audience examined. *Journal of Advertising Research*, 34(4): S1–S1, 1994. 93

K. S. Corts. Third-degree price discrimination in oligopoly: All-out competition and strategic commitment. *The RAND Journal of Economics*, pages 306–323, 1998. 95

A. T. Coughlan and D. A. Soberman. Strategic segmentation using outlet malls. *International Journal of Research in Marketing*, 22(1):61–86, 2005. 95

D. Court, D. Elzinga, S. Mulder, and O. J. Vetvik. The consumer decision journey. *McKinsey Quarterly*, June 2009. 51, 55

B. Dalessandro, O. Stitelman, C. Perlich, and F. Provost. Causally motivated attribution for online advertising. *Proceedings of the Sixth International Workshop on Data Mining for Online Advertising and Internet Economy*, pages 7:1–7:9, 2012. 52, 55

J. A. Dearden and G. L. Lilien. On optimal salesforce compensation in the presence of production learning effects. *International Journal of Research in Marketing*, 7(2-3):179–188, 1990. 53

P. S. Desai. Quality segmentation in spatial markets: When does cannibalization affect product line design? *Marketing Science*, 20(3):265–283, 2001. 95

P. S. Desai and D. Purohit. "Let me talk to my manager": Haggling in a competitive environment. *Marketing Science*, 23(2):219–233, 2004. 95

P. S. Desai and K. Srinivasan. Demand signalling under unobservable effort in franchising: Linear and nonlinear price contracts. *Management Science*, 41(10):1608–1623, 1995. 14

Economist. Invisible ads, phantom readers. May 2016. 51

L. Einav, C. Farronato, J. D. Levin, and N. Sundaresan. Sales mechanisms in online markets: What happened to internet auctions? Technical report, National Bureau of Economic Research, 2013. 6

J. C. Ely and T. Hossain. Sniping and squatting in auction markets. *American Economic Journal: Microeconomics*, 1(2):68–94, 2009. 2, 5

M. Eswaran and A. Kotwal. The moral hazard of budget-breaking. *The Rand Journal of Economics*, pages 578–581, 1984. 53

J. Falkinger. A welfare analysis of "junk" information and spam filters. Technical report, Working Paper, Socioeconomic Institute, University of Zurich, 2008. 95

H. Fang and S. O. Parreiras. On the failure of the linkage principle with financially constrained bidders. *Journal of Economic Theory*, 110(2):374–392, 2003. 5

D. Fudenberg and J. Tirole. Game theory, 1991. 16

E. Gal-Or, M. Gal-Or, and A. Dukes. Optimal information revelation in procurement schemes. *The RAND Journal of Economics*, 38(2):400–418, 2007. 5

B. Glover and Y. Raviv. Revenue non-equivalence between auctions with soft and hard closing mechanisms: New evidence from yahoo! *Journal of Economic Behavior & Organization*, 81(1):129–136, 2012. 5, 16

V. Goel. In advertising, it's about who gets credit for the sale. *New York Times*, December 2014. URL `http://bits.blogs.nytimes.com/2014/12/26/in-advertising-its-about-who-gets-credit-for-the-sale/`. 50

K.-Y. Goh, K.-L. Hui, and I. P. Png. Privacy and marketing externalities: Evidence from do not call. *Management Science*, 61(12):2982–3000, 2015. 94

D. G. Goldstein, S. Suri, R. P. McAfee, M. Ekstrand-Abueg, and F. Diaz. The economic and cognitive costs of annoying display advertisements. *Journal of Marketing Research*, 51(6):742–752, 2014. 95

Google. About the default attribution models, 2017a. URL `https://support.google.com/analytics/answer/1665189?hl=en&ref_topic=3205717`. Accessed: March 2017. 50, 55, 61

Google. Data-driven attribution methodology, 2017b. URL `https://support.google.com/analytics/answer/3191594`. Accessed: March 2017. 50, 55

I.-H. Hann, K.-L. Hui, S.-Y. T. Lee, and I. P. Png. Consumer privacy and marketing avoidance: A static model. *Management Science*, 54(6):1094–1103, 2008. 94

B. Holmstrom. Moral hazard in teams. *The Bell Journal of Economics*, pages 324–340, 1982. 50, 53

B. Holmstrom and P. Milgrom. Aggregation and linearity in the provision of intertemporal incentives. *Econometrica: Journal of the Econometric Society*, pages 303–328, 1987. 53

T. Hossain. Learning by bidding. *The RAND Journal of Economics*, 39(2):509–529, 2008. 2, 5

S. Jain. Digital piracy: A competitive analysis. *Marketing Science*, 27(4):610–626, 2008. 95

B. Jiang, K. Jerath, and K. Srinivasan. Firm strategies in the “mid tail” of platform-based retailing. *Marketing Science*, 30(5):757–775, 2011. 14

J. P. Johnson. Targeted advertising and advertising avoidance. *The RAND Journal of Economics*, 44(1):128–144, 2013. 94

P. Jordan, M. Mahdian, S. Vassilvitskii, and E. Vee. The multiple attribution problem in pay-per-conversion advertising. In *Proceedings of the 4th international conference on Algorithmic game theory*, SAGT'11, pages 31–43, 2011. 52

E. Koutsoupias and C. Papadimitriou. Worst-case equilibria. In *Annual Symposium on Theoretical Aspects of Computer Science*, pages 404–413. Springer, 1999. 52, 63

V. Krishna. *Auction theory*. Academic Press, 2009. 5

T. Lan, D. Kao, M. Chiang, and A. Sabharwal. An axiomatic theory of fairness in network resource allocation. In *INFOCOM, 2010 Proceedings IEEE*, pages 1–9, March 2010. 55

H. Li and P. Kannan. Attributing conversions in a multichannel online marketing environment: An empirical model and a field experiment. *Journal of Marketing Research*, 51(1): 40–56, 2014. 53

S. Li, K. Srinivasan, and B. Sun. Internet auction features as quality signals. *Journal of Marketing*, 73(1):75–92, 2009. 5

W. Li and Z. Huang. The research of influence factors of online behavioral advertising avoidance. *American Journal of Industrial and Business Management*, 6(09):947, 2016. 93

D. Mayzlin and J. Shin. Uninformative advertising as an invitation to search. *Marketing Science*, 30(4):666–685, 2011. 5

R. P. McAfee and J. McMillan. Optimal contracts for teams. *International Economic Review*, pages 561–577, 1991. 53

P. R. Milgrom and R. J. Weber. A theory of auctions and competitive bidding. *Econometrica: Journal of the Econometric Society*, pages 1089–1122, 1982. 3, 5, 12, 13

S. Moorthy and K. Srinivasan. Signaling quality with a money-back guarantee: The role of transaction costs. *Marketing Science*, 14(4):442–466, 1995. 5

S. Mulpuru. The purchase path of online buyers. *Forrester Report*, March 2011. 51, 55

P. A. Naik, K. Raman, and R. S. Winer. Planning marketing-mix strategies in the presence of interaction effects. *Marketing Science*, 24(1):25–34, Jan. 2005. 50

A. Ockenfels and A. E. Roth. Late and multiple bidding in second price internet auctions: Theory and evidence concerning different rules for ending an auction. *Games and Economic behavior*, 55(2):297–320, 2006. 2, 5

PageFair. The state of the blocked web. *Global Adblock Report*, February 2017. 89, 90, 91

PageFair and Adobe. The cost of ad blocking. *Ad Blocking Report*, 2015. 88, 91

A. Pazgal, D. Soberman, and R. Thomadsen. Profit-increasing consumer exit. *Marketing Science*, 32(6):998–1008, 2013. 95

M. Peitz and T. M. Valletti. Content and advertising in the media: Pay-tv versus free-to-air. *international Journal of industrial organization*, 26(4):949–965, 2008. 95

M. Perry and P. J. Reny. On the failure of the linkage principle in multi-unit auctions. *Econometrica*, 67(4):895–900, 1999. 5

A. Prasad, V. Mahajan, and B. Bronnenberg. Advertising versus pay-per-view in electronic media. *International Journal of Research in Marketing*, 20(1):13–30, 2003. 95

PWC. Cross channel attribution is needed to drive marketing effectiveness. Technical report, 2014. 50

M. Rabin. Incorporating fairness into game theory and economics. *The American Economic Review*, 83(5):1281–1302, 1993. 55

V. Ramaswamy, W. S. Desarbo, D. J. Reibstein, and W. T. Robinson. An empirical pooling approach for estimating marketing mix elasticities with pims data. *Marketing Science*, 12 (1):103–124, 1993. 50

E. B. Rasmusen. Strategic implications of uncertainty over one's own private value in auctions. *Advances in Theoretical Economics*, 6(1):1–22, 2006. 2, 4

A. E. Roth and A. Ockenfels. Last-minute bidding and the rules for ending second-price auctions: Evidence from ebay and amazon auctions on the internet. *American economic review*, 92(4):1093–1103, 2002. 2

T. Roughgarden and É. Tardos. How bad is selfish routing? *Journal of the ACM (JACM)*, 49(2):236–259, 2002. 52, 63

S. Scott. The $8.2 billion adtech fraud problem that everyone is ignoring. *Techcrunch*, January 2016. 51

Z. Seyedghorban, H. Tahernejad, and M. J. Matanda. Reinquiry into advertising avoidance on the internet: A conceptual replication and extension. *Journal of Advertising*, 45(1): 120–129, 2016. 93

G. Shaffer and Z. J. Zhang. Competitive coupon targeting. *Marketing Science*, 14(4):395–416, 1995. 95

X. Shao and L. Li. Data-driven multi-touch attribution models. KDD '11, pages 258–264, 2011. 52

L. S. Shapley. *A Value for n-Person Games.* Princeton University Press, 1953. 55, 85

D. Simester. Signalling price image using advertised prices. *Marketing Science*, 14(2):166–188, 1995. 14

D. A. Soberman. Simultaneous signaling and screening with warranties. *Journal of Marketing Research*, 40(2):176–192, 2003. 5

P. S. Speck and M. T. Elliott. Predictors of advertising avoidance in print and broadcast media. *Journal of Advertising*, 26(3):61–76, 1997. 93

T. Stühmeier and T. Wenzel. Getting beer during commercials: Adverse effects of ad-avoidance. *Information Economics and Policy*, 23(1):98–106, 2011. 95

U. Subramanian and R. C. Rao. Leveraging experienced consumers to attract new consumers: An equilibrium analysis of displaying deal sales by daily deal websites. *Management Science*, 62(12):3555–3575, 2016. 5

J. Tåg. Paying to remove advertisements. *Information Economics and Policy*, 21(4):245–252, 2009. 95

R. van den Brink. An axiomatization of the shapley value using a fairness property. *International Journal of Game Theory*, 30(3):309–319, 2002. 55

N. Vratonjic, M. H. Manshaei, J. Grossklags, and J.-P. Hubaux. Ad-blocking games: Monetizing online content under the threat of ad avoidance. In *The Economics of Information Security and Privacy*, pages 49–73. Springer, 2013. 95

R. Wang. Auctions versus posted-price selling. *The American Economic Review*, pages 838–851, 1993. 6

K. C. Wilbur. A two-sided, empirical model of television advertising and viewing markets. *Marketing science*, 27(3):356–378, 2008. 94

R. T. Wilcox. Experts and amateurs: The role of experience in internet auctions. *Marketing Letters*, 11(4):363–374, 2000. 2, 4

L. Wintr. Some evidence on late bidding in ebay auctions. *Economic Inquiry*, 46(3):369–379, 2008. 5

L. Xu, J. A. Duan, and A. Whinston. Path to purchase: A mutually exciting point process model for online advertising and conversion. *Management Science*, 60(6):1392–1412, 2014. 53

H. Yoganarasimhan. The value of reputation in an online freelance marketplace. *Marketing Science*, 32(6):860–891, 2013. 4

J. Zhang. Policy and inference: The case of product labeling. Technical report, Working Paper, MIT, 2013. 5