Title
Essays on the Role of Externalities in Information Goods

Author
Dao, Duy Duc

Publication Date
2017

UNIVERSITY OF CALIFORNIA, SAN DIEGO

Essays on the Role of Negative Externalities in Information Goods

A dissertation submitted in partial satisfaction of the
requirements for the degree
Doctor of Philosophy

in

Management

by

Duy Duc Dao

Committee in charge:

Professor Terrence August, Chair
Professor Hyoduk Shin, Co-Chair
Professor Ery Arias-Castro
Professor Sanjiv Erat
Professor Vincent Nijs
Professor Kevin Xiaoguo Zhu

2017

The dissertation of Duy Duc Dao is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

Co-Chair

Chair

University of California, San Diego

2017

DEDICATION

To the mentors I've had during different stages of my life.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I would like to thank Terry and Shin for their guidance and patience throughout my time here at Rady.

I am also very thankful for the support from all of my committee members.

I would also like to thank Kihoon for his support and feedback.

I am very lucky to have had the love and emotional support of my parents, my sister Chi, and my girlfriend Janell.

Chapter 1, in full, is a reprint of the material as it appears in "Optimal Timing of Sequential Distribution: The Impact of Congestion Externalities and Day-and-Date Strategies," 2015, *Marketing Science*, **34**(5), 755-774. Terrence August, Duy Dao, and Hyoduk Shin. The dissertation author was the primary researcher and author of this material.

Chapter 2, in full, has been submitted for publication of the material as it may appear in "Market Segmentation and Software Security: Pricing Patching Rights," in *Management Science*. Terrence August, Duy Dao, and Kihoon Kim. The dissertation author was the primary researcher and author of this material.

Chapter 3, in part, is currently being prepared for submission for publication. Terrence August, Duy Dao, and Kihoon Kim. The dissertation author was the primary researcher and author of this material.

VITA

2011	B. S. in Mathematics *cum laude*, University of California, San Diego
2016-2017	Graduate Teaching Assistant, Rady School of Management; University of California, San Diego
2017	Ph. D. in Management, Rady School of Management; University of California, San Diego

PUBLICATION

Optimal Timing of Sequential Distribution: The Impact of Congestion Externalities and Day-and-Date Strategies (with Terrence August and Hyoduk Shin), 2015, *Marketing Science*, **34**(5), 755-774.

FIELD OF STUDY

Innovation, Technology, and Operations

Studies in Information Systems

Professors Terrence August and Hyoduk Shin

ABSTRACT OF THE DISSERTATION

Essays on the Role of Negative Externalities in Information Goods

by

Duy Duc Dao

Doctor of Philosophy in Management

University of California, San Diego, 2017

Professor Terrence August, Chair
Professor Hyoduk Shin, Co-Chair

In a wide variety of settings, the value one derives from purchasing and consuming a product depends on choices other consumers make. In particular, negative externalities such as congestion costs and cybersecurity risks not only impact consumers but they also affect a product's profitability. My dissertation examines novel content release strategies and pricing mechanisms that can be used in settings where choices consumers make can negatively affect a product's value.

In Chapter 1, I study a movie studio's pricing and channeling decisions when releasing a product to congestion-sensitive consumers. Over the past fifteen years, the time between theatrical release to a movie's release on another channel has decreased from seven months on average to about 110 days. In recent years,

more films have gone directly to home video without theatrical release. Studios are even experimenting with "day-and-date" strategies, distributing a release across distinct channels on the same day. In this chapter, I develop a game-theoretic model of film distribution and consumption to understand how studios should optimally price and time the release of video versions of their films, given that consumers are making strategic decisions about how to consume the product. I characterize conditions under which direct-to-video, day-and-date, and delayed release strategies maximize profitability for a studio.

In Chapters 2 and 3, I investigate pricing mechanisms to improve cybersecurity. In Chapter 2, I establish how software vendors can differentiate their products by offering "patching rights" and how the optimal pricing of these rights can segment the market in a manner that leads to both greater security and greater profitability. I characterize the price for these rights, the discount provided to those who relinquish rights, and the consumption and protection strategies taken by users as they strategically interact due to the security externality associated with product vulnerabilities. In Chapter 3, I study the ability of taxes to achieve an analogous effect in the open-source domain. In this domain, I demonstrate why large populations of unpatched users remain even when automatic updating is available, and then characterize how taxes on patching rights should optimally be structured.

Chapter 1

Optimal Timing of Sequential Distribution: The Impact of Congestion Externalities and Day-and-Date Strategies

MARKETING SCIENCE
Vol. 34, No. 5, September–October 2015, pp. 755–774
ISSN 0732-2399 (print) | ISSN 1526-548X (online)

http://dx.doi.org/10.1287/mksc.2015.0936

Optimal Timing of Sequential Distribution: The Impact of Congestion Externalities and Day-and-Date Strategies

Terrence August
Rady School of Management, University of California, San Diego, La Jolla, California 92093;
and Korea University Business School, Seoul 136-701, Korea, taugust@ucsd.edu

Duy Dao, Hyoduk Shin
Rady School of Management, University of California, San Diego, La Jolla, California 92093
{dddao@ucsd.edu, hdshin@ucsd.edu}

The window between a film's theatrical and video releases has been steadily declining with some studios now testing day-and-date strategies (i.e., when a film is released across multiple channels at once). We present a model of consumer choice that examines trade-offs between substitutable products (theatrical and video forms), the possibility of purchasing both alternatives, a congestion externality affecting consumption at theaters with heterogeneous consumer groups, and a decay in the quality of the content over time. Our model permits a normative study of the impact of shorter release windows (zero–three months) for which there is a scarcity of relevant data. We characterize the market conditions under which a studio makes video release time and price selections indicative of direct-to-video, day-and-date, and delayed video release tactics. During seasons of peak congestion, we establish that day-and-date strategies are optimal for high-quality films with high content durability (i.e., films whose content tends to lead consumers to purchase both alternatives) whereas prices are set to perfectly segment the consumer market for films with low content durability. We find that lower congestion effects provide studios with incentives to delay release and price the video to induce multiple purchasing behavior for films with higher content durability. However, an increase in congestion effects can, in certain cases, actually lead to higher studio profitability. We also show that, at the lower range of quality, an increase in movie quality should often be accompanied by a later video release time. Surprisingly, however, we observe the opposite result at the upper range of movie quality: an increase in quality can justify an earlier release of the video.

Keywords: channel relationships; game theory; marketing-operations interface; film industry
History: Received: February 16, 2011; accepted: April 21, 2015; Preyas Desai served as the editor-in-chief and John Zhang served as associate editor for this article.

1. Introduction

Since the late 1990s, a dramatic change has taken place in how movie studios manage film distribution. In these years, the average time between the initial theatrical release of a film and its debut on digital video has dropped almost 40%—from 179 days in 1999 to only 115 days in 2013 (see Table 1). This forward shift of two months has partly been driven by advances in technology, especially video quality and home theater capabilities, but also by an industry that is now challenging antiquated norms (Grover 2005, Cole 2007). Is it profit maximizing for a release window between sequential distribution channels to *always* exist? In recent years, more films have been skipping theatrical release entirely and going directly to home video as part of a direct-to-video strategy (Barnes 2008); in some cases, studios are even entertaining "day-and-date" strategies, which strike at the heart of the matter (Miller 2012, Tartaglione 2013). A day-and-date strategy typically means that a product is released across two or more distinct channels on the same day. For example, in 2006, the film *Bubble* was simultaneously released across all channels by 2929 Entertainment, a company founded by Mark Cuban and Todd Wagner, that has vertically integrated across production, distribution, and exhibition—an opportune proving ground for testing such strategies (Kirsner 2005). In fact, many argue that release windows are inherently inefficient since the positive impacts of early promotional spending are not fully captured (Gross 2006). Disney CEO Robert Iger even comments that a film should be released faster on digital video since it has "... more perceived value to the consumer because it's more fresh" (Marr 2005); perhaps not surprisingly, Disney announced an early video release of *Alice in Wonderland* only

Table 1 **Shrinking of Industry Average Video-Release Window from 1998 to 2013**

Year	Release window (days)	Year	Release window (days)
1998	200.4	2006	129.2
1999	179.1	2007	126
2000	175.7	2008	127.8
2001	165.4	2009	123.2
2002	171.4	2010	121.4
2003	153	2011	120.2
2004	145.8	2012	114.7
2005	141.8	2013	115.1

Source. Tribbey (2013).

12.5 weeks after its theatrical release instead of the typical 16.5 weeks at the time (Smith and Schuker 2010). Vogel (2007) predicts that further changes in studios' sequential distribution strategies will continue to occur; as a result, there is a strong need for new research aiming to provide a better understanding of these strategies.

Although the release window is gradually narrowing and a few films have been released using day-and-date strategies, it remains difficult to ascertain the impact of a substantial reduction in the release window on profitability. From an empirical standpoint, the average release window is still approximately three to four months, and there is very little data on any films with windows ranging from zero to three months. In prior studies, researchers have also commented on the low variance observed in the release window measure (see, e.g., Lehmann and Weinberg 2000). Thus, accurately predicting the effect of releasing a film simultaneously on video or even one month after theatrical release continues to be quite difficult, although some studies have sought to close this gap using surveys (Grover 2006, Hennig-Thurau et al. 2007).

However, one can gain significant insights into how various release strategies would tend to affect consumption and hence profitability if we enhance our understanding of the economic trade-offs consumers face when choosing between theatrical and video alternatives. In this paper, we take a normative approach to studying the theater-video windowing problem. We model the primary economic incentives of consumers making film consumption decisions and subsequently analyze how consumers behave, in equilibrium, as the video release time is varied over its full span. By taking into account strategic consumer behavior, we can explore how studios should set the video release time and price based on market conditions, movie characteristics, and operational factors. Although consumption by moviegoers can be affected by a wide range of other considerations, in our model we restrict our attention to four primary considerations: (i) quality decay of the film content over time, (ii) the quality/price gap between theatrical and video alternatives, (iii) residual value of the video alternative in addition to movie consumption (multiple purchases), and (iv) theater congestion.

For example, all else being equal, consumers prefer to see a film earlier rather than later (Marr 2005). Whether the content is viewed in a theater or at home on video, due to "buzz" generated by marketing, film critics, and social circles, consumers derive the highest perceived value at launch, which then decreases over time (Thompson 2006, Cole 2007, Smith and Schuker 2010). Because consumers who watch a film in theaters might also purchase videos of the film, the studios can affect such multiple purchasing behavior by moving up the video release date and pricing it accordingly (Moul and Shugan 2005).

On the other hand, there are two effects that can incentivize viewers to postpone consumption. First, if a film is sequentially distributed through separate channels, some consumers may prefer to wait for lower prices in the secondary channel even though the value derived from the film decays during that time. Thus, a *substitution effect* tends to shift consumption later because of lower prices in the subsequent channel. Second, a *congestion effect* can also shift consumption later in time. Many popular films sell out screenings right after their theatrical release. In 2008, *The Dark Knight* was selling out so many screens at midnight on opening day that exhibitors hurried to boost capacity by adding screenings at 3 A.M. and 6 A.M. (Cieply 2008). Furthermore, when Michael Jackson's comeback concert footage was assembled into a film titled, *This Is It*, it sold out over 1,000 screens (Jurgensen 2009). Anticipating sold-out screenings, some consumers may prefer to delay their viewing. However, a film need not sell out to induce consumers to delay. Even higher utilization of theaters leads to longer waits at ticketing and concessions, poorer seat choices in auditoriums, and other undesirable crowd externalities (e.g., crying babies, noise from conversations, and temperature problems). Hence, some consumers may avoid viewing a movie at the theater because of congestion-generated problems. Because congestion also strongly interacts with the studio's choice of video release time and its price due to substitution effects, optimal management of the entire system requires careful coordination.

In this paper, our research objective is to develop an understanding of how a studio should coordinate the video release time and price for its film as part of a comprehensive strategy to manage the film's total profitability across the theater and home video channels. Because earlier release times greatly impact cannibalization, it is critical that we capture the preferences of consumers who make multiple purchases, i.e., pay to see the film in a theater as well as purchase the video, as part of the central trade-off. Such preferences tend to mitigate cannibalization losses, but we also parameterize the delay after which multiple purchasing consumers actually obtain residual value to study regions of limited mitigation. Because delaying consumption due to congestion becomes more appealing with earlier video release times, we capture different classes of congestion sensitivity in the consumer population. Using our model, we fully characterize market conditions under which the optimal video release time and price give rise to direct-to-video, day-and-date, perfect segmentation, and delayed release tactics that maximize profitability for the studio. In each case, we provide an analytical characterization of the optimal video release time and price, and we offer studios practical insights on how to utilize each lever to shape consumer demand toward preferable outcomes.

2. Literature Review

Eliashberg et al. (2006) provide an extensive review of research related to the motion picture industry. In their discussion of the distribution stage, they pose several questions in regard to the substitutability of DVDs for theatrical consumption and how consumers make trade-offs between the two product forms. We investigate these topics at the consumer level to study various video release strategies while generating broader implications on how to optimally manage sequential distribution; thus, our work is closest in nature to research that examines the time window between theatrical and video release.

Frank (1994) studies the timing of sequential distribution by constructing a theoretical model in which potential revenue functions for both product forms linearly decrease in time. Minimizing the sum of the opportunity cost of a video release in the theatrical market and the opportunity cost of a delay of the video release, he characterizes the optimal, positive video release time. Lehmann and Weinberg (2000) construct a reduced-form model in the context of a video rental firm managing inventory ordering decisions and also analytically characterize the optimal release. In addition to inventory ordering decisions, Gerchak et al. (2006) also consider shelf-retention time, i.e., when to remove a video from front shelves, for a video rental chain. They show that a revenue sharing contract with wholesale price augmented by a licensing fee is optimal for a studio and coordinates the channel. Studying potential consumption of both versions in a model of intertemporal movie distribution, Calzada and Valletti (2012) show that versioning can be optimal for information goods with zero marginal costs. They further establish that a monopolist, or a centralized channel, will often price and simultaneously release both versions, whereas a sequential release strategy may be driven by a vertical, decentralized channel structure found in the motion picture industry. Their work is closest in spirit to the current work, and we discuss our relative contribution in greater detail below.

The empirical analysis in Luan and Sudhir (2006) is also related to our work since they account for buzz decay and multiple purchases in their utility specification. In their study, they find that both highly rated films and animated films tend to be less substitutable and that, on average, the optimal release window should be 2.5 months. In another empirical study, Hennig-Thurau et al. (2007) examine multiple channels and release orderings. Although the introduction of DVD rentals is country dependent, they similarly find that DVD sales should optimally be delayed by three months. Nelson et al. (2007) study the time gap between the end of a film's theatrical run and its release on DVD, finding that about 30% of films have DVD versions released while the film is still in theaters and that the time gap is generally declining.

Our work is also related to papers that study the impact of a social presence on consumers. Harrell et al. (1980) find that perceived crowding negatively affects shopping behavior as consumers employ adaptation strategies. Hui et al. (2009) study shoppers' path behavior and zone density, finding that consumers might be attracted to higher density zones but shop less in them. Argo et al. (2005) demonstrate that increasing social presence tends to positively affect emotions initially and then to have a more negative effect as the presence gets larger. In line with their finding, we focus on the impact of congestion on the theater-viewing experience since theater owners maximize their profits by allocating screens based on movie demand. Specifically, theater owners have economic incentives to keep their capacity highly utilized, and congestion-sensitive consumers are likely to be negatively affected at these higher levels of congestion.

In a broader context, sequential product introduction is related to intertemporal price discrimination (see, e.g., Coase 1972, Bulow 1982, Gul et al. 1986, Besanko and Winston 1990, Desai and Purohit 1998, and Desai et al. 2004). The papers in this literature demonstrate that a firm competes against itself by selling across different time periods without commitment. Exploring intertemporal pricing under capacity constraint, Su (2007) shows that markdown pricing can be optimal when high valuation customers are less patient and that, otherwise, optimal prices are increasing in time. Moorthy and Png (1992) consider a seller with two substitutable and differentiated products and show that sequential introduction can both effectively reduce cannibalization and be more profitable than simultaneous introduction. Studying films with short runs, Waterman and Weiss (2010) find that the release window for videos is still long and invariant to theatrical run length. Their results suggest that studios can credibly commit to release windows. We examine sequential introduction where, on theatrical movie release, the studio commits to a video release time and price.

In this paper, we build on the ideas summarized above with the goal of developing a theoretical framework in which we start from the consumer's choice problem to clarify the trade-offs consumers make and how demand for each product form arises as the result of equilibrium strategic behavior. Augmenting extant literature, we explicitly consider additional relevant factors on the studio's optimal video release time and pricing strategy: (i) an endogenously arising congestion externality at theaters with heterogeneity in consumer sensitivity to congestion; (ii) a time delay after which consumers who buy both versions obtain residual value associated with the second purchase; and (iii) a decaying quality over time. By studying the decision problems faced by consumers and the studio in the presence of these factors, we provide several new insights into how optimal film release strategies should be adapted to the factors across diverse conditions.

Although the model employed in Calzada and Valletti (2012) shares some of our model's features, our focus on the above factors leads to an enriched understanding of decision making in this context. Congestion is as important of a factor on the consumer's decision as video release time, and it is imperative to assess how congestion effects influence the studio's optimal release strategy. In a survey of a sample of the moviegoer population, we found that 45% out of 209 respondents of the moviegoer population chose either "very likely to delay" or "definitely will delay" in response to a question of whether they would consider delaying film consumption because of crowds at theaters (73% when including "somewhat likely to delay"). Moreover, 103 respondents ranked congestion as providing greater delay incentives than the video release time itself. Our study is the first to examine this important factor, and we formally demonstrate that congestion has a subtle, moderating effect between film content durability and optimal release time. We show that for high-quality films, under low congestion effects, an increase in film durability should be coupled with a *delayed* video release. However, when congestion effects are substantial, it is a more profitable strategy to release the video immediately under both high and low film durability. Even though congestion provides consumers with additional incentives to substitute toward what is typically characterized as the less profitable video channel, a studio may still find it optimal to release the video immediately under high congestion; this can be true despite substantial movie revenues being cannibalized and even in conditions where no multiple purchases are induced in equilibrium.

Calzada and Valletti (2012) establish that for information goods, versioning is optimal as long as the degree of substitutability between alternatives is not too large. We model congestion as an endogenously determined externality in equilibrium that stems from the size of moviegoer demand. When consumers can strategically respond to congestion effects and effectively separate, the studio has greater incentives to version. In fact, we establish that a versioning strategy continues to be optimal for studios on a much broader region, including the region where the degree of substitutability between alternatives is quite high.

Because the consumer value derived from watching a film drops quickly in the typical three–four month release window, whereas financial discounting over the same time period is relatively negligible in comparison, it is important to separate out the film's decaying quality over time in the model. By doing so, in contrast to prior work, we find that sequencing (i.e., delaying the video release) can be optimal even if the studio and consumers are homogeneous with regard to financial discounting. Furthermore, by capturing a time delay after which consumers subsequently derive value in the case of multiple purchases, we characterize how the sequencing decision interacts with time delays in consumer preference. Finally, because of the additional context features captured by our model, we formally establish how the optimal video release time surprisingly responds in a nonmonotonic manner to changes in primitives such as the degree of substitutability and the congestion cost factor; this has not been established in prior work and highlights the importance of our focal factors. As a consequence, we add to this body of literature by providing a nuanced view of the interactive effects that film quality, congestion, and content durability have on the optimal strategy that should be pursued by studios.

3. Model and Consumer Market Equilibrium

There is a continuum of consumers who are heterogeneous in their sensitivities to the quality of a cinematic production. Each consumer's sensitivity (i.e., her type) is uniformly distributed on $\mathcal{V} \triangleq [0, 1]$. We assume that the product can be consumed in theaters (the *movie*) for a given price $p_m > 0$ at time zero, and consumed in digital form at home (the *video*) for $p_d > 0$. In the movie industry, there are many other channels for obtaining film content including pay-per-view on-demand services (e.g., Vudu), video rental services (e.g., Netflix, Blockbuster, and Redbox), and cable services (e.g., Time Warner and Comcast). Our model simplifies this setting and focuses on clarifying the main trade-offs between a primary and secondary channel. However, the insights derived from our analysis can be readily applied to the more complex setting.

Two of the primary factors identified in §1 that affect consumption are the quality/price gap between alternatives and quality decay. First, to capture the former factor, we adapt a standard model of vertical product differentiation to our specific setting (Shaked and Sutton 1983, 1987). The product consumed in theaters has an inherent level of quality given by $\gamma_m > 0$. For example, the box office hit *Avatar* would be associated with a higher value for γ_m because of its special effects and 3-D features. If a consumer with quality sensitivity $v \in \mathcal{V}$ views the movie in a theater, her maximum willingness to pay is given by $\gamma_m v$. Similarly, the inherent quality of the corresponding video is given by $\gamma_d > 0$. Second, as discussed earlier, there is substantial decay in the quality of the film itself over time not because the content has changed but because it is steadily losing its relevance (akin to *vintage-use* depreciation in Desai and Purohit 1998). In that sense, we can consider the video to be essentially a different product at each moment in time. If the video is released at time $T \in [0, 1]$ and consumer v purchases only the video, her maximum willingness to pay is $\gamma_d(1 - T)v$.

The third critical factor we identified relates to the residual value of the video for consumers who make multiple purchases. Because consumers often purchase videos of films they have already seen in theaters, it is important to permit consumption of both movie and video alternatives in the model. Should a consumer opt for multiple purchases by consuming both the movie and the video, her willingness to pay for the video is modified to $\delta\gamma_d(1 - \max(T, \xi))v$, where $\delta \in [0, 1]$ represents the *durability* of the film in terms of its content, and $\xi \in [0, 1]$ denotes the minimum time beyond which the consumer is again willing to pay for the durable content. For example, a larger value for δ indicates that consumers still derive considerable value from watching the film on video again after viewing it in a theater. As δ becomes smaller, the video becomes more of a substitute for the movie since the residual value associated with multiple purchasing diminishes. Because a consumer has less incentive to consume both alternatives, she tends toward the one providing higher net utility.

In our model, the content durability is a film characteristic, but it directly interacts with consumers' heterogeneous types. Specifically, consumers with the highest types are the ones with the greatest incentive to engage in multiple purchases. However, the durability of the content itself is dependent on the type of film produced. For example, children's movies such as *Pocahontas*, *Aladdin*, and *Cars* would likely be associated with a higher δ because their content maintains large residual value for repeated viewings. Similarly, films that have established subcultures (e.g., *Star Wars* and *Star Trek*) would also have higher durability. On the other hand, documentary films and historical dramas like *Hotel Rwanda*, in which the focus lies on being informative, may have relatively lower residual value after a first viewing in comparison to highly entertaining films. Luan and Sudhir (2006) also find that films with lower overall consumer ratings from reviews as well as films that are *R*-rated tend to have lower content durability.

We can now specify the timing of decisions and formally define the consumer strategy set. At the beginning stage, the production studio determines when to open its video distribution channel and sets the video price. The studio announces and commits to this video release time, which is denoted by T as well as the video price p_d. Subsequent to the announcement, each consumer decides whether to consume only the movie ($s = M$), only the video ($s = D$), both the movie and video ($s = B$), or neither alternative ($s = N$). The strategy set is thus denoted by $S \triangleq \{M, D, B, N\}$, and each consumer chooses the action $s \in S$ that maximizes her payoff.

When the product is consumed in a theater (i.e., either $s = M$ or $s = B$), congestion externalities arise because of the theater's fixed capacity and limited resources. For example, as the number of patrons seeing a movie at a theater grows large, there is an increased risk of screenings being sold out, having only poor seats remaining, and the viewing experience being degraded because of congestion externalities. Congestion is the fourth factor we capture in our model that critically affects moviegoer consumption. We use the term "congestion" in a vein similar to that of Vickrey (1955, p. 39–40) in his study of New York City's subway system where he states, "...where congestion occurs, the fare may fail to reflect the relatively high cost either of providing additional service at such times, or of the added discomfort to existing passengers occasioned by the crowding in of additional passengers." We highlight this point since traditional congestion costs in operations management literature stem from longer waiting times in service processes. In our context, however, consumers do not simply wait at theaters and incur costs until a screen becomes free; rather, they adapt by altering their consumption decision. For these reasons, we model congestion costs as proportional to the mass of consumers viewing the movie in a theater. Ceteris paribus, a consumer strives to consume earlier to increase her surplus due to quality decay, but congestion provides incentives to delay and substitute to alternative content forms.

Because consumers may vary in their sensitivity to congestion, we examine two classes of consumers: $\mathscr{C} = \{H, L\}$. Class H refers to consumers who are sensitive to congestion, and thus may have a positive congestion cost parameter denoted as $\alpha > 0$. We let σ: $\mathscr{V} \times \mathscr{C} \to S$ be a strategy profile of consumer actions and denote the mass of consumers choosing in-theater consumption with $D_m(\sigma)$. A class H consumer with quality sensitivity v obtains a net payoff of $\gamma_m v - \alpha D_m(\sigma) - p_m$ if she consumes the movie only (i.e., $\sigma(v, H) = M$), for example. On the other hand, a class L consumer is less sensitive to congestion and for convenience we assume her congestion parameter is zero such that her net payoff analogously becomes $\gamma_m v - p_m$. Finally, we denote the probability any given consumer $v \in \mathscr{V}$ belongs to class L and H with $\rho \in (0, 1)$ and $1 - \rho$, respectively. Fixing all other consumers to the strategy prescribed by σ_{-v}, we can summarize the net payoff to the consumer with quality sensitivity v when undertaking action s by

$$V(v, H, s, \sigma_{-v}) \triangleq \begin{cases} \gamma_m v - \alpha D_m(\sigma) - p_m \\ \quad + \delta\gamma_d(1 - \max(T, \xi))v - p_d, & \text{if } s = B; \\ \gamma_m v - \alpha D_m(\sigma) - p_m, & \text{if } s = M; \\ \gamma_d(1 - T)v - p_d, & \text{if } s = D; \\ 0, & \text{if } s = N; \end{cases} \quad (1)$$

for class H, and

$$V(v, L, s, \sigma_{-v}) \triangleq \begin{cases} \gamma_m v - p_m \\ \quad + \delta\gamma_d(1 - \max(T, \xi))v - p_d, & \text{if } s = B; \\ \gamma_m v - p_m, & \text{if } s = M; \\ \gamma_d(1 - T)v - p_d, & \text{if } s = D; \\ 0, & \text{if } s = N; \end{cases} \quad (2)$$

for class L. We focus our study on the parameter region where $\gamma_m > \gamma_d$ is satisfied so that the inherent quality of the theatrical experience is higher than the video (Vogel 2007). For example, going to the theater can be thought of as a complementary event that includes viewing the film and thus carries a higher quality. We also focus on the region where $\gamma_m > p_m$ is satisfied such that there exist consumers who can obtain a positive surplus at the theater. Finally, we assume that the movie price p_m is high enough that the video pricing decision is not constrained from above.

3.1. Consumer Market Equilibrium and the Studio's Problem

Taking the video release time T, video price p_d, and other model parameters as given, we derive the consumer market equilibrium. We can classify consumers by the product forms they consume: both the movie and the video (*both*), only the movie (*movie*), only the video (*video*), or nothing (*none*). First, we develop an understanding of what types of consumption outcomes occur under the various market conditions. For example, if a blockbuster movie is coupled with a fast video release time, to what extent will theater demand be cannibalized, particularly for high content durability films? By gaining insight into how moviegoers adjust their consumption patterns in response to durability, video release times, and congestion, we can more clearly see how release timing and pricing affect profitability, a subject we address in §4.

Thus, taking into account a film's quality decay over time, a congestion externality, and the availability of a video alternative, each consumer chooses an action that maximizes her own surplus. An equilibrium strategy profile σ^* must satisfy the following for each $v \in \mathscr{V}$ and $c \in \mathscr{C}$:

$$V(v, c, \sigma^*(v, c), \sigma^*_{-v}) \geq V(v, c, s, \sigma^*_{-v}), \quad \text{for all } s \in S. \quad (3)$$

Because of the monotone properties of (1) and (2), it follows that the equilibrium strategy profile σ^* is characterized by thresholds. In particular, there exist threshold values ω_b^c, ω_m^c, $\omega_d^c > 0$ (where $c \in \mathscr{C}$ refers to the consumer class) such that the equilibrium consumer strategy profile is given by

$$\sigma^*(v, c) = \begin{cases} B, & \text{if } \omega_b^c \leq v \leq 1; \\ M, & \text{if } \omega_m^c \leq v < \omega_b^c; \\ D, & \text{if } \omega_d^c \leq v < \omega_m^c; \\ N, & \text{if } v < \omega_d^c; \end{cases} \quad (4)$$

noting that (i) consumers with the lowest sensitivity to quality (i.e., low types) remain out of the market; (ii) consumers with slightly higher quality sensitivity purchase only the video alternative; (iii) consumers with even higher sensitivity choose to view the movie in theaters; and finally (iv) consumers with the highest quality sensitivity consume both the movie and video alternatives. This threshold structure holds for both consumer classes, H and L, who vary with regard to congestion costs. Whether any particular consumer segment *both*, *movie*, *video*, or *none* is present in equilibrium critically hinges on the underlying parameter region as well as how the studio strategically sets the video release time and price.

Taking into consideration the equilibrium consumer strategies developed above, we next lay out the studio's decision problem. We compute the demand for the movie by

$$D_m \triangleq \int_{\mathcal{V}} \left[\rho \mathbf{1}_{\{\sigma^*(v,L)\in\{B,M\}\}} + (1-\rho)\mathbf{1}_{\{\sigma^*(v,H)\in\{B,M\}\}}\right] dv, \quad (5)$$

which measures the population of consumers whose equilibrium strategy includes viewing the film in a theater. Similarly, we denote the aggregate demand for the video by

$$D_d \triangleq \int_{\mathcal{V}} \left[\rho \mathbf{1}_{\{\sigma^*(v,L)\in\{B,D\}\}} + (1-\rho)\mathbf{1}_{\{\sigma^*(v,H)\in\{B,D\}\}}\right] dv. \quad (6)$$

We denote the studio's share of movie and video revenues with λ_m and λ_d, respectively, where $\lambda_m, \lambda_d \in [0,1]$. Since the marginal cost of satisfying consumers, whether in a theater or by providing a video, is fairly small, we make a simplifying assumption that it is zero. Thus, the studio's profit function can be written as

$$\Pi(T, p_d) \triangleq \lambda_m p_m D_m + \lambda_d p_d D_d. \quad (7)$$

The main objective of this paper is to develop an understanding of how consumers adapt their consumption choices to changing video release times and price, and, subsequently, to characterize how studios can manage them by optimizing video release time and pricing. Hence, we take movie prices as fixed and exogenous to the model especially because uniform pricing has been the standard in this industry since the 1970s (Orbach and Einav 2007). The studio's problem can then be written as

$$\begin{aligned} \max_{T\in[0,1],\, p_d>0} \quad & \{\Pi(T, p_d)\} \\ \text{s.t.} \quad & \sigma^*(\cdot \mid T, p_d) \text{ satisfies (3).} \end{aligned} \quad (8)$$

As can be seen in (7), prices have a direct effect on studio profits whereas all parameters, including release time and prices, indirectly influence profitability through their impact on the strategic consumption behavior of consumers.

4. Optimal Release Time and Pricing

In this section, we develop the solution to the studio's profit maximization problem. By the formulation in (8), the solution to the studio's problem is a couple (T^*, p_d^*) corresponding to the optimal video release time and price. The studio can vastly change the equilibrium consumer market structure induced in both consumer classes by changing its release time and video price. In the full characterization of the consumer market equilibrium, there are 15 unique structure pairs that arise in equilibrium as T and p_d are varied. As an example, we will use shorthand notation such as [L: B-M-N] and [H: D-N] to conveniently express that *both*, *movie*, and *none* segments are represented in equilibrium in class L, whereas only *video* and *none* segments are present in class H. The optimal strategy, (T^*, p_d^*) together with the induced consumer market structure $\boldsymbol{\sigma}^*(\cdot \mid T^*, p_d^*)$ jointly can be thought of as a *tactic* being employed by the studio in a given parameter range.

First, we briefly describe how a studio should handle a movie with a low quality parameter. Because $\gamma_m > p_m$, consumers are always guaranteed positive surplus from consuming the movie, and the studio faces a trade-off. On one hand, the inherent quality of its movie offering is higher and can earn the studio a price premium. On the other hand, to incentivize moviegoers to consume the movie version, the studio necessarily must either increase the price of the video or delay its release to limit cannibalization of movie revenues. In either case, the studio's video revenues associated with the *both* and *video* consumer segments will be negatively affected. When γ_m is low, the potential market for the movie is smaller, and the trade-off shifts in favor of enhancing video revenues. It is straightforward to show that the studio's optimal strategy is given by $(T^*, p_d^*) = (0, \gamma_d/2)$, and because no one will consume the movie in equilibrium under this strategy, the studio need not release it in theaters. In this sense, the studio essentially pursues a direct-to-video tactic. In the film industry, the number of direct-to-DVD films has grown 36% since 2005 with 675 films being released in 2008, according to Adams Media Research, and the direct-to-DVD market generates approximately \$1 billion in annual revenues (Barnes 2008). Oftentimes, studios pursue a direct-to-video tactic for films that are of lower quality.

For the remainder of the paper, to simplify the presentation to the reader while retaining the main insights, we employ a binary discretization for several parameters at high and low values. Specifically, film content durability will take on either a high (δ_H) or low (δ_L) value; the congestion parameter will similarly take on either a high (α_H) or low (α_L) value; and the movie quality parameter will also take on either a high (γ_m^H) or intermediate (γ_m^I) value. We already argued that studios will employ a direct-to-video strategy for sufficiently low-quality movies so we focus our study on movies with sufficient quality that are released in theaters. Finally, because the effect of congestion is paramount to the current study, we will restrict our focus to a limited population of congestion-insensitive, class L customers, i.e., ρ will be kept at a lower level. To keep the mathematical analysis simple and clear, we will take $\delta_L = 0$, $\delta_H = 1$, and $\alpha_L = 0$ in the proofs in Appendix A, but all results generalize to regions

Table 2 **Studio's Optimal Video Release Time and Pricing Strategy**

	δ_H			δ_L		
	α_H		α_L	α_H	α_L	
γ_m^H	$T^*=0$ $p_d^*=\dfrac{\delta_H\gamma_d(1-\xi)}{2(\delta_H(1-\xi)(1-\rho)+\rho)}$ Day-and-date [L: B-M-N] [H: D-N]		$T^*=\xi$ $p_d^*=\dfrac{\delta_H\gamma_d(1-\xi)}{2}$ Delayed release [L: B-M-N] [H: B-M-N]	$T^*=0$ $p_d^*=\dfrac{\gamma_d}{2}$ Perfect segmentation [L: M-N] [H: D-N]	$\underline{\lambda_m<\lambda_d}$ $T^*=0$ p_d^* in (10) Day-and-date [L: M-D-N] [H: M-D-N]	$\underline{\lambda_m\geq\lambda_d}$ $T^*=1$ $p_d^*=\gamma_d$ Movie only [L: M-N] [H: M-N]
γ_m^l	$\underline{\lambda_m<\lambda_d\Phi}$ $T^*=0$ $p_d^*=\dfrac{\gamma_d}{2}$ Day-and-date	$\underline{\lambda_m\geq\lambda_d\Phi}$ $T^*=1-\delta_H(1-\xi)-\dfrac{\gamma_m^l}{\gamma_d}+\dfrac{2p_m}{\gamma_d}\sqrt{\dfrac{\lambda_m\rho}{\lambda_d}}$ $p_d^*=\dfrac{\gamma_m^l+\delta_H\gamma_d(1-\xi)}{2}-p_m\sqrt{\dfrac{\lambda_m\rho}{\lambda_d}}$ Delayed release	$T^*=\xi$ $p_d^*=\dfrac{\gamma_d(1-\xi)}{2}$ Delayed release [L: B-D-N] [H: B-D-N]	$T^*=0$ p_d^* in (11) Day-and-date [L: M-D-N] [H: D-N]		
	[L: B-D-N] [H: D-N]					

of low δ_L and α_L and high δ_H (generalized proofs are included in the online supplement (available as supplemental material at http://dx.doi.org/10.1287/mksc.2015.0936)). Sufficient bounds on γ_m^l and ρ, and other simplifying technical conditions are detailed in Appendix B. For the case of α_H, we employ asymptotic analysis as required.[1]

To lay out how our results are organized, we initially group them into these two cases of film durability. Within each case, we cover two subcases of movie quality. In the formal propositions, we then study outer case/subcase combinations as we vary the operational congestion factor. We also provide additional insights by holding the congestion factor and durability constant, and discussing how the studio's optimal strategy is affected as movie quality varies. Similarly, we then hold congestion and movie quality constant and vary film durability, which reveals interesting comparative statics on the studio's optimal release time behavior. Our results are summarized in Table 2, which we refer to throughout the analysis.

4.1. High Film Content Durability

We begin by examining the case of high content durability (δ_H) where the film retains high residual video value for consumers who also consume the movie format.

4.1.1. High Movie Quality. We first consider the subcase in which the quality of the theatrical offering is fairly high, i.e., γ_m^H, and the theatrical movie offering becomes a very lucrative channel for the studio.

Proposition 1. *For a film with high content durability, δ_H, and high quality, γ_m^H:*

(i) *Under a high congestion cost factor, α_H, the studio optimally releases the video immediately at $T^*=0$ and sets its price to $p_d^*=\delta_H\gamma_d(1-\xi)/(2(\delta_H(1-\xi)(1-\rho)+\rho))$. The studio's optimal strategy induces a consumer market structure characterized by [L: B-M-N] and [H: D-N], i.e., a day-and-date tactic is employed.*

(ii) *Under a low congestion cost factor, α_L, the studio optimally delays video release until $T^*=\xi$ and sets its price to $p_d^*=\delta_H\gamma_d(1-\xi)/2$. The studio's optimal strategy induces a consumer market structure characterized by [L: B-M-N] and [H: B-M-N], i.e., a delayed release tactic is employed.*

Proposition 1 highlights that as the congestion factor increases, a studio should optimally adjust its (T^*, p_d^*) strategy such that it makes a switch from a delayed video release to a day-and-date tactic where the video is released simultaneously in conjunction with the theatrical version. To see why congestion and the optimal release time are negatively associated in this region, we first discuss part (i) of the proposition. Congestion affects class H consumers by

[1] Because of the complexity in the analysis of the problem, asymptotic analysis has been commonly used in microeconomic studies, e.g., Li et al. (1987), Laffont and Tirole (1988), Muller (2000), Tunca and Zenios (2006), August and Tunca (2006), Pei et al. (2011), and August and Tunca (2011) among many others. Furthermore, comprehensive treatments of the mathematical techniques in asymptotic analysis are provided in Miller (2006).

decreasing their utility for the movie and increasing their incentive to substitute toward the video. A delayed release can help deter substitution from the movie to the video, however it will also drive low valuation video consumers out of the market and reduce the number of multiple purchases. When congestion costs are high, the delay would need to be substantial to effectively garner movie revenues from class H, and quite detrimental to these other two segments. Hence, the trade-off favors the opposite direction toward an earlier release to preserve class H video purchases and class L multiple purchases. In particular, the studio employs the strategy $(T^*, p_d^*) = (0, \delta_H \gamma_d (1-\xi)/(2(\delta_H(1-\xi)(1-\rho)+\rho)))$ that induces no class H consumer to view the movie in theaters in equilibrium, i.e., [H: D-N], but also induces a *both* segment from class L consumers, i.e., [L: B-M-N], which increases profitability. This provides insight into how the studio adapts its pricing strategy as the composition of consumers changes. For example, as ρ decreases (larger class H), the studio increases p_d^* to boost video revenues from class H. On the other hand, as ρ increases (larger class L), the studio decreases p_d^* to induce a larger *both* segment from class L in equilibrium. Notably, high content durability is a critical driver of multiple purchasing behavior and integral to these arguments.

In totality, we say that the studio pursues a day-and-date tactic in this region because it releases the video immediately and sets price to have all consumer segments represented in equilibrium, with some consumers making multiple purchases. Part (i) of Proposition 1 implies that a day-and-date tactic can be attractive for the studio for high content durability films with high theatrical quality, when the effects of congestion are substantial. One fitting example is a blockbuster children's movie released during periods of peak demand such as the holiday season. Commenting on the motion picture industry, Disney CEO Robert Iger said, "I don't think it's out of the question that a DVD can be released in effect in the same window as a theatrical release," suggesting that day-and-date video release strategies are being considered. Noting that childrens' movies often have high content durability, in an interview with *The Wall Street Journal* (Donaldson-Evans 2006), Iger also suggested selling DVDs of *Chicken Little* in theaters in which the movie was playing.

One issue with employing a day-and-date tactic is potential push back from theater owners, who are concerned that early video availability will cannibalize theatrical movie demand. A solution consistent with Iger is to sell early released DVDs *only* in theaters at patron exit areas such that these DVDs can be targeted to the *both* consumer market segment, which will not cannibalize movie demand. Another option that studios may consider to alleviate theater owners' concerns is to implement revenue sharing for video sales with theater owners, which has also been suggested by 2929 Entertainment (Grover 2006). In addition, one ancillary benefit of using a day-and-date tactic relates to advertisement costs, which amount to half of the total production cost on average (Vogel 2007). When a studio releases a film's video a few months after being released in theaters, it must once again incur additional advertising expenditures. Under a day-and-date tactic, a studio can consolidate a film's marketing budget into a single, shorter period, leveraging the initial buzz effectively across both movie and video offerings.

Next, we examine the studio's decision problem as capacity constraints play a lesser role. For instance, the shadow price of capacity is likely to be lower in winter and spring in comparison to summer and holiday seasons (Einav 2007). Other factors that may decrease the congestion parameter include how the number of screens has increased over time, as well as possible local changes in the number of seats per screen (NATO 2013). Recently, small theater chains such as Cinépolis have competed by focusing on a higher quality experience through offering luxury seating, alcoholic beverages, full-service cafes, and an extensive selection of menu options for dining (Abate 2012, Luna 2012). In such cases, although there are fewer seats per screen, every seat is a "good seat" with unobstructed views, adequate spacing, and the ability to recline, which effectively reduces the congestion cost.

When the congestion cost factor diminishes, class H consumers do not have as strong of an incentive to substitute from movie to video. Therefore, in contrast to the high congestion cost case, even a small delay in the release time can be an effective deterrent. As part (ii) of Proposition 1 conveys, in this case the studio delays release to preclude substitution by either class of consumers toward the video and focus its strategy on expanding movie revenues. Specifically, the studio's optimal strategy is $(T^*, p_d^*) = (\xi, \delta_H \gamma_d (1-\xi)/2)$. Notably, the studio still must limit the extent of the delay to protect the multiple purchasing behavior associated with the highly durable film content. Analytically we establish in this region that when $T \leq \xi$, p_d hinges only on ξ, hence profits are weakly increasing in T. However, when $T > \xi$, profits are decreasing in T, hence $T^* = \xi$. In Table 2, we summarize the results from Proposition 1 on the top row, left two columns.

4.1.2. Intermediate Movie Quality. Next, we study the studio's release time and pricing problem when the quality of the movie is at an intermediate level, γ_m^I. We shall see that the studio has increased incentives to delay the timing of video release in this intermediate subcase, which leads to an intriguing finding: in aggregate, the studio's optimal release time is nonmonotonic as the quality of the movie increases through its feasible space, from low to intermediate to high. To see why this nonmonotonicity arises, we first formalize the studio's optimal behavior for an intermediate region of movie quality.

Proposition 2. *For a film with high content durability, δ_H, and intermediate quality, γ_m^I:*

(i) *Under a high congestion cost factor, α_H, the studio optimally adjusts its video release and pricing strategy depending on its relative revenue share:*

- *If $\lambda_m \geq \lambda_d \Phi$, then the studio delays release until $T^* = 1 - \delta_H(1-\xi) - \gamma_m^I/\gamma_d + (2p_m/\gamma_d)\sqrt{\lambda_m \rho/\lambda_d} < \xi$ and sets its price to $p_d^* = (\gamma_m^I + \delta_H \gamma_d(1-\xi))/2 - p_m\sqrt{\lambda_m \rho/\lambda_d}$, i.e., a delayed release tactic is employed.*
- *If $\lambda_m < \lambda_d \Phi$, then the studio releases immediately with $T^* = 0$ and sets its price to $p_d^* = \gamma_d/2$, i.e., a day-and-date tactic is employed, where*

$$\Phi = \left(\frac{\gamma_m^I - \gamma_d(1-\delta_H(1-\xi))}{2p_m\sqrt{\rho}}\right)^2.$$

In both cases, the studio's optimal strategy induces a consumer market structure characterized by [L: B-D-N] and [H: D-N].

(ii) *Under a low congestion cost factor, α_L, the studio optimally delays video release until $T^* = \xi$ and sets its price to $p_d^* = \gamma_d(1-\xi)/2$. The studio's optimal strategy induces a consumer market structure characterized by [L: B-D-N] and [H: B-D-N].*

Part (i) of Proposition 2 stands in partial contrast to the direct-to-video tactic used by the studio under low-quality γ_m^L and the day-and-date tactic employed under γ_m^H when congestion cost is high; both of these tactics involve the studio optimally releasing the video at $T^* = 0$ as part of its optimal strategy. When the movie quality is at an intermediate level, consumers have increased incentives to substitute from the movie to the video. Similar to before, the video release would need to be significantly delayed to deter this substitution, which remains inefficient, and again an earlier video release can improve the video market for class H and multiple purchases in class L. The critical difference here is that when movie quality is only at an intermediate level, the studio has to be concerned with a different kind of substitution: consumers shifting from multiple purchases to the video. Here the equilibrium consumer market structure for class L consumers satisfies

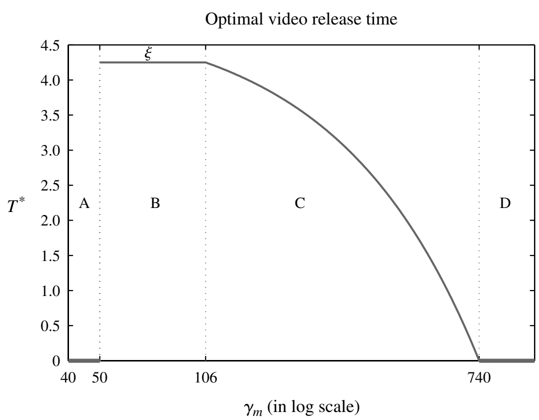
$$\sigma^*(v,L) = \begin{cases} B, & \text{if } \dfrac{p_m}{\gamma_m^I - \gamma_d((1-T)-\delta_H(1-\max(T,\xi)))} \leq v < 1; \\ D, & \text{if } \dfrac{p_d}{\gamma_d(1-T)} \\ & \quad \leq v < \dfrac{p_m}{\gamma_m^I - \gamma_d((1-T)-\delta_H(1-\max(T,\xi)))}; \\ N, & \text{if } v < \dfrac{p_d}{\gamma_d(1-T)}, \end{cases} \tag{9}$$

from which it can be seen how reducing T shifts some consumers from N (nothing) to D (video) but also others from B (both) to D (video). In this case, the studio optimally delays release to protect revenues from multiple purchases while sacrificing some video revenues at the low end of the consumer market, provided its share of movie revenues (λ_m) is high enough. The extent to which the video release time is optimally delayed depends critically on the composition of the consumer population. As the proportion of class H consumers increases (lower ρ), video revenues at the lower end become more impactful so the studio either delays release to a lesser extent or pursues a day-and-date tactic as before.

When the congestion cost parameter is low, even the congestion-sensitive class H consumers now have increased incentives to go to theaters. In contrast to part (ii) of Proposition 1, because the movie quality (γ_m^I) is now closer to the video quality (γ_d), it is relatively more difficult to get consumers to prefer the movie over the video. To do so, the video release time would need to be significantly delayed. Doing so would hamper the studio's ability to benefit from multiple purchases that can be induced in cases of high durability. Thus, in this case, the studio takes a balanced approach, using a moderate delay $T^* = \xi$. This is not a sufficiently large delay to induce a *movie* segment, but it does help prevent substitution from *both* to *video*, thus protecting multiple purchasing behavior. In equilibrium, the studio's strategy yields [L: B-D-N] and [H: B-D-N], and it prices the video at the monopoly level $p_d^* = \gamma_d(1-\xi)/2$ associated with this optimal delay.

4.1.3. Impact of Movie Quality. As technology rapidly evolves, implementation in theater equipment often precedes consumer electronics. For example, the last few years have seen a rebirth of 3-D theatrical releases driven by improvements in 3-D technology, as seen with *Avatar*, *Alice in Wonderland*, and *Clash of the Titans*. However, 3-D televisions have just

Figure 1 **The Impact of Theatrical Quality on the Optimal Video Release Time Under High Congestion Cost**



recently become available and have not yet achieved widespread adoption (Bonnington 2012). Thus, the relative quality of theatrical and video offerings can vary over time. Figure 1 illustrates how the optimal video release time changes in theatrical quality under a high congestion cost parameter, which summarizes some of the results in this section. People commonly believe that low theatrical quality films are the ones that should be released earlier to video (Epstein 2005). This intuition is often associated with the observation that lower quality films sometimes bypass theatrical release entirely and appear directly on home video. In particular, if γ_m is sufficiently low, the studio releases the video immediately as part of a direct-to-video tactic, which is illustrated in the left-hand portion of Figure 1, labeled as region A.

Although it seems reasonable that higher quality movies are more likely to have longer release windows, this conclusion is not always justified. For instance, for a hit film with extremely high γ_m, the studio releases the video immediately, utilizing a day-and-date tactic as demonstrated in part (i) of Proposition 1; this behavior is also illustrated in region D in Figure 1. As γ_m takes intermediate values, as in regions B and C, the studio optimally delays the video release time. Consistent with part (i) of Proposition 2, T^* moves from ξ as seen in region B to a release time strictly less than ξ as seen in region C. In these regions, not only has the video release time become strictly positive, but it also continuously decreases toward a day-and-date release tactic. This result has an important empirical implication. Because of reasonably high variability on the many dimensions that characterize films, we can expect that γ_m varies significantly film by film. As the movie industry moves toward having more day-and-date releases, our model suggests that we should see video release times spanning the feasible range. As Disney CEO Robert Iger commented, "we're not doing a one-size-fits-all approach" (Smith and Schuker 2010); we should therefore not expect videos to continue to be released according to an almost binary standard: either immediately or three–four months later. Instead, we anticipate seeing a more complete timespan utilized because of the diversity of film characteristics. This is a testable empirical implication of our model. Another important implication of our model is that an increase in the inherent quality of a movie, e.g., because of the 3-D theatrical releases, should sometimes be coupled with an *earlier* video release time as illustrated in regions B, C, and D of Figure 1. Hence, the optimal video release time is nonmonotonic in the inherent quality of the theatrical version of the film.

4.2. Low Film Content Durability

A less durable film is one that does not carry as much residual value for another viewing. A film's genre, its targeted demographic, and other characteristics can all affect its level of durability. In this section, we study how a studio's optimal video release timing and pricing strategy should be adjusted when a film's content durability is at a lower level, δ_L. We compare and contrast the studio's optimal strategy under δ_L to its strategy under δ_H, and discuss the manner in which the level of content durability affects the studio's incentives.

4.2.1. High Movie Quality. We again begin by analyzing the γ_m^H subcase.

Proposition 3. *For a film with low content durability, δ_L, and high quality, γ_m^H:*

(i) *Under a high congestion cost factor, α_H, the studio optimally releases the video immediately at $T^* = 0$ and sets its price to $p_d^* = \gamma_d/2$. The studio's optimal strategy induces a consumer market structure characterized by* [L: M-N] *and* [H: D-N], *i.e., a perfect segmentation tactic is employed.*

(ii) *Under a low congestion cost factor, α_L, the studio optimally adjusts its video release and pricing strategy depending on its relative revenue share:*

- *If $\lambda_d \leq \lambda_m$, then the studio prefers not to release a video version. The studio's optimal strategy induces a consumer market structure characterized by* [L: M-N] *and* [H: M-N].
- *If $\lambda_d > \lambda_m$, then $T^* = 0$ and*

$$p_d^* = \frac{\gamma_d(p_m(\lambda_m + \lambda_d) + \alpha_L\lambda_d(1-\rho))}{2\lambda_d(\gamma_m^H + \alpha_L(1-\rho))}. \tag{10}$$

The studio's optimal strategy induces a consumer market structure characterized by [L: M-D-N] *and* [H: M-D-N], *i.e., a day-and-date release tactic is employed.*

Under a large congestion factor, we see that similar to part (i) of Proposition 1 the studio also finds it preferable to release the video immediately even under δ_L. However, the driving force of its overall strategy is markedly different. As before, there exist incentives for releasing early to expand the *both* market for class L and the video market for class H. However, it is important to note that the optimal video price in the δ_H subcase satisfies $p_d^* = \delta_H \gamma_d(1-\xi)/(2(\delta_H(1-\xi)(1-\rho)+\rho))$. Thus, in that case, coupled with an immediate release the studio also needs to adjust price in a manner dependent on δ_H to incentivize multiple purchases. Under δ_L, the requisite price reduction would need to be significant because of lower content durability, and this would drastically hurt revenues generated from class H video consumers. In this case, the studio finds it more profitable to forgo the *both* segment and pursue a different strategy. Because of class H consumers' high congestion sensitivity, the studio is again less concerned with trying to deter cannibalization of movie demand. Instead, the studio releases the video immediately ($T^* = 0$) and sets the corresponding monopoly price for the video ($p_d^* = \gamma_d/2$). As a result, the classes separate; class L consumers compose the *movie* segment, and class H consumers compose the *video* segment. In this sense, the studio uses a perfect segmentation tactic under these conditions because of the lower content durability.

Part (ii) of Proposition 3 has both similarities and differences with part (ii) of Proposition 1, which underscore the impact of content durability on the studio's decisions. Under δ_H, we established that the studio can effectively deter substitution by delaying release because congestion costs are low, but it limits the extent of delay to maintain multiple purchases. However, as we saw above, under low content durability, it is not profit maximizing to induce multiple purchases. However, if the revenue share from the movie market is more lucrative (i.e., $\lambda_d \le \lambda_m$), the studio will still optimally delay video release to deter substitution. Moreover, in this case, it need not limit the amount of delay to protect multiple purchases; therefore, it instead delays video release to the point where only the movie is consumed in equilibrium. On the other hand, if the revenue share from the video market is more lucrative (i.e., $\lambda_d > \lambda_m$), the studio significantly adapts its strategy. Because it cannot induce multiple purchases with low content durability, its strategy must focus more on expanding video revenues while protecting movie revenues. Under high movie quality and low congestion costs, it is relatively difficult to induce video purchases. To achieve this expansion, the studio must both (i) release the video immediately at $T^* = 0$ and (ii) price it strategically lower at $p_d^* = \gamma_d(p_m(\lambda_m+\lambda_d)+\alpha_L\lambda_d(1-\rho))/(2\lambda_d(\gamma_m^H+\alpha_L(1-\rho)))$ to provide the necessary incentive to induce more video purchases. The greater the video revenue share λ_d, the more the studio is willing to cut its video price and expand the video market.

The movie industry has changed drastically over the past 15 years with many new channels (e.g., Netflix, Redbox, and online VOD such as Vudu and Hulu), new technologies (e.g., Blu-ray, HD streaming, and iTunes), and evolving consumers (e.g., higher broadband household penetration, pirated content availability, and customer impatience for content). In this dynamic environment, the studio's share of revenues in different channels also varies as players enter and exit and as negotiations take place. Part (ii) of Proposition 3 gives insight into how a studio may need to adapt its strategies when industry changes critically affect the revenue share it obtains in each channel. The results from this section are summarized in Table 2 on the top row, right two columns. The table illustrates how a decrease in congestion has different effects on the studio's strategy under high and low content durability. Under δ_H, the studio responds with a measured delay in the video release to protect the *both* market segment. However, under δ_L, because of its inability to profitably induce multiple purchases, the studio either institutes an extreme delay (completely deterring substitution to video) or switches to a video market expansion strategy depending on its video revenue share.

4.2.2. Impact of Congestion. We have shown how varying the congestion factor leads to different optimal studio strategies and their associated consumer market structures in equilibrium. Further, we also demonstrated that the effect of congestion is quite different, depending on the quality of the movie and durability of content. Next, we examine how the level of congestion impacts profitability. Because of theaters' capacity constraints, the congestion parameter can often be higher during periods when big-budget movies are released (e.g., Memorial Day, Fourth of July, Thanksgiving, and Christmas) as these films compete for a limited number of screens (Einav 2007). On the surface, one may think that an increase in the congestion cost parameter *always* leads to lower profits because it directly hurts the utility of class H consumers. However, in the following corollary, we establish that a higher congestion cost parameter can sometimes be beneficial to the studio.

Corollary 1. *For a film with low content durability, δ_L, and high quality, γ_m^H, an increase in the congestion cost parameter from α_L to α_H increases the studio's profit if either of the following conditions holds*: (i) $\lambda_d > \lambda_m \max(1, \bar{\lambda}_d)$; *or* (ii) $\lambda_m > \lambda_d > 4\lambda_m p_m/\gamma_d$, *where $\bar{\lambda}_d$ is characterized in the appendix.*

Corollary 1 carries an important message: congestion can sometimes increase the profitability of a film. In particular, this profit improvement can occur if the studio has negotiated an increased share of the revenue in the video channel in comparison to the theatrical channel; moreover, it can also occur when the video revenue share is *less than* the movie revenue share as long as the video revenue share is not too low. We saw in part (ii) of Proposition 3 that for a high-quality movie with low congestion costs, it takes an early release coupled with a price reduction to incentivize video purchases, which is costly to the studio. However, under high congestion costs, the congestion-sensitive consumers in class H react to the externality imposed on them by class L consumers. This provides much stronger incentives for class H to consume the video instead of the movie. With a larger congestion externality, the studio can increase its profits because it can perfectly segment the consumer market and charge class H consumers a price reflecting monopoly power over its content. Essentially, the studio leverages endogenously determined congestion as a tool to separate the market and increase profits, but, importantly, such a strategy only makes sense for films with low content durability. Corollary 1 suggests that releasing high-quality films with lower content durability during peak seasons has the potential to help increase returns to the studio.

The role of congestion here is connected to Desai and Purohit (1998), which studies the profitability of leasing versus selling strategies for a monopolist. A central element in their model is the mean depreciation factors under each strategy that determine how much residual value remains in bought and leased products as time passes. They demonstrate that differences in depreciation rates give rise to the optimality of a combined leasing and selling strategy. An important point being made is that a high rate of depreciation of the product being sold is helpful to the firm because it makes the good effectively more of a nondurable one. In our paper, congestion negatively affects the utility associated with the movie alternative, but, similar to Desai and Purohit (1998), it relaxes incentive compatibility constraints enabling the firm to better segment the movie and video markets. An interesting point made in our paper is that such an outcome can still prevail even when consumers themselves determine the equilibrium level of congestion by their consumption behavior (e.g., even under α_H, if no consumer prefers to watch the movie, then there are zero congestion costs).

4.2.3. Intermediate Movie Quality. Finally, for the case of low content durability, we turn our attention to the final subcase: movies with intermediate quality γ_m^I.

Proposition 4. *For a film with low content durability, δ_L, and intermediate quality, γ_m^I:*

(i) *Under a high congestion cost factor, α_H, the studio optimally releases the video immediately at $T^* = 0$ and sets its price to*

$$p_d^* = \frac{\gamma_d(\lambda_d(1-\rho)(\gamma_m^I - \gamma_d) + \rho p_m(\lambda_m + \lambda_d))}{2\lambda_d(\gamma_m^I - (1-\rho)\gamma_d)}. \tag{11}$$

The studio's optimal strategy induces a consumer market structure characterized by [L: M-D-N] and [H: D-N], i.e., a day-and-date tactic is employed.

(ii) *Under a low congestion cost factor, α_L, the studio optimally adjusts its video release and pricing strategy depending on its relative revenue share, in the same manner (i.e., under γ_m^I) given in part* (ii) *of Proposition 3.*

First, we discuss part (i) of Proposition 4 in relation to part (i) of Proposition 3, where we found that the studio prefers to release the video immediately as part of a perfect segmentation tactic. As the quality of the movie decreases from γ_m^H to γ_m^I, the studio still has strong incentives to release earlier; however, it becomes more difficult to induce consumption of only the *movie* option and attain perfect segmentation. In particular, if the studio releases earlier, now some class L consumers will shift consumption from *movie* to *video*. Thus, the studio faces a clear trade-off between maintaining a larger *video* market (with regard to both consumer classes) by releasing earlier and preventing cannibalization of its more valuable channel (in class L) by releasing later. In part (i) of Proposition 4, we establish that the studio's optimal strategy should still be to release immediately but then mitigate cannibalization of movie revenues by increasing the price of the video, as is characterized in (11). Releasing early maintains the quality of the video offering, which is important to class H consumers. This also enables it to price the video high and, in turn, achieves two purposes: (i) increasing revenues from the *video* segments, and (ii) reducing cannibalization from the *movie* segment to the *video* segment. Notably, this result holds true even if the revenue share for the movie is higher than for the video; that is, the studio will sacrifice movie demand and this margin to some degree by releasing early.

Second, it is worthwhile to contrast this result to that obtained under high content durability but for the same subcase (i.e., intermediate movie quality and high congestion costs). Under high content durability, part (i) of Proposition 2 shows that the studio employs a delayed release tactic when its movie revenue share is lucrative. However, under low content durability, part (i) of Proposition 4 shows that it is never in the best interest of the studio to delay release. This difference in optimal release timing is attributable to how the studio manages cannibalization between *both* and *video* segments when multiple purchasing behavior occurs (δ_H). Because both of these market segments consume the video and incur p_d, the studio necessarily needs to use its video release timing lever to throttle substitution. On the other hand, when a *both* segment does not arise in equilibrium (δ_L), the studio can control substitution between *movie* and *video* market segments primarily using its video pricing lever.

By combining all results in Table 2, we gain a better understanding of how changes in content durability affect the video release strategy, when holding the movie quality and congestion factor classifications fixed. As an example, it is worth considering the case of a film with intermediate quality γ_m^I and under a high congestion factor α_H. Provided that the studio obtains a large share of movie revenues, we see that $T^* = 0$ for δ_L, whereas $T^* = 1 - \delta_H(1-\xi) - \gamma_m^I/\gamma_d + (2p_m/\gamma_d)\sqrt{\lambda_m \rho/\gamma_d}$ for δ_H. An immediate insight derived from this analysis is that the optimal video release time is also nonmonotonic in content durability. That is, T^* increases with a shift from δ_L to δ_H, and then decreases with higher δ_H. This finding complements the work of Calzada and Valletti (2012) by demonstrating conditions under which the optimal video release time increases in content durability instead of decreases. More generally, our model suggests that if there exists sufficient variation in content durability among movies, we can expect to see video release times become more film specific and possibly span feasible window lengths.

5. Discussion and Concluding Remarks

In this paper, we present a model of film distribution and consumption to gain insight into how studios should optimally price and time the release of video versions of their films when accounting for strategic behavior of consumers in product choice. We take a normative approach to the studio's problem and highlight the critical factors that can motivate the studio to choose video release and pricing strategies; these strategies can be characterized as direct-to-video, day-and-date, perfect segmentation, and delayed release tactics. In developing the consumer model, we incorporate several important factors: (i) quality decay of the film content over time; (ii) the quality/price gap between theater and video alternatives; (iii) consumption of both the theatrical and video version and quantification of these preferences using a notion of content durability (including delayed value realization for consumers who purchase both); and (iv) either negative consumption externalities associated with congestion at theaters or the absence of related costs for heterogeneous classes of consumers. These factors, which are in turn influenced by video release timing and pricing, together affect consumption behavior. Table 2 provides an overall summary of our findings. We analyze a variety of dimensions organized by case (two levels of content durability), subcase (two levels of movie quality), and factor (two levels of congestion costs).

Using our model, we establish a wide range of relevant insights for studios. Our results also have several testable implications.

1. Focusing on intermediate-to-high movie quality: as theatrical movie quality increases, the video release time decreases.

2. Focusing on lower movie quality: as a film's theatrical movie quality decreases, it is more likely to be released using a direct-to-video tactic. Overall, the optimal video release time is nonmonotonic in the quality of the movie.

3. For high-quality movies with high content durability, as congestion in theater increases, e.g., because of peak season capacity constraints, more day-and-date strategies will be implemented.

4. When congestion effects are low, films with higher content durability are more likely to be released later; however, films with lower content durability should be released using a day-and-date tactic if the studio's revenue share of the video channel is more attractive than the theater channel.

These testable implications require parameters such as content durability to be measured. Such parameters can be forecasted from data on consumer viewing habits. For example, data from Nielsen can help determine how durable different genres of film are as well as how content durability might relate to a film's target audience.

A number of important questions remain for future research. In this work, we focus our attention on the optimal video release timing and pricing strategy for a single movie. This setting is important since studios have some degree of freedom in the timing of video releases, and the simpler model more readily clarifies the central trade-offs. However, competition is certainly an important topic and can even be a factor for deciding when to schedule theatrical release, particularly during peak seasons. Notably, the effect of competition on *video* release times is much weaker (Goldberg 1991). Another aspect worth investigating is the release of multiple films by the same studio over time, i.e., the repeated game aspect. Along these lines, Prasad et al. (2004) specifically consider the impact of consumer expectations over time in an aggregate model. In that context, the authors commented that an interesting avenue of future research would be to start from a consumer utility model and capture the impact of consumer expectations that arise endogenously. Our paper may serve as a building block for further research in this direction.

Piracy is a pervasive problem in digital goods markets and is often facilitated by online systems that support illegal file sharing. Researchers who study the economics of digital piracy identify how consumers, when they pirate, often still face costs that are akin to price. For example, pirates are subject to government imposed penalties, search and learning costs, and moral costs (Connor and Rumelt 1991, August and Tunca 2008, Danaher et al. 2010, Lahiri and Dey 2013). Because would-be pirates are an important source of revenue to studios, the more general pricing problem they face must also account for the incentive compatibility constraints of consumers considering piracy. An important direction for future research is to study how movie piracy would affect the studio's setting of video release time and price controls. In particular, formal analysis of the studio's behavior as its pricing problem becomes more constrained as a consequence of combatting piracy may lead to important insights (Pogue 2012).

In this paper, we study the video release time and pricing decisions based on the studio's perspective, i.e., one decision maker. Calzada and Valletti (2012) study a decentralized channel in which an exhibitor sets the movie price and the retailer sets the video price, and show that if the studio's bargaining power relative to the exhibitor is high and content durability is relatively low, a delayed release is an optimal strategy for the studio. Furthermore, they demonstrate that if the studio's bargaining power is sufficiently higher than the exhibitor, the day-and-date strategy becomes optimal. In our model, if the channel becomes decentralized, the studio loses direct control over the video price and must then resort more toward utilizing the video release time to improve its revenue by inducing appropriate consumer market structures. This loss of direct video price control can lead the studio to delay the video release and sacrifice a loss in video quality so that it can ensure lucrative movie revenues. Consequently, in our setting, we expect that channel decentralization would result in greater use of a delayed released strategy for videos.

We consider a fixed simple revenue sharing contract between studios and exhibitors. In practice, conventional contracts involve time-dependent revenue sharing (i.e., a sliding, increasing percentage of revenues to exhibitors) after allowance for exhibitors' expenses. But recently, some studios and exhibitors have begun to implement "aggregate settlement," a simple revenue split without sliding scales (Vogel 2007, p. 148). Our model can serve as a reasonable approximation for that scenario. We have made simplifying assumptions to focus primarily on the underlying trade-offs relevant to our research questions. Nevertheless, our results are robust and satisfied for wide ranges of revenue shares. One fruitful extension would be to rigorously analyze how the various contracts between studios, exhibitors, and video retailers should be designed, particularly in light of our results on the profitability of day-and-date release strategies.

We assume that movie quality is certain and common knowledge, noting that empirical evidence concerning demand uncertainty associated with quality is not as strong as popularly argued (see, e.g., Orbach 2004, Orbach and Einav 2007). Furthermore, most of the uncertainty is revealed after the first weekend of release. Consequently, when distributors set video release dates after uncertainties are almost resolved and consumers are well informed, most of our results are preserved. If consumers are not informed about the quality of films, then video windows determined by distributors may signal quality, a scenario that may merit study. In this direction, our paper provides an interesting observation: contrary to conventional wisdom, longer video release windows do not necessarily signal higher movie quality.

Finally, in our study, we assume that the studio announces the video release time and price immediately and commits to both of them, which seems credible given the repeated nature of the context. Commitment is a nonissue when either day-and-date or direct-to-video strategies are optimal since the video channel is opened immediately. However, a studio's inability to announce immediately can affect consumption early on during the run of a film in theaters. Extending the model to permit delayed announcements may be worth studying, especially for cases where video release times are determined to optimally be less than a month out.

This paper is a first step toward analyzing the trade-offs faced by studios as they determine an appropriate video release time and pricing strategy. Studios are demonstrably interested in the prospects of earlier release times and even day-and-date strategies, and can benefit from a better understanding of how congestion, film content durability, and movie quality interact with such strategies. Given the speed of technological advances and the enduring use of a channeling system, the film industry now has a significant opportunity to design products and make decisions on delivery systems that cater more effectively to consumer preferences. By establishing how effective day-and-date strategies are at boosting studio profits, we hope that our work helps to initiate part of this progress.

Supplemental Material

Supplemental material to this paper is available at http://dx.doi.org/10.1287/mksc.2015.0936.

Acknowledgments

The authors thank the editor-in-chief, the associate editor, and the anonymous reviewers for their helpful suggestions throughout the review process. The authors also thank the participants at the INFORMS 2013 Annual Meeting, the Theory in Economics of Information Systems 2013 Workshop, the Production and Operations Management Society 2014 Annual Conference, as well as the participants in the research seminars at the University of Calgary, Stanford University, Santa Clara University, University of Washington, University of California, Los Angeles, Zaragoza Logistics Center, Georgia Institute of Technology, University of Southern California, and McGill University for their helpful comments and discussions.

Appendix A. Proofs of Propositions

For simplicity and greater clarity of argument, we focus on limiting values: $\alpha_L = 0$, $\delta_L = 0$, $\delta_H = 1$, and $\alpha_H \to \infty$ in the proofs in the main appendix. We provide extended proofs for more general parameter values in the online supplement.

Proof of Proposition 1. For part (i), for class H, a high congestion cost factor α_H induces the consumer market structure [H: D-N], i.e.,

$$\sigma^*(v,H)=\begin{cases} D, & \text{if } \dfrac{p_d}{\gamma_d(1-T)} \le v < 1;\\ N, & \text{if } v < \dfrac{p_d}{\gamma_d(1-T)}. \end{cases} \tag{A1}$$

For $c = L$, under high content durability δ_H, high theatrical movie quality γ_m^H, and a high congestion cost factor α_H, we show that the optimal video price and the release time induce the consumer market structure [L: B-M-N]; specifically

$$\sigma^*(v,L)=\begin{cases} B, & \text{if } \dfrac{p_d}{\gamma_d(1-\xi)} \le v < 1;\\ M, & \text{if } \dfrac{p_m}{\gamma_m^H} \le v < \dfrac{p_d}{\gamma_d(1-\xi)};\\ N, & \text{if } v < \dfrac{p_m}{\gamma_m^H}. \end{cases} \tag{A2}$$

To demonstrate that these equilibrium consumer market structures arise under the studio's optimal strategy requires careful examination of numerous cases and extensive algebraic comparisons. We provide these complete and rigorous arguments in an online supplement, while focusing on the characterization of the release time and price in the main appendix.[2]

Given that the studio prefers to induce this equilibrium, its corresponding profit function is written as

$$\begin{aligned}\Pi(T,p_d) = {}& \lambda_m p_m \rho\left(1-\frac{p_m}{\gamma_m^H}\right)+\lambda_d p_d\left(\rho\left(1-\frac{p_d}{\gamma_d(1-\xi)}\right)\right.\\ &\left.+(1-\rho)\left(1-\frac{p_d}{\gamma_d(1-T)}\right)\right).\end{aligned} \tag{A3}$$

Taking the derivative of $\Pi(T,p_d)$ in (A3) with respect to p_d, we then obtain

$$\frac{\partial\Pi}{\partial p_d}=\frac{\lambda_d(1-\xi)(\gamma_d(1-T)-2p_d)-2\rho\lambda_d p_d(1-T-(1-\xi))}{\gamma_d(1-T)(1-\xi)}. \tag{A4}$$

[2] For each of the propositions, we have separated the proof in this manner between the appendix and the online supplement.

Furthermore, the second-order condition becomes

$$\frac{\partial^2\Pi}{\partial p_d^2}=-\frac{2\lambda_d}{\gamma_d}\left(\frac{1-\rho}{1-T}+\frac{\rho}{1-\xi}\right)<0, \tag{A5}$$

which is satisfied, and hence the first-order condition is sufficient. Thus, $p_d^*(T)$ is an interior solution given T; specifically, given T, $p_d^*(T)=\gamma_d(1-T)(1-\xi)/(2((1-\xi)(1-\rho)+\rho(1-T)))$. Plugging this expression into (A3), we obtain

$$\begin{aligned}\Pi(T,p_d^*(T)) = {}& \frac{\rho\lambda_m p_m(\gamma_m^H-p_m)}{\gamma_m^H}\\ &+\frac{\lambda_d\gamma_d(1-T)(1-\xi)}{4((1-\rho)(1-\xi)+\rho(1-T))}.\end{aligned} \tag{A6}$$

Taking the derivative of this profit function with respect to T, we have

$$\frac{d\Pi(T,p_d^*(T))}{dT}=-\frac{\lambda_d\gamma_d(1-\xi)^2(1-\rho)}{4((1-\xi)(1-\rho)+\rho(1-T))^2}<0. \tag{A7}$$

Hence, $T^*=0<\xi$, and plugging this back into $p_d^*(T^*=0)$, we then obtain the optimal p_d^*.

Similarly, for part (ii), under a low congestion cost factor, one can prove that the optimal release time and the video price induce the consumer market structure of [L: B-M-N] and [H: B-M-N] (see the online supplement), and the corresponding profit can be written as

$$\Pi(T,p_d)=\frac{\lambda_m p_m(\gamma_m^H-p_m)}{\gamma_m^H}+\lambda_d p_d\left(1-\frac{p_d}{\gamma_d(1-\xi)}\right). \tag{A8}$$

Note that this profit function does not depend on T. Because video consumption occurs at $T=\xi$ for *both* consumers, profits are weakly maximized at $T^*=\xi$. Resolving indifference at $T^*=\xi$ maintains continuity in the optimal strategy as γ_m^H decreases and is also more likely to arise because of production lead times associated with videos. Therefore, in this case, $T^*=\xi$, and by maximizing the studio's profit function over p_d, we obtain $p_d^*=\gamma_d(1-\xi)/2$. Furthermore, if $T>\xi$, then the studio's profit is decreasing in T in this parameter region. Hence, under high content durability, $T^*=\xi$ and $p_d^*=\gamma_d(1-\xi)/2$. □

Proof of Proposition 2. For part (i), similar to the proof of Proposition 1, a high congestion cost factor α_H induces the consumer market structure [H: D-N]. Moreover, for $c=L$, under high content durability δ_H, intermediate theatrical movie quality γ_m^I, and a high congestion cost factor α_H, the optimal video price and the release time ($T\le\xi$) induce the consumer market structure [L: B-D-N]. In this case, the corresponding studio's profit function can be written as

$$\begin{aligned}&\Pi(T,p_d)\\ &=p_m\lambda_m\rho\left(1-\frac{p_m}{\gamma_m^I-\gamma_d(1-T-(1-\xi))}\right)\\ &\quad+p_d\lambda_d\left(\rho\left(1-\frac{p_d}{(1-T)\gamma_d}\right)+(1-\rho)\left(1-\frac{p_d}{(1-T)\gamma_d}\right)\right).\end{aligned} \tag{A9}$$

Differentiating (A9) with respect to p_d and solving a first-order condition, it follows that

$$p_d^*(T)=\frac{(1-T)\gamma_d}{2}. \tag{A10}$$

The second-order condition is satisfied in this case, which guarantees the optimality of (A10). Plugging (A10) into (A9), we then obtain

$$\Pi(T, p_d^*(T)) = \frac{\lambda_d \gamma_d (1-T)}{4} + \rho \lambda_m p_m \left(1 - \frac{p_m}{\gamma_m^I - \gamma_d(\xi - T)}\right). \quad \text{(A11)}$$

Differentiating $\Pi(T, p_d^*(T))$ with respect to T, we obtain

$$\frac{d\Pi(T, p_d^*(T))}{dT} = \gamma_d \left(-\frac{\lambda_d}{4} + \frac{\rho \lambda_m p_m^2}{(\gamma_m^I - \gamma_d(\xi - T))^2}\right). \quad \text{(A12)}$$

It follows that $d\Pi(T, p_d^*(T))/dT$ is decreasing in T, hence the second-order condition is satisfied. Furthermore, if $\lambda_d \Phi \le \lambda_m$, where $\Phi = ((\gamma_m^I - \gamma_d \xi)/(2\sqrt{\rho} p_m))^2$, the first-order condition is satisfied at $T^* = \xi - \gamma_m^I/\gamma_d + (2p_m/\gamma_d)\sqrt{\lambda_m \rho/\gamma_d}$ and by replacing this optimal T^*, we then obtain p_d^*. Otherwise, if $\lambda_m < \lambda_d \Phi$, then (A12) is strictly negative for all $T < \xi$ under δ_H. Therefore, the studio optimally sets $T^* = 0$ and $p_d^* = \gamma_d/2$, which results in day-and-date tactic.

For part (ii), under high content durability δ_H, intermediate theatrical movie quality γ_m^I, and a low congestion cost factor α_L, the optimal video price and release time induce the consumer market structure [*L*: *B-D-N*] and [*H*: *B-D-N*]. Consequently, the studio's resulting profit function is

$$\Pi(T, p_d) = \lambda_d p_d \left(1 - \frac{p_d}{\gamma_d(1-T)}\right) + \frac{\lambda_m p_m (\gamma_m^I - p_m - \gamma_d(\xi - T))}{\gamma_m^I - \gamma_d(\xi - T)}, \quad \text{(A13)}$$

for $T \le \xi$. From the first-order condition on p_d, we obtain $p_d^*(T) = \gamma_d(1 - T)/2$. By plugging this optimal expression into (A13), it follows that

$$\Pi(T, p_d^*(T)) = \frac{(\gamma_m^I - \gamma_d(\xi - T))(\lambda_d \gamma_d(1-T) + 4\lambda_m p_m) - 4\lambda_m p_m^2}{4(\gamma_m^I - \gamma_d(\xi - T))}, \quad \text{(A14)}$$

which is increasing in T for $T \le \xi$ under the bounds on γ_m^I (see Appendix B). Similarly, for $T > \xi$, the corresponding studio's profit function for the consumer market structure [*L*: *B-D-N*] and [*H*: *B-D-N*] is

$$\Pi(T, p_d) = \lambda_d p_d \left(1 - \frac{p_d}{\gamma_d(1-T)}\right) + \frac{\lambda_m p_m (\gamma_m^I - p_m)}{\gamma_m^I}. \quad \text{(A15)}$$

The optimal video price given T is the same as before, i.e., $p_d^*(T) = \gamma_d(1 - T)/2$. Replacing this optimal price in (A15), we obtain

$$\Pi(T, p_d^*(T)) = \frac{\gamma_m^I(\lambda_d \gamma_d(1-T) + 4\lambda_m p_m) - 4\lambda_m p_m^2}{4\gamma_m^I}, \quad \text{(A16)}$$

which is decreasing in T under bounds on γ_m^I. Therefore, the optimal release time is $T^* = \xi$, and the video price is $p_d^* = \gamma_d(1 - \xi)/2$. □

Proof of Proposition 3. First, for part (i), for a film with low content durability δ_L and high movie quality γ_m^H, under a high congestion cost factor α_H, the studio sets the release time and the video price in such a way that the consumer market structure at the optimum will be [*L*: *M-N*] and [*H*: *D-N*] (see the online supplement for a detailed proof that other consumer market structures are dominated by this perfect segmentation market structure). In this structure, the studio's profit can be written as

$$\Pi(T, p_d) = \rho \lambda_m p_m \left(1 - \frac{p_m}{\gamma_m^H}\right) + (1-\rho)\lambda_d p_d \left(1 - \frac{p_d}{\gamma_d(1-T)}\right). \quad \text{(A17)}$$

By taking the derivative of this profit function with respect to p_d, it follows that

$$\frac{\partial \Pi(T, p_d)}{\partial p_d} = \frac{(1-\rho)\lambda_d(\gamma_d(1-T) - 2p_d)}{\gamma_d(1-T)}, \quad \text{(A18)}$$

from which we obtain $p_d^*(T) = \gamma_d(1 - T)/2$. By plugging $p_d^*(T)$ into (A17), we have

$$\Pi(T, p_d^*(T)) = \frac{(1-\rho)\lambda_d \gamma_m^H \gamma_d(1-T) + 4\rho \lambda_m p_m(\gamma_m^H - p_m)}{4\gamma_m^H}, \quad \text{(A19)}$$

which is decreasing in T. Hence, $T^* = 0$ and consequently, $p_d^* = \gamma_d/2$.

For part (ii), under a low congestion cost factor α_L and $\lambda_d > \lambda_m$, the consumer market structure at the optimal release time and video price is [*H*: *M-D-N*] and [*L*: *M-D-N*]. In this case, the studio's profit function is written as

$$\Pi(T, p_d) = \lambda_m p_m \frac{\gamma_m^H - p_m - (\gamma_d(1-T) - p_d)}{\gamma_m^H - \gamma_d(1-T)} + \lambda_d p_d \left(\frac{p_m - p_d}{\gamma_m^H - \gamma_d(1-T)} - \frac{p_d}{\gamma_d(1-T)}\right). \quad \text{(A20)}$$

From the first-order condition on p_d, we obtain

$$p_d^*(T) = \frac{\gamma_d p_m (1-T)(\lambda_d + \lambda_m)}{2\lambda_d \gamma_m^H}. \quad \text{(A21)}$$

Substituting the optimal video price $p_d^*(T)$ into the studio's profit function, we obtain $\Pi(T, p_d^*(T))$, which is decreasing in T; as a result, $T^* = 0$. Therefore, $p_d^* = \gamma_d p_m(\lambda_m + \lambda_d)/(2\lambda_d \gamma_m^H)$. If $\lambda_d \le \lambda_m$, the proof is very similar to that above. The difference is that in the corresponding parameter region, the condition $\lambda_m \ge \lambda_d$ leads to the equilibrium outcome of not releasing the video at all because of the negative impact of demand cannibalization on the studio's profits. □

Proof of Corollary 1. We prove that under γ_m^H and δ_L,

$$\Pi^*|_{\{\alpha = \alpha_L\}} < \Pi^*|_{\{\alpha = \alpha_H\}}, \quad \text{(A22)}$$

if either of the following conditions hold:

(i) $\lambda_d > \lambda_m \max(1, \bar{\lambda}_d)$ where $\bar{\lambda}_d$ is the unique positive root of λ_d that solves

$$\gamma_d(\gamma_m^H(\gamma_m^H - \gamma_d) - p_m^2)\lambda_d^2 - 2p_m(2(\gamma_m^H)^2 - 2(\gamma_d + p_m)\gamma_m^H + p_m \gamma_d)\lambda_d - \gamma_d p_m^2 = 0, \quad \text{(A23)}$$

or

(ii) $\lambda_m > \lambda_d > 4\lambda_m p_m/\gamma_d$.

First, suppose that $\lambda_d > \lambda_m$. Under α_H, δ_L, and γ_m^H, the outcome corresponds to the perfect segmentation tactic in part (i) of Proposition 3, and the corresponding studio's optimal profit is

$$\Pi^*|_{\{\alpha=\alpha_H\}} = \lambda_m p_m \rho\left(1 - \frac{p_m}{\gamma_m^H}\right) + \lambda_d(1-\rho)\frac{\gamma_d}{4}. \quad \text{(A24)}$$

On the other hand, under α_L, the outcome corresponds to the day-and-date tactic in part (ii) of Proposition 3 under the condition of $\lambda_d > \lambda_m$. The corresponding optimal profit for the studio can be written as

$$\Pi^*|_{\{\alpha=\alpha_L\}} = p_m\left(\lambda_m - p_m\frac{4\gamma_m^H\lambda_m\lambda_d - \gamma_d(\lambda_m+\lambda_d)^2}{4\gamma_m^H\lambda_d(\gamma_m^H - \gamma_d)}\right), \quad \text{(A25)}$$

for $\alpha_L = 0$. By comparing the optimal studio's profits in (A24) and (A25), we obtain that if $\lambda_d > \lambda_m\bar{\lambda}_d$, the studio's profit under perfect segmentation in (A24) is higher than the profit under the day-and-date tactic in (A25). Moreover, there exists a unique, positive solution of λ_d in (A23) for γ_m^H sufficiently high.

Next, suppose that $\lambda_d < \lambda_m$. In this case, under α_H, it is the same as the previous, i.e., it corresponds to perfect segmentation and the studio's profit is given in (A24). The difference is that if $\lambda_d < \lambda_m$, then under α_L, it corresponds to a movie only release strategy in part (ii) of Proposition 3. In this region where a movie-only structure is optimal, the studio's corresponding profits become

$$\Pi^*|_{\{\alpha=\alpha_L\}} = \lambda_m p_m(1 - p_m/\gamma_m^H). \quad \text{(A26)}$$

By comparing optimal profits in those two cases, we obtain that if $\lambda_d > 4\lambda_m p_m \gamma_d$, the studio's profit under perfect segmentation given in (A24) is higher than the profit under a movie-only release presented in (A26). As a result, an increase in the congestion cost parameter α can increase studio profits, which completes the proof.

Proof of Proposition 4. For part (i), a high congestion cost factor α_H yields [H: D-N] in equilibrium. Furthermore, under δ_L and γ_m^I, the consumer market structure for class L at optimality is [L: M-D-N]. Under this consumer market structure, the studio's profit can be written as

$$\begin{aligned}\Pi(T, p_d) = {} & \lambda_d p_d\left(\rho\left(\frac{p_m - p_d}{\gamma_m^I - (1-T)\gamma_d} - \frac{p_d}{(1-T)\gamma_d}\right)\right. \\ & \left. + (1-\rho)\left(1 - \frac{p_d}{(1-T)\gamma_d}\right)\right) \\ & + \rho\lambda_m p_m\left(1 - \frac{p_m - p_d}{\gamma_m^I - (1-T)\gamma_d}\right). \quad \text{(A27)}\end{aligned}$$

Optimizing its profit over p_d, we obtain

$$p_d^*(T) = \frac{(1-T)\gamma_d(\lambda_d(1-\rho)(\gamma_m^I - (1-T)\gamma_d) + \rho p_m(\lambda_m + \lambda_d))}{2\lambda_d(\gamma_m^I - (1-T)(1-\rho)\gamma_d)}. \quad \text{(A28)}$$

Plugging (A28) into (A27), and taking a derivative of $\Pi(T, p_d^*(T))$ with respect to T, we find that it is decreasing in T. Thus, it follows that $T^* = 0$. Replacing $T^* = 0$ into (A28), we obtain (11). The proof of part (ii) directly follows from the proof of part (ii) in Proposition 3. □

Appendix B. Characterization of Bounds

In this section, we provide detailed expressions for the bounds of γ_m^I and characterize the parameter regions that we focus on.

First, $\gamma_m^I \in [\underline{\gamma}_m, \bar{\gamma}_m]$, where $\underline{\gamma}_m = \max(\underline{\gamma}_1, \underline{\gamma}_2, \ldots, \underline{\gamma}_6)$, and $\bar{\gamma}_m = \min(\bar{\gamma}_1, \bar{\gamma}_2, \ldots, \bar{\gamma}_7)$, in which

$$\underline{\gamma}_1 = \frac{p_m}{1-\xi},$$

$$\underline{\gamma}_2 = \frac{1}{2}\left(p_m + \gamma_d + \sqrt{p_m^2 + \gamma_d^2}\right),$$

$$\begin{aligned}\underline{\gamma}_3 = {} & \frac{1}{2\lambda_d(3-2\xi)} \\ & \cdot\Big(\gamma_d\lambda_d(-2+\xi)(-3+2\xi) + p_m\big(\lambda_d(7-6\xi) + \lambda_m(-3+2\xi)\big) \\ & - \big[(p_m^2(7\lambda_d - 3\lambda_m - 6\lambda_d\xi + 2\lambda_m\xi)^2 + 2p_m\gamma_d\lambda_d(1-\xi)(3-2\xi) \\ & \times(-\lambda_d - 3\lambda_m + 2(\lambda_d+\lambda_m)\xi) + \gamma_d^2\lambda_d^2(3 - 5\xi + 2\xi^2)^2)\big]^{1/2}\Big),\end{aligned}$$

$$\begin{aligned}\underline{\gamma}_4 = {} & \frac{1}{2\lambda_d}\Big(3p_m\lambda_d + \gamma_d\lambda_d - p_m\lambda_m - \gamma_d\lambda_d\xi \\ & - \sqrt{(p_m(-3\lambda_d + \lambda_m) + \gamma_d\lambda_d(-1+\xi))^2 + 8p_m\gamma_d\lambda_d^2(-1+\xi)}\Big),\end{aligned}$$

$$\underline{\gamma}_5 = \frac{1}{4}\left((3 + 4(1-\xi))\gamma_d + 2p_m\left(1 - \frac{\lambda_m}{\lambda_d}\right)\right),$$

$$\underline{\gamma}_6 = 2p_m\frac{\lambda_m}{\lambda_m + \lambda_d},$$

$$\bar{\gamma}_1 = \gamma_d(1 - \delta(1-\xi)) + 2p_m\delta(1-\xi),$$

$$\bar{\gamma}_2 = 2p_m\sqrt{\frac{\lambda_m\rho}{\lambda_d}} - \gamma_d\delta(1-\xi)\left(1 - \sqrt{\frac{\lambda_d}{\lambda_m\rho}}\right),$$

$$\bar{\gamma}_3 = 2p_m\left(1 + \frac{\xi\rho}{1-\xi}\right),$$

$$\bar{\gamma}_4 = 2\sqrt{2p_m\gamma_d(1-\xi)} - \gamma_d(1-\xi),$$

$$\bar{\gamma}_5 = 2p_m\lambda_m,$$

$$\begin{aligned}\bar{\gamma}_6 = {} & \frac{1}{2\lambda_d(3-2\xi)} \\ & \cdot\Big(\gamma_d\lambda_d(-2+\xi)(-3+2\xi) + p_m\big(\lambda_d(7-6\xi) + \lambda_m(-3+2\xi)\big) \\ & + \big[(p_m^2(7\lambda_d - 3\lambda_m - 6\lambda_d\xi + 2\lambda_m\xi)^2 + 2p_m\gamma_d\lambda_d(1-\xi)(3-2\xi) \\ & \times(-\lambda_d - 3\lambda_m + 2(\lambda_d+\lambda_m)\xi) + \gamma_d^2\lambda_d^2(3 - 5\xi + 2\xi^2)^2)\big]^{1/2}\Big),\end{aligned}$$

$$\begin{aligned}\bar{\gamma}_7 = {} & \frac{1}{2\lambda_d}\Big[3p_m\lambda_d + \gamma_d\lambda_d - p_m\lambda_m - \gamma_d\lambda_d\xi \\ & + \sqrt{(p_m(-3\lambda_d + \lambda_m) + \gamma_d\lambda_d(-1+\xi))^2 + 8p_m\gamma_d\lambda_d^2(-1+\xi)}\Big],\end{aligned}$$

and

$$\bar{\gamma}_8 = p_m + \gamma_d(1-\delta)(1-\xi).$$

Second, we also have $\rho < \bar{\rho} = \min(\lambda_d(\gamma_m - \gamma_d(1-\delta)(1-\xi))^2/(4p_m^2\lambda_m), (1-\xi)/2)$. Also, we impose the following set of restrictions on the parameter region, which guarantees nonemptiness of the interval $[\underline{\gamma}_m, \bar{\gamma}_m]$ and helps analytical tractability: $p_m > \frac{3}{4}\gamma_d$, $\xi<1/2$, $\lambda_m + \lambda_d > 1$,

$$\gamma_d\lambda_d(1-\lambda_m)(1-\rho) + p_m\lambda_m\big(\lambda_m\rho + \lambda_d(-1-(1-\rho)(1-2\lambda_m))\big) < 0,$$

$$\big(p_m(-3\lambda_d+\lambda_m)+\gamma_d\lambda_d(-1+\xi)\big)^2 + 8p_m\gamma_d\lambda_d^2(-1+\xi) > 0,$$

$$\big(1+(1-\xi)(4+5(1-\xi))\big)\gamma_d^2\lambda_d^2 + p_m^2(\lambda_d-\lambda_m)^2 - 2p_m(1-\xi)\gamma_d\lambda_d(3\lambda_d+\lambda_m) < 0,$$

and

$$2p_m\gamma_d\lambda_d(-1+\xi)(-3+2\xi)\big(-\lambda_d-3\lambda_m+2(\lambda_d+\lambda_m)\xi\big) + \gamma_d^2\lambda_d^2(3-5\xi+2\xi^2)^2 + p_m^2\big(\lambda_m(3-2\xi)+\lambda_d(-7+6\xi)\big)^2 > 0.$$

The derivation of these bounds is shown in the online supplement.

References

Abate MB (2012) Dinner at the movies: Cinepolis Del Mar luxury cinema. *San Diego Reader*. www.sandiegoreader.com/weblogs/feast/2012/jan/03/x/.

Argo JJ, Dahl DW, Manchanda RV (2005) The influence of a mere social presence in a retail context. *J. Consumer Res.* 32(2): 207–212.

August T, Tunca TI (2006) Network software security and user incentives. *Management Sci.* 52(11):1703–1720.

August T, Tunca TI (2008) Let the pirates patch? An economic analysis of software security patch restrictions. *Inform. Systems Res.* 19(1):48–70.

August T, Tunca TI (2011) Who should be responsible for software security? A comparative analysis of liability policies in network environments. *Management Sci.* 57(5):934–959.

Barnes B (2008) Direct-to-DVD releases shed their loser label. *New York Times* (January 28), www.nytimes.com/2008/01/28/business/media/28dvd.html.

Besanko D, Winston WL (1990) Optimal price skimming by a monopolist facing rational consumers. *Management Sci.* 36(5): 555–567.

Bonnington C (2012) Why 3D TV went from CES darling to consumer reject. *Wired* (January 5), www.wired.com/2012/01/state-of-3-d-technology/.

Bulow JI (1982) Durable-goods monopolist. *J. Political Econom.* 90(2):314–332.

Calzada J, Valletti TM (2012) Intertemporal movie distribution: Versioning when customers can buy both versions. *Marketing Sci.* 31(4):649–667.

Cieply M (2008) Many movie theaters decide to leave the bat signal on till dawn. *New York Times* (July 9), www.nytimes.com/2008/07/09/movies/09dark.html.

Coase RH (1972) Durability and monopoly. *J. Law Econom.* 15(1): 143–149.

Cole G (2007) Release windows: Now you can have it on DVD in no time at all. *One to One*.

Connor KR, Rumelt RP (1991) Software piracy: An analysis of protection strategies. *Management Sci.* 37(2):125–139.

Danaher B, Dhanasobhon S, Smith MD, Telang R (2010) Converting pirates without cannibalizing purchasers: The impact of digital distribution on physical sales and Internet piracy. *Marketing Sci.* 29(6):1138–1151.

Desai P, Purohit D (1998) Leasing and selling: Optimal marketing strategies for a durable goods firm. *Management Sci.* 44(11): S19–S34.

Desai P, Koenigsberg O, Purohit D (2004) Strategic decentralization and channel coordination. *Quant. Marketing Econom.* 2(1):5–22.

Donaldson-Evans C (2006) Doomsayers predict death of movie theater. *FoxNews.com* (January 16), http://www.foxnews.com/story/2006/01/16/doomsayers-predict-death-movie-theater.html.

Einav L (2007) Seasonality in the U.S. motion picture industry. *RAND J. Econom.* 38(1):127–145.

Eliashberg J, Elberse A, Leenders MA (2006) The motion picture industry: Critical issues in practice, current research, and new research directions. *Marketing Sci.* 25(6):638–661.

Epstein EJ (2005) Hollywood's death spiral. *Slate.com* (July 25), http://www.slate.com/articles/arts/the_hollywood_economist/2005/07/hollywoods_death_spiral.html.

Frank B (1994) Optimal timing of movie releases in ancillary markets: The case of video rentals. *J. Cultural Econom.* 18(2): 125–133.

Gerchak Y, Cho RK, Ray S (2006) Coordination of quantity and shelf-retention timing in the video movie rental industry. *IIE Trans.* 38(7):525–536.

Goldberg F (1991) *Motion Picture Marketing and Distribution* (Butterworth-Heinemann, Stoneham, MA).

Gross D (2006) The broken windows theory. *Slate.com* (January 18), http://www.slate.com/articles/business/moneybox/2006/01/the_broke_windows_theory.html.

Grover R (2005) What's driving the box office batty? *BusinessWeek* (July 10), www.bloomberg.com/bw/stories/2005-07-10/whats-driving-the-box-office-batty.

Grover R (2006) Will *Bubble* burst a Hollywood dogma? *BusinessWeek* (January 24), https://web.archive.org/web/20060806150954/http://www.businessweek.com/bwdaily/dnflash/jan2006/nf20060124_4959_db011.htm.

Gul F, Sonnenschein H, Wilson R (1986) Foundations of dynamic monopoly and the Coase conjecture. *J. Econom. Theory* 39(1): 155–190.

Harrell GD, Hutt MD, Anderson JC (1980) Path analysis of buyer behavior under conditions of crowding. *J. Marketing Res.* 17(1): 45–51.

Hennig-Thurau T, Henning V, Sattler H, Eggers F, Houston MB (2007) The last picture show? Timing and order of movie distribution channels. *J. Marketing* 71(4):63–83.

Hui SK, Bradlow ET, Fader PS (2009) Testing behavioral hypotheses using an integrated model of grocery store shopping path and purchase behavior. *J. Consumer Research* 36(3):478–493.

Jurgensen J (2009) Saving Jackson's last steps. *Wall Street Journal* (October 23), www.wsj.com/articles/SB10001424052748704224004574487742094843078.

Kirsner S (2005) Maverick mogul. *FastCompany.com* (December 1), www.fastcompany.com/54534/maverick-mogul.

Laffont J-J, Tirole J (1988) The dynamics of incentive contracts. *Econometrica* 56(5):1153–1175.

Lahiri A, Dey D (2013) Effects of piracy on quality of information goods. *Management Sci.* 59(1):245–264.

Lehmann DR, Weinberg CB (2000) Sales through sequential distribution channels: An application to movies and videos. *J. Marketing* 64(3):18–33.

Li L, McKelvey RD, Page T (1987) Optimal research for Cournot oligopolists. *J. Econom. Theory* 42(1):140–166.

Luan JY, Sudhir K (2006) Optimal inter-release time between sequentially released products. Working paper, Dartmouth University, Hanover, NH, and Yale University, New Haven, CT.

Luna N (2012) Luxury theater with "on call waiters" opens Friday. *Orange County Register* (May 30), www.ocregister.com/articles/theater-592968-cinepolis-movie.html.

Marr M (2005) Redirecting Disney. *Wall Street Journal* (December 5), www.wsj.com/articles/SB113374725252213739.

Miller D (2012) Sundance 2012: The day-and-date success story of "Margin Call." *Hollywood Reporter* (January 18), www.hollywoodreporter.com/news/sundance-2012-margin-call-video-on-demand-zach-quinto-283033.

Miller PD (2006) *Applied Asymptotic Analysis*, Graduate Studies in Mathematics, Vol. 75 (American Mathematical Society, Providence, RI).

Moorthy KS, Png IPL (1992) Market segmentation, cannibalization, and the timing of product introductions. *Management Sci.* 38(3):345–359.

Moul CC, Shugan SM (2005) Theatrical release and the launching of motion pictures. Moul CC, ed. *A Concise Handbook of Movie Industry Economics* (Cambridge University Press, New York), 80–137.

Muller HM (2000) Asymptotic efficiency in dynamic principal-agent problems. *J. Econom. Theory* 91(2):292–301.

NATO (2013) Number of U.S. movie screens. *NATOOnline.org* (January), http://www.natoonline.org/data/us-movie-screens/.

Nelson RA, Reid CE, Gilmore O (2007) An analysis of the out-of-market gap for DVDs in the U.S. *Rev. Indust. Organ.* 30(4): 303–323.

Orbach BY (2004) Antitrust and pricing in the motion picture industry. *Yale J. Regulation* 21(2):317–367.

Orbach BY, Einav L (2007) Uniform prices for differentiated goods: The case of the movie-theater industry. *Internat. Rev. Law Econom.* 27(2):129–153.

Pei PP-E, Simchi-Levi D, Tunca TI (2011) Sourcing flexibility, spot trading, and procurement contract structure. *Oper. Res.* 59(3): 578–601.

Pogue D (2012) Hollywood killed the video star. *Scientific Amer.* (August 12), www.nature.com/scientificamerican/journal/v307/n3/full/scientificamerican0912-36.html.

Prasad A, Bronnenberg B, Mahajan V (2004) Product entry timing in dual distribution channels: The case of the movie industry. *Rev. Marketing Sci.* 2(1):1–18.

Shaked A, Sutton J (1983) Natural oligopolies. *Econometrica* 51(5): 1469–1483.

Shaked A, Sutton J (1987) Product differentiation and industrial structure. *J. Indust. Econom.* 36(2):131–146.

Smith E, Schuker LA (2010) Studios unlock DVD release dates. *Wall Street Journal* (February 12), http://www.wsj.com/articles/SB10001424052748704337004575059713216224640.

Su X (2007) Intertemporal pricing with strategic customer behavior. *Management Sci.* 53(5):726–741.

Tartaglione N (2013) European union backs pan-European day-and-date distribution trials. *Deadline.com* (February 28), http://deadline.com/2013/02/european-union-backs-pan-european-day-and-date-distribution-trials-441271/.

Thompson A (2006) Distributors hold firm against day-and-date. *Hollywood Reporter* (March 17), https://web.archive.org/web/20060518010224/http://www.hollywoodreporter.com/thr/columns/risky_business_display.jsp?vnu_content_id=1002198452.

Tribbey R (2013) DVD and Blu-ray release report. *DVDRelease Report.com*.

Tunca TI, Zenios SA (2006) Supply auctions and relational contracts for procurement. *Manufacturing Service Oper. Managment* 8(1):43–67.

Vickrey WS (1955) A proposal for revising New York's subway fare structure. *Oper. Res.* 3(1):38–68.

Vogel HL (2007) *Entertainment Industry Economics*, 7th ed. (Cambridge University Press, New York).

Waterman D, Weiss AA (2010) Time consistency and seller commitment in intertemporal movie distribution: An empirical study of the video window. *J. Indust. Econom.* 58(3):717.

Chapter 1, in full, is a reprint of the material as it appears in "Optimal Timing of Sequential Distribution: The Impact of Congestion Externalities and Day-and-Date Strategies," 2015, *Marketing Science*, **34**(5), 755-774. Terrence August, Duy Dao, and Hyoduk Shin. The dissertation author was the primary researcher and author of this material.

Chapter 2

Market Segmentation and Software Security: Pricing Patching Rights

2.1 Introduction

Attacks on unpatched software continue to be a major problem. The Department of Homeland Security's United States Computer Emergency Readiness Team indicates that systems running unpatched versions of software from providers such as Microsoft, Adobe, and OpenSSL are consistently attacked through patchable vulnerabilities (US-CERT 2015). The NotPetya and WannaCry ransomware attacks are recent examples of large-scale attacks that exploited such vulnerabilities. Microsoft had released a patch on March 14, 2017, following revelation of the vulnerability's existence by The Shadow Brokers hacker group (Microsoft 2017b). Two months later, despite the patch having been made available, the WannaCry ransomware attack struck over 200,000 computers across over 150 countries (Lohr and Alderman 2017, Greenberg 2017). Even one month after WannaCry was on the news around the world, many users and organizations had still not patched and, as a result, the NotPetya malware was able to spread using the same exploit (Microsoft 2017a). In addition to being exposed to ransomware threats, unpatched systems can be leveraged in a variety of other criminal activities, including those associated with spam (Levchenko et al. 2011) and distributed denial-of-service services (Fitzgerald 2015).

Observing today's cybersecurity attack landscape, the current patching process for security has been less effective than desired (August et al. 2014). Many systems remain unpatched long after patches are released. HP indicates in its recent *Cyber Risk Report 2015*, that "... the majority of exploits discovered by our teams attempt to exploit older vulnerabilities. By far the most common exploit is CVE-2010-2568, which roughly accounts for a third of all discovered exploit samples" (HP Security Research 2015). CVE-2010-2568, a Windows shell vulnerability that allows for remote code execution, was discovered in June of 2010. The patch for this vulnerability was released weeks later in August of 2010. Despite the patch being available for over six years, this vulnerability is *still* being exploited by attackers. According to HP, 64% of the top exploit samples in 2014 targeted vulnerabilities from 2012 and prior. OPSWAT, which collects data from software users through its security platform, finds that less than 29% of Windows operating systems in their data are up to date (OPSWAT 2014). Similarly, the Canadian Cyber Incident Response Centre states that patching applications and patching operating system vulnerabilities (in addition to whitelisting and access control strategies) would prevent as much as 85% of targeted security attacks (CCIRC 2014).

Why aren't users patching? The growing reality is that security is not a technical problem, it's an economic one. Even though security patches are available, many users are not deploying them because it is not in their economic best interest to do so. For organizations, enterprise deployment of patches is a costly process. Extensive testing of patches in development and staging environments, roll-out of updates onto production servers, and final testing is both time consuming and resource intensive. Moreover, in aggregate there is a deluge of patches that system administrators must continuously monitor and process. For end users, the situation is regrettably similar because security patching is often not considered to be a priority. The deployment of updates and system rebooting is instead viewed as an inconvenience, particularly when users feel their own productivity is of greater concern. Taking into consideration these costs, we offer an approach on how a software vendor can substantially increase security in an incentive-compatible way by encouraging improved user behavior.

In particular, we argue that a vendor should differentiate its software product by pricing *patching rights*. Specifically, the vendor should charge users for the right to choose for themselves whether patches are installed or not installed on their systems. The status quo is that all users are endowed with patching rights, and a substantial portion of them elect not to do so as a result. By charging for patching rights, users who would otherwise have elected not to patch under the status quo must now examine whether it is worth paying for this right to remain unpatched. This decision is non-trivial as the expected security losses one would incur when retaining rights and remaining unpatched depends on the security behaviors of all other users in aggregate. On the flip side, by foregoing patching rights, users will have their systems automatically (and immediately[1]) patched by the vendor, and

[1]In our study, we focus on automatic updates provided by vendors who intend for them to be applied either immediately or within a reasonably short time frame so as to reduce security

pay a different price. However, automatic deployment of patches also comes with risk because patches are not always stable. For example, some users encountered the "Blue Screen of Death" when applying a patch from Microsoft (Westervelt 2014). More recently, the cryptocurrency exchange QuadrigaCX lost $14 million when inadequate patch testing failed to reveal that their process of transferring ether (a cryptocurrency) was incompatible with their Ethereum client software update (Higgins 2017). Considering these trade-offs, we will establish that vendors have an incentive to target certain users with lower prices in equilibrium in exchange for their patching rights, hence these users will ultimately cause less of a security externality while benefitting from discounted software prices.

We study the impact of optimally priced patching rights (PPR) for a software product on security, vendor profitability, and the overall value of the product to the economy. While software vendors tend to be painted as not caring about security, realistically their strategic motivations are much more complex. In fact, riskier software is not just harmful to the users who bear losses. The presence of a large unpatched user population creates disincentives for software usage which, in turn, certainly hurts profitability for the vendor. We construct a model of security where users can choose whether to purchase a software product and additionally whether to remain patched, unpatched, or have their systems automatically updated in a timely manner by a software vendor. By characterizing equilibrium consumer behavior in this setting, we can explore the potential benefits associated with strategically differentiating on patching rights.

We aim to make a contribution in an area where many researchers are working toward improving security and understanding why patch availability alone has not led to very secure outcomes. Importantly, the insights generated with our model have promising practical implications to the software industry. Historically software companies have not differentiated on patching rights, but recent activities by firms suggest a greater willingness to do so. These observations may impel

risk. This type of automatic update is consistent with those provided by many leading software vendors (e.g., Microsoft, Oracle, Adobe, Google, etc.). There are examples where firms manage automatic updates but do so in an delayed manner, such as was the case with Apple taking three months to patch the vulnerability exploited by the Flashback trojan (Hick 2012). This behavior is not in the spirit of the automated patching being modeled here.

vendors to discover that a more profitable, more secure ecosystem can be achieved as a result. In the discussion that follows our main results, we address how the essence of a PPR policy can be implemented as a component of a software vendor's overall versioning strategy.

2.2 Literature Review

This work is related to three broad areas in the literature: (i) product differentiation and market segmentation, (ii) economics of information security, and (iii) economics of product bundling. With regard to the first stream, to the best of our knowledge, our paper is the first to examine how beneficial segmentation in software markets can be constructed by differentiation on software patching rights. More specifically, we examine a monopolist's product mix/line decisions when the quality-differentiated dimension concerns patching rights, as opposed to competition-driven product differentiation. Within the second area, our paper is most closely related to a strand that studies the management of security patches. For the third area, our paper adds to a strand that examines the impact of mixed bundling on a monopolist's profit and social welfare. We make several contributions to these areas. Our work is based on the original idea that *patching rights* should be managed. Having been first introduced qualitatively in a perspectives piece (August et al. 2014), our paper is the first to formally model and analyze the value of patching rights and clarify the impact that a PPR policy has on security and the value of software products.

In the following, we will detail how our paper fits with each of these areas in the literature, beginning with the first. We study how a monopolist producer of software can expand its product mix by differentiating based on the patching rights. There is a well-developed literature in economics, marketing, operations management and information systems that examines the monopolist's problem of whether to offer and how to price quality-differentiated goods. In this vein, the classic papers focus on characterizing the general optimal non-linear price schedule that incentivizes each consumer to select the quality that is designed for her specific type (Mussa and Rosen 1978, Maskin and Riley 1984). Subsequent literature explored how these schedules react to various changes in consumer characteristics such as income dispersion and taste preferences (Moorthy 1984, Gabszewicz et al. 1986, Desai 2001, Villas-Boas 2009), firm characteristics such as production technology and marketing costs (Villas-Boas 2004, Debo et al. 2005, Netessine and Taylor 2007), and non-intrinsic characteristics such as strategic delays (Moorthy and Png 1992, Chen 2001).

Within this body of work, our paper is most closely related to those that focus on consumer characteristics in that an automated patching version essentially modifies the costs incurred by consumers. There are two critical contributions that our model adds to this portion of the literature. First, the qualities that comprise the product mix in our setting are endogenously determined in equilibrium by consumption decisions; this is a significant context-specific trait that we capture where users who do not patch their systems weakly reduce the quality of the software product for other users. Second, and an important focus of our work is that the price schedule is not based on the product that gets consumed but rather the *rights* retained by the consumer. Thus, our model admits interesting possibilities where consumers opt for the same rights (giving access to the same quality level) but then separate based on their own subsequent patching decision (resulting in different effective qualities) – effective quality of the product is inclusive of security attacks which are endogenously determined by the strategic protection behaviors employed in equilibrium.

With regard to the second area of literature, our paper is close to work that studies the management of security patches. Within that area, researchers examine the timing of security patch release and its application (Beattie et al. 2002, Cavusoglu et al. 2008, Dey et al. 2015), vulnerability disclosure policy (Cavusoglu et al. 2007, Arora et al. 2008, Ransbotham et al. 2012), vendor patch policy (Lahiri 2012 and Kannan et al. 2013), and users' patching incentives (August and Tunca 2006, Choi et al. 2010). Our work is closest to the latter group of papers which construct models that endogenize users' patching decisions. We build on this body of work and develop an originative model that includes an automated patching option for users within a game theoretical context accenting negative externalities stemming from unpatched behavior. Consistent with this work and models of vaccination (see, e.g., Brito et al. 1991), we capture these negative externalities and further generalize the model to include risk that is independent of patching populations. While interesting in their own right, our model does not include other attack-related effects such as hiding effects (Gupta and Zhdanov 2012), zero-day vulnerabilities (August and Tunca 2011), or strategic attackers (Png and Wang 2009, Kannan et al. 2013). However, our model is the first to include both standard and automated patching options for users while also modeling the security externality. The inclusion of automated patching is significant. First, it permits a characterization of the natural consumer market segmentation that arises in equilibrium as users strategically respond to security risk and expanded patching options. An understanding of equilibrium consumption and security behavior serves to inform how security enhancements should be marketed. Second, an automated patching option is the logical choice for the baseline product in a policy where patching rights are contracted, which is the focus of our work.

Toward addressing users' incentives, August and Tunca (2006) examine the efficacy of patching rebates. Patching rebates work by having the vendor subsidize patching costs in order to get more users to patch rather than remaining unpatched and contributing to security risk. August and Tunca (2006) show that these rebates can be very effective at improving behavior and security. They also show that if standard patching costs are large, it is not efficient to incentivize lower valuation users to incur these costs. On that dimension, one nice feature of the PPR policy we study in this work is it does not require lower valuation users to incur these patching costs. Rather, it incentivizes them to have patches automatically deployed and instead incur potential system instability losses due to automated deployment. Since these losses tend to be low, PPR works by incentivizing these users to engage in more appropriate and economical patching behaviors.

Finally, our work is related to a third body of literature examining a monopolist's decision regarding pure and mixed bundling. Under a PPR policy, software is unbundled with the right to decide whether or not to apply security patches.[2] There exists a well-developed literature on the bundling of physical products and information goods that spans the fields of economics, marketing, and information systems. We establish that under moderate costs associated with automated patching, our proposed partial mixed bundling scheme (PPR) can simultaneously improve the software vendor's profit as well as security relative to the pure bundling alternative status quo. Several related works share the similar qualitative conclusion in which mixed bundling is favored over pure bundling and unbundled sales (see, e.g., Adams and Yellen 1976, Schmalensee 1984, McAfee et al. 1989, Venkatesh and Kamakura 2003). These works show that bundling effectively extracts consumer surplus under various distributions of reservation values. In our work, we show that partial mixed bundling when involving patching rights can possibly result in a slight decrease in social welfare, but it can also drive increases in social welfare depending on the quality of automated patching solutions and to what extent security risk is reduced.

The marginal costs of individual goods and their asymmetric values and externalities strongly influence a monopolist's bundling decisions. In comparison to physical goods, information goods are typically assumed to have zero marginal costs, which enable the monopolist to bundle many information goods economically; this makes sense particularly when their valuations are correlated (Bakos and Brynjolfsson 1999). If two information goods provide highly asymmetric values to consumers, partial mixed bundling is optimal; the higher valuation good should not be sold separately not to cannibalize the sale of the bundle (Eckalbar 2010, Bhargava 2013). Idiosyncratic to our context, patching rights cannot be sold separately because they only have value to those who buy software. Unlike prior work on bundling, our model involves a security externality from unpatched software usage. Partial mixed bundling of two information goods has been shown to be optimal when only one good has a direct externality on consumer utility

[2] The status quo of providing software bundled with patching rights is, in this sense, pure bundling. PPR is partial mixed bundling: a bundle of software with patching rights and software alone without patching rights (Stremersch and Tellis 2002). Because patching rights have no standalone value, mixed bundling does not include the sale of patching rights.

(Prasad et al. 2010). However, in our model, as more consumers purchase the automated patching version, other consumers become willing to pay more for the bundled version with patching rights due to the increased level of security.

2.3 Model Description and Consumer Market Equilibrium

2.3.1 Model

There is a continuum of consumers whose valuations of a software product lie uniformly on $\mathcal{V} = [0, 1]$. The software is used in a network setting, thus exposing consumers to security risks associated with its use. In particular, a vulnerability can arise in the software in which case the vendor makes a security patch available to all users of the software. Because the security vulnerability can be used by malicious hackers to exploit systems, users who do not apply the security patch are at risk.

The vendor offers two options for users to protect their respective systems. In doing so, the vendor prices the software based on whether patching rights are granted to the consumer. Specifically, if a consumer elects to purchase the software and retain full patching rights, she pays the price $p \geq 0$. Having this right means she can choose whether to patch the software or not patch the product and do so according to her own preferences. If she decides to patch the software, she will incur an expected cost of patching denoted $c_p > 0$. This *standard patching* cost accounts for the money and effort that a consumer must exert in order to verify, test, and roll-out patched versions of existing systems.[3]

If she decides not to patch the software, then she faces the risk of an attack. We model two classes of security losses that can be incurred: (i) those that are

[3]Standard patching processes require considerable care, essentially coming down to labor costs associated with system administrators and developers spending time to complete all of the tasks in the patching process (Beres and Griffin 2009). Studies find that standard patching costs tend to be on the order of one thousand dollars per server (Bloor 2003, Forbath et al. 2005, Beres and Griffin 2009). Modeling the cost of standard patching as a constant is common in the literature that examines topics related to patching costs as can be seen in Beattie et al. (2002), August and Tunca (2006), Choi et al. (2010) and Cavusoglu et al. (2008).

dependent on the size of the unpatched population of users, and (ii) those that are independent of the unpatched population size. For the dependent case, if she decides not to patch the software, the probability she is hit by a security attack is given by $\pi_s u$, where $\pi_s > 0$ is the probability an attack appears on the network and u is the size of the unpatched population of users.[4] This reflects the negative security externality imposed by unpatched users of the software. If she is successfully attacked, she will incur expected security losses that are positively correlated with her valuation. That is, consumers with high valuations will suffer higher losses than consumers with lower valuations due to opportunity costs, higher criticality of data and loss of business. For simplicity, we assume that the correlation is of first order, i.e., the loss that a consumer with valuation v suffers if she is hit by an attack is $\alpha_s v$ where $\alpha_s > 0$ is a constant. We refer to $\pi_s \alpha_s$ as the *dependent risk* throughout the paper. The dependent case directly captures propagation type of attacks that spread through vulnerable populations - this is consistent with the case of the WannaCry ransomware attack described earlier (Lohr and Alderman 2017, Greenberg 2017). The dependent case also indirectly captures any type of security attack where the incentives of the malicious individual for constructing the attack is positively related to the unpatched population size. For example, if large vulnerable populations are more attractive to hackers because it becomes easier to penetrate hosts or the return on their efforts becomes higher when infecting more hosts, then the dependent case of our model will apply. On the other hand, unpatched users can also face risk from attacks which are simply independent of the size of the unpatched population. This class can include targeted attacks and other forms of background risk. In a similar fashion, we denote the likelihood of an independent attack with π_i, and refer to $\pi_i \alpha_i$ as the *independent risk*, where $\alpha_i > 0$ is a constant.

If the consumer instead elects to purchase the software and relinquish patching rights, she pays the price δp, where $\delta \geq 0$. In this case, the vendor retains full control over patching the software and will automatically and immediately do so

[4]The size of the unpatched population u is determined by the consumer strategies in equilibrium. Therefore, by the definition of $\mathcal{V}$, $u \in [0, 1]$.

to better protect the network.[5] From an implementation point of view, this software version would not give users much or any control over patch deployment (e.g., the typical options can be grayed out in this version). The user incurs a cost of *automated patching*, $c_a > 0$, which is associated with both inconvenience and configuration of the system to handle automatic deployment of security patches. Our model can examine any relationship between c_p and c_a. For example, it can capture the commonly observed situation in which users are choosing between: (i) completing all tasks associated with the rigorous, standard patching approach and incurring c_p, or (ii) doing the bare minimum tasks to deploy patches automatically without verification and incurring a lower cost, $c_a < c_p$, related to deployment. The model can also handle situations where $c_a \geq c_p$, to study scenarios in which users aim to achieve all the tasks associated with standard patching but in an automated manner. In general, a software vendor who releases a security patch cannot test for compatibility of the patch with every possible user system configuration. Thus, there is always some risk associated with an automatically deployed patch causing a user's system to become unstable or even crash. We denote the probability that the automated patch fails with $\pi_a > 0$. We assume that the loss associated with an automated patch deployment failure is again positively correlated with her valuation, and that this correlation is of first order, denoted as $\alpha_a > 0$. Thus, her expected loss associated with automated patching is given by $\pi_a \alpha_a v$. The loss factor $\pi_a \alpha_a$ captures in expectation major patch failures that would lead to severe backlash against the vendor. An increased likelihood of such events is represented by a higher $\pi_a \alpha_a$, which will affect the value of a PPR policy.

Each consumer makes a decision to buy, B, or not buy, NB. Similarly, the patching decision is denoted by one of patch, P, not patch, NP, and automatically patch, AP. In order to choose P or NP, the consumer must pay the premium price p to retain patching rights. By choosing AP, the consumer delegates patching

[5] In interconnected networks, this is fairly easy to enforce; for example, vendors such as Adobe and Matlab enforce real-time license checks for their subscription-based offerings. While it is always possible to circumvent protections, most paying customers are unlikely to break the license agreement. We assume immediate patch deployment for simplicity. The essence of our results only require that patchers, whether automated or standard, do so in a relatively timely manner that distinguishes them from those who do not patch.

rights to the vendor for a potentially discounted price δp, and again we focus on vendors who deploy those patches on users' systems in an expeditious manner. The consumer action space is then given by $S = (\{B\} \times \{P, NP, AP\}) \cup (NB, NP)$. In a consumer market equilibrium, each consumer maximizes her expected utility given the equilibrium strategies for all consumers. For a given strategy profile $\sigma : \mathcal{V} \to S$, the expected utility for consumer v is given by:

$$U(v,\sigma) \triangleq \begin{cases} v - p - c_p & if \quad \sigma(v) = (B,P)\,; \\ v - p - \pi_s\alpha_s u(\sigma)v - \pi_i\alpha_i v & if \quad \sigma(v) = (B,NP)\,; \\ v - \delta p - c_a - \pi_a\alpha_a v & if \quad \sigma(v) = (B,AP)\,; \\ 0 & if \quad \sigma(v) = (NB,NP)\,, \end{cases} \tag{2.1}$$

where

$$u(\sigma) \triangleq \int_{\mathcal{V}} \mathbb{1}_{\{\sigma(v) = (B,NP)\}}\, dv\,. \tag{2.2}$$

To avoid trivialities and without loss of generality, we reduce the parameter space to $c_p, c_a \in (0,1)$, $\pi_i, \pi_s, \pi_a \in (0,1]$, $\alpha_i, \alpha_s, \alpha_a \in (0,\infty)$, and $\pi_a\alpha_a \in (0, 1 - c_a)$. The latter restriction, $\pi_a\alpha_a + c_a < 1$, ensures automated patching is economical.

2.3.2 Consumer Market Equilibrium

Before examining how patching rights should be priced, we first must characterize how consumers segment across strategies for an arbitrary set of prices in equilibrium. Complicating the situation is that the level of security risk is endogenously determined by the actions of consumers, thus we first focus on understanding the effect of their strategic interactions on equilibrium behavior due to the security externality imposed by unpatched users. The consumer with valuation v selects an action that solves the following maximization problem:

$$\max_{s \in S} \quad U(v,\sigma)\,, \tag{2.3}$$

where the strategy profile σ is composed of σ_{-v} (which is taken as fixed) and the choice being made, i.e., $\sigma(v) = s$. We denote her optimal action that solves (2.3) with $s^*(v)$. Further, we denote the equilibrium strategy profile with σ^*, and it satisfies the requirement that $\sigma^*(v) = s^*(v)$ for all $v \in \mathcal{V}$.

Lemma 1 *There exists a unique equilibrium consumer strategy profile* σ^* *that is characterized by thresholds* v_b, v_a, $v_p \in [0,1]$. *For each* $v \in \mathcal{V}$, *it satisfies either*

$$\sigma^*(v) = \begin{cases} (B,P) & if \quad v_p < v \leq 1\,; \\ (B,NP) & if \quad v_b < v \leq v_p\,; \\ (B,AP) & if \quad v_a < v \leq v_b\,; \\ (NB,NP) & if \quad 0 \leq v \leq v_a\,, \end{cases} \tag{2.4}$$

or

$$\sigma^*(v) = \begin{cases} (B,P) & if \quad v_p < v \leq 1\,; \\ (B,AP) & if \quad v_a < v \leq v_p\,; \\ (B,NP) & if \quad v_b < v \leq v_a\,; \\ (NB,NP) & if \quad 0 \leq v \leq v_b\,. \end{cases} \tag{2.5}$$

Lemma 1 establishes that if a population of patched consumers arises in equilibrium, it will consist of a segment of consumers with the highest valuations. These consumers prefer to shield themselves from any valuation-dependent losses seen with either remaining unpatched and bearing security losses or selecting automated patching and bearing patch instability losses. Importantly, this segment need not arise, and $v_p = 1$ in cases where the valuation-dependent losses are smaller than the patching costs. For the middle segment on the other hand, the segment of consumers who elect for automated patching and the segment of consumers who elect to remain unpatched can be ordered either way. This ordering depends on the relative strength of the losses under each strategy.

2.4 Pricing Patching Rights

Since the value of patching is most applicable under higher security risk, our study centers on a region where either $\pi_i\alpha_i$ or $\pi_s\alpha_s$ is reasonably high such that patching is worthwhile. Having motivated both independent and dependent security risks, we begin our study by examining the case where independent risk becomes high. There are several merits to starting with this case. First, as we shall see, based on user incentives this case is inherently simpler and ultimately admits closed-form solutions of the equilibrium consumer market thresholds and prices. This will help the reader to build an intuition into how priced patching rights tend to affect equilibrium behaviors. Second, the impact of higher independent risk is similar in nature to that of higher dependent risk in that both tend to reduce unpatched populations as users become unwilling to bear higher risk. In this light, certain limit effects on thresholds and profitability will be the same and can be characterized more easily in this simplified setting. Third, examination of this case underscores why capturing dependent risk is essential to a comprehensive understanding of interdependent security settings which propels the remainder of the paper.

2.4.1 High $\pi_i\alpha_i$

In this section, we study independent risk satisfying $\pi_i\alpha_i > 1$. In Section A.2 of the Appendix, we provide a complete characterization of the parameter conditions and thresholds for each possible consumer market structure that can arise in equilibrium under high independent risk. There are three possible structures, with the two most relevant to the current discussion being $0 < v_a < v_p < 1$ and $0 < v_p < 1$. These two structures obtain under broad parameter conditions, and as will be shown, those conditions can be satisfied under equilibrium pricing decisions. Importantly, under high independent risk, we establish that no user will elect to be unpatched which is to say there are no unpatched users causing an security externality in equilibrium. Thus, examination of this region will provide initial insights into more complex equilibria involving limited unpatched populations as well as the discriminatory role of priced patching rights.

To characterize the pricing equilibrium, we first denote the vendor's profit function by

$$\Pi(p,\delta) = p\int_{\mathcal{V}} \mathbb{1}_{\{\sigma^*(v|p,\delta)\in\{(B,NP),(B,P)\}\}}dv + \delta p\int_{\mathcal{V}} \mathbb{1}_{\{\sigma^*(v|p,\delta)=(B,AP)\}}dv\,, \tag{2.6}$$

noting that marginal costs are assumed to be negligible for information goods. Because we are interested in determining the benefit of optimally pricing the *right* for a user to determine whether or not to install patches on her system, it is useful to first present a characterization of the equilibrium when this right is not priced. In this reference case, referred to throughout the paper as the *status quo*, $\delta = 1$, which is standard practice for the industry. In this case, the vendor sets a price p for use of the software by solving the following problem:

$$\begin{aligned} \max_{p \in [0,\infty)} \quad & \Pi(p,\delta) \\ \text{s.t.} \quad & (v_b, v_a, v_p) \text{ are given by } \sigma^*(\cdot \mid p, \delta), \\ & \delta = 1. \end{aligned} \tag{2.7}$$

Given a price p^* that solves (2.7), we denote the profits associated with this optimal price by $\Pi_{SQ} \triangleq \Pi(p^*, 1)$. Since the value of automated patching options on security is most applicable under higher security risk, our study centers on a region where either $\pi_i \alpha_i$ or $\pi_s \alpha_s$ is suitably high such that patching is worthwhile.

Lemma 2 (Status Quo) *Suppose that $\pi_i \alpha_i > 1$ and $\delta = 1$ (i.e., when patching rights are not priced).*

(i) If $c_p - \pi_a \alpha_a < c_a \leq 1 - \pi_a \alpha_a - (1 - c_p)\sqrt{1 - \pi_a \alpha_a}$, then

$$p^* = \frac{1 - \pi_a \alpha_a - c_a}{2}, \tag{2.8}$$

and σ^* is characterized by $0 < v_a < v_p < 1$ such that the lower tier of users prefers automated patching.

(ii) On the other hand, if $c_a > 1 - \pi_a \alpha_a - (1 - c_p)\sqrt{1 - \pi_a \alpha_a}$, then

$$p^* = \frac{1 - c_p}{2}, \tag{2.9}$$

and σ^* is characterized by $0 < v_p < 1$ such that there is no user of automated patching in equilibrium.

Lemma 2 presents the equilibrium behavior under the status quo reference case. In part (i), we see that as long as the cost of standard patching (c_p) and automated patching (c_a) satisfy conditions where c_a is *relatively* moderate, both (B, P) and (B, AP) are observed strategies in equilibrium. On the other hand, as c_a increases to a higher level, only standard patching is observed in equilibrium. Lemma 2 highlights that a higher independent risk essentially squeezes out unpatched behaviors, i.e., (B, NP), leading to the absence of a security externality in equilibrium. In this context, it is useful to understand the role of priced patching rights as it will serve as a contrastable reference point for when externalities become a driving force.

When patching rights are priced, the vendor jointly selects (p, δ) to maximize his profits. The premium $p(1-\delta)$ is charged for patching rights, regardless of whether the patching rights are exercised. Alternatively, $p(1-\delta)$ can be considered the "discount" given to users who agree to have their systems automatically updated to reduce security risk on the network. When patching rights are priced in this fashion, the vendor's pricing problem is formulated as follows:

$$\begin{aligned} &\max_{(p,\delta)\in[0,\infty)^2} && \Pi(p,\delta) \\ &\text{s.t.} && (v_b, v_a, v_p) \text{ are given by } \sigma^*(\cdot|p,\delta). \end{aligned} \tag{2.10}$$

Similarly, under the optimal $(p^*,\ \delta^*)$ which solve (2.10), we denote the associated profits of pricing patching rights by $\Pi_P \triangleq \Pi(p^*,\ \delta^*)$.[6]

Lemma 3 (Priced Patching Rights) *Suppose that $\pi_i\alpha_i > 1$ and patching rights are priced by the vendor.*

(i) If $c_p - \pi_a\alpha_a < c_a \le c_p(1-\pi_a\alpha_a)$, then

$$p^* = \frac{1-c_p}{2}, \tag{2.11}$$

$$\delta^* = \frac{1-c_a-\pi_a\alpha_a}{1-c_p}, \tag{2.12}$$

[6] Going forward, we will use subscripts "SQ" and "P" to indicate a particular measure refers to the outcome under the status quo and under priced patching rights, respectively, for consistency.

and σ^* is characterized by $0 < v_a < v_p < 1$ such that the lower tier of users prefers automated patching.

(ii) On the other hand, if $c_a > c_p(1 - \pi_a \alpha_a)$, then

$$p^* = \frac{1 - c_p}{2}\,, \tag{2.13}$$

and σ^* is characterized by $0 < v_p < 1$ such that there is no user of automated patching in equilibrium.

Lemma 3 formally establishes that pricing patching rights can greatly expand the region of the parameter space on which automated patching is observed, relative to the status quo. Specifically, when $1 - \pi_a\alpha_a - (1 - c_p)\sqrt{1 - \pi_a\alpha_a} \le c_a \le c_p(1 - \pi_a\alpha_a)$ is satisfied, then $0 < v_a < v_p < 1$ is induced when patching rights are priced whereas $0 < v_p < 1$ is induced in the status quo. Additionally, if part (i) of Lemma 3 is satisfied, then patching rights are priced in a way that consumers who select automated patching in equilibrium form the lower tier of the consumer market. By (2.11) and (2.12), it is easy to see that the premium charged for patching rights $p^*(1 - \delta^*)$ is only greater than c_a when c_p is small enough. When these standard patching costs are small, the software vendor has an incentive to charge a high price for his software. One can think of it as being better software that is easily maintained via a cost-effective, rigorous patching process. In this case, even with a high price, the vendor can still achieve a sizable user population, most of which elects for standard patching. Because of the attractiveness of standard patching, it is necessary to provide a significant discount to incentivize users to prefer the automated patching option. This is reflected in (2.12); δ^* decreases as c_p decreases. From the other perspective, the patching rights premium $p^*(1 - \delta^*)$ is substantial and can be an incentive-compatible option only for high valuation users. Low valuation users necessarily find the patching rights premium too high to bear, instead opting for automated patching in equilibrium.

Next, we examine the value of pricing patching rights for the software vendor.

Proposition 1 *When $\pi_i\alpha_i > 1$, if $c_p - \pi_a\alpha_a < c_a < 1 - \pi_a\alpha_a - (1-c_p)\sqrt{1-\pi_a\alpha_a}$, the percentage increase in profitability of pricing patching rights is given by*

$$\frac{\Pi_P - \Pi_{SQ}}{\Pi_{SQ}} = \frac{(1-\pi_a\alpha_a)(c_a - c_p + \pi_a\alpha_a)^2}{\pi_a\alpha_a(1-c_a-\pi_a\alpha_a)^2}\,, \tag{2.14}$$

where

$$\Pi_{SQ} = \frac{(1-c_a-\pi_a\alpha_a)^2}{4(1-\pi_a\alpha_a)}\,. \tag{2.15}$$

Proposition 1 formally establishes the extent to which a priced patching rights policy can increase profitability for the vendor. In the context of our overall study, what is important to emphasize here is that the vendor can have strong incentives to leverage an automated patching version toward discriminatory purposes. In particular, under high independent risk, Lemmas 2 and 3 establish that no unpatched usage arises in equilibrium and so it is not the case that the priced patching rights policy being employed aims to reduce any externality on the network. The segmentation behavior seen here solely targets extraction of surplus from high valuation users by inducing them to pay the patching rights premium. These users have much to lose in the event of any system failure occurring due to patch instability, and thus are willing to pay to retain control and continue to exercise diligence in their patching processes.

Users whose valuations are not high will find it incentive-compatible to relinquish patching rights. In fact, a sizable segment of consumers switch from standard patching towards automated patching when patching rights are priced. As the proofs of part (i) of Lemmas 2 and 3 establishes, the user type indifferent between using automated patching and not using, v_a, is identical both in the status quo and under priced patching rights. Viewed in that light, priced patching rights does not expand usage in the market in this case, instead only serving to encourage some users to make less loss-efficient security choices yet benefitting vendor profitability.

This proposition also shows that a priced patching rights strategy always outperforms a mandated automated patching policy here. The reason is because mandating automated patching is a special case of a priced patching rights policy.

Specifically, the decision problem when mandating automatic patching can be formulated as the priced patching rights problem, subject to the constraint that $p = 1$. As seen in the utility function given by (1), setting $p = 1$ makes the strategies (B, NP) and (B, P) infeasible to consumers, and the vendor chooses δ to maximize profits with all consumers now only considering (B, AP). Despite $p = 1$ being feasible, it is never chosen, showing that mandating automatic patching leads to strictly lower profits here.

A vendor's patch release frequency impacts the value of pricing patching rights in terms of relative profitability. A frequent patch release policy imposes additional burden on companies who follow a standard patching policy.[7] In our model abstraction, higher frequency corresponds to higher standard patching costs (i.e., higher c_p). By equation (2.14), we see that the relative profitability of priced patching rights is decreasing in c_p. Because higher standard patching costs naturally incentivize users to shift toward automated patching usage rather than unpatched usage (due to high $\pi_i \alpha_i$), the upside of priced patching rights becomes limited. More frequently released patches can also reduce costs associated with patch instability because problems are much easier to diagnose when scope is narrower. Examining (2.14), we see that relative profitability is similarly decreasing as $\pi_a \alpha_a$ decreases, which is to say that the relative value of priced patching rights is generally higher when patches tend to be more bundled.

Proposition 1 highlights the discriminatory forces at work when the vendor can separately price an automated patching version of his product without being concerned about security externalities. On the other hand, equilibrium consumer market outcomes marked by no user being unpatched call attention to the source of the risk. In particular, one might ponder why $\pi_i \alpha_i$ is high if nobody is unpatched? Even if a few users out of a large network were unpatched, should we expect them to face high risk? This line of thought suggests that security risk and the size of the unpatched population may naturally have some dependence, and in the following

[7]The class of software product determines the type of users: individuals and/or businesses. Our model concerns itself with the valuation of the software product by the user regardless of user type. For software classes with both types of users, it is likely that being a business user is correlated with higher valuation, in which case the lower valuation, individual users will be the ones targeted by the priced patching rights policy.

section we examine regions where both independent and dependent risks arise in equilibrium. By doing so and contrasting with the current section's results, we will more clearly see the relative impact of security externalities on the vendor's pricing of patching rights and its profitability. The case of a high $\pi_s \alpha_s$ has some similarities to when $\pi_i \alpha_i$ is high but also some significant differences which stem from unpatched usage. Importantly, we will demonstrate the efficacy of priced patching rights on reducing the security externalities that have plagued network software.

2.4.2 Low $\pi_i \alpha_i$

In this section, we turn our attention to a region where independent risk is more moderate in nature (i.e., explicitly bounded above) and study equilibria as dependent risk becomes high. In contrast to the prior section, we show that there exists a segment of unpatched users despite the increased risk. The characterization of the thresholds that emerge in the equilibrium consumer market structure in this case become significantly more complex, satisfying a nonlinear system of equations. Therefore, for this region we employ asymptotic analysis which is commonly employed in microeconomic studies. Its use can be expected here due to the complexity of the game and corresponding equilibrium characterization (some examples of studies using asymptotic analysis include Li et al. 1987, Laffont and Tirole 1988, MacLeod and Malcomson 1993, Pesendorfer and Swinkels 2000, Muller 2000, Tunca and Zenios 2006, August and Tunca 2006, Pei et al. 2011 among many others). Miller (2006) and Cormen et al. (2009) provide comprehensive treatments of the mathematical foundation underlying asymptotic analysis. Due to model complexity in this region, some boundaries do not have explicit functional forms, in contrast to the prior the section. However, the objective of the analysis is the identification of regions of applicability in terms of parameter characteristics, which is the focus of our formalized results.

As before, we first characterize the consumer market equilibrium when patching rights are freely included.

Lemma 4 (Status Quo) *There exists ω such that when $\pi_s\alpha_s > \omega$, if $\pi_i\alpha_i < \min\left[\frac{c_p\pi_a\alpha_a}{1+c_p-c_a}, \frac{c_p}{1+c_p}\right]$, $\delta = 1$ (i.e., when patching rights are not priced),*

(i) if $c_p - \pi_a\alpha_a < c_a \leq 1 - \pi_a\alpha_a - (1-c_p)\sqrt{1-\pi_a\alpha_a}$, then

$$p^* = \frac{1}{2}(1-\pi_a\alpha_a - c_a) + \Bigg(2c_a^2(\pi_a\alpha_a - 1)\big((\pi_a\alpha_a - 1)(2\pi_a\alpha_a - \pi_i\alpha_i - 1) + c_a(2\pi_a\alpha_a + \pi_i\alpha_i - 3)\big)\Bigg)\Bigg(\big(-\pi_a\alpha_a + c_a + 1\big)^3\pi_s\alpha_s\Bigg)^{-1} + K_a\,, \tag{2.16}$$

and σ^* is characterized by $0 < v_b < v_a < v_p < 1$ such that the lower tier of users remain unpatched and the middle tier prefers automated patching.

(ii) On the other hand, if $c_a > 1 - \pi_a\alpha_a - (1-c_p)\sqrt{1-\pi_a\alpha_a}$, then

$$p^* = \frac{1-c_p}{2} - \frac{2c_p^2(1 - 3c_p + \pi_i\alpha_i(1+c_p))}{(1+c_p)^3\pi_s\alpha_s} + K_b\,, \tag{2.17}$$

and σ^* is characterized by $0 < v_b < v_p < 1$ such that there is no user of automated patching in equilibrium.[8]

Similar to Lemma 2, again we first examine the reference case when the cost of automated patching is relatively moderate. An immediate observation is that when patching rights are free as in the status quo, the consumer segment whose equilibrium strategy is to use automated patching is always the middle tier. This occurs because when a user compares an automated patching strategy (B, AP) to an unpatched strategy (B, NP), the price is the same for both options, i.e., $p = \delta p$ when $\delta = 1$. Therefore, the strategy (B, AP) is preferred to (B, NP) as long as $v[\pi_s\alpha_s u(\sigma^*) + \pi_i\alpha_i - \pi_a\alpha_a] > c_a$ is satisfied. If this inequality is satisfied for any user with valuation v, it will also be satisfied for any user with a valuation higher than v. Thus, the automated patching segment of users will always form the middle tier. Contrasting this to the previous section, under high independent security risk the automated patching segment formed the lowest tier. This cannot happen in

[8]The existence of ω is proven in the Appendix. The characterization of constants denoted by K and enumerated by a subscript are similarly provided.

the current region because patching rights are endowed.[9]

This observation highlights an important potential impact of priced patching rights; if the premium charged for patching rights, $p(1-\delta)$, is greater than c_a and the unpatched population, $u(\sigma^*)$, decreases enough in equilibrium, then the lower tier can instead be composed of users who strategically choose automated patching (see (2.4) in Lemma 1). In this sense, a priced patching rights policy can fundamentally change segmentation behavior in the consumer market which in turn can have a significant impact on security and profitability. The following lemma formalizes the equilibrium strategies under priced patching rights.

Lemma 5 (Priced Patching Rights) *Suppose that $\pi_s\alpha_s > \omega$, that patching rights are priced by the vendor, and that $\pi_i\alpha_i < \min\left[\frac{c_p\pi_a\alpha_a}{1+c_p-c_a}, \frac{c_p}{1+c_p}\right]$.*

(i) If $c_a < \min\left[\pi_a\alpha_a - c_p, c_p(1-\pi_a\alpha_a)\right]$, then

$$p^* = \tilde{p} + \Big(2\pi_a\alpha_a c_p\big(\pi_i\alpha_i c_a^2 + c_a(-\pi_a\alpha_a\pi_i\alpha_i + 3\pi_a\alpha_a c_p - 2c_p\pi_i\alpha_i) + c_p(\pi_a\alpha_a(\pi_a\alpha_a + \pi_i\alpha_i) + c_p(\pi_i\alpha_i - 3\pi_a\alpha_a))\big)\Big) \Big(\pi_s\alpha_s(-\pi_a\alpha_a + c_a - c_p)^3\Big)^{-1} + K_c, \quad (2.18)$$

$$\delta^* = \tilde{\delta} - \Big(4c_p\pi_a\alpha_a(\pi_a\alpha_a + c_a - 1)\big(\pi_i\alpha_i c_a^2 + c_a(-\pi_a\alpha_a\pi_i\alpha_i + 3\pi_a\alpha_a c_p - 2\pi_i\alpha_i c_p) + c_p(\pi_a\alpha_a(\pi_a\alpha_a + \pi_i\alpha_i) + c_p(\pi_i\alpha_i - 3\pi_a\alpha_a))\big)\Big) \Big(\pi_s\alpha_s(c_p-1)^2(\pi_a\alpha_a - c_a + c_p)^3\Big)^{-1} + K_d, \quad (2.19)$$

and σ^* is characterized by $0 < v_a < v_b < v_p < 1$ under optimal pricing, where $\tilde{p} = \frac{1-c_p}{2}$ and $\tilde{\delta} = \frac{1-\pi_a\alpha_a-c_a}{1-c_p}$, such that the lower tier of users prefer automated patching and the middle tier remains unpatched;

[9] Under low independent risk whenever a standard patching population arises in equilibrium, there must also be a population of unpatched users. Otherwise, a standard patching user would deviate to being unpatched and bear no risk.

(ii) if $|\pi_a\alpha_a - c_p| < c_a < c_p(1-\pi_a\alpha_a)$, *then*

$$p^* = \tilde{p} + \frac{c_a(c_a - c_p\pi_a\alpha_a + c_p)(c_a(1-\pi_i\alpha_i) + (1-\pi_a\alpha_a)(-\pi_i\alpha_i + 2c_p - 1))}{\pi_s\alpha_s(1+c_a-\pi_a\alpha_a)^3} + K_e, \quad (2.20)$$

$$\delta^* = \tilde{\delta} - \Big(2c_a\big(c_a^2 + (\pi_a\alpha_a - 1)\big(\pi_a\alpha_a + c_p^2 - 2c_p\pi_a\alpha_a\big)\big)\big(c_a(\pi_i\alpha_i - 1) + (\pi_a\alpha_a - 1)(-\pi_i\alpha_i + 2c_p - 1)\big)\Big)\Big((c_p - 1)^2\pi_s\alpha_s(-\pi_a\alpha_a + c_a + 1)^3\Big)^{-1} + K_f, \quad (2.21)$$

and σ^* *is characterized by* $0 < v_b < v_a < v_p < 1$ *under optimal pricing such that the lower tier of users remains unpatched and the middle tier prefers automated patching;*

(iii) if $c_a > c_p(1-\pi_a\alpha_a)$, *then*

$$p^* = \tilde{p} - \frac{2c_p^2(1 - 3c_p + \pi_i\alpha_i(1+c_p))}{(c_p+1)^3\pi_s\alpha_s} + K_g, \quad (2.22)$$

$$\delta^* = 1, \quad (2.23)$$

and σ^* *is characterized by* $0 < v_b < v_p < 1$ *under optimal pricing.*

Lemma 5 demonstrates that a restructuring of the consumer market can indeed be the equilibrium outcome when patching rights are priced. Specifically, if the patching costs are small such that part (i) of Lemma 5 is satisfied, then the equilibrium patching rights are priced in a way that consumers who select automated patching in equilibrium form the lower tier of the consumer market. This outcome more closely resembles the structure that emerges in part (i) of Lemma 3, with a similar driving force. Specifically, small standard patching costs prompt a high patching rights premium that low valuation users are unwilling to assume.

On the other hand, when the patching rights premium is limited, the equi-

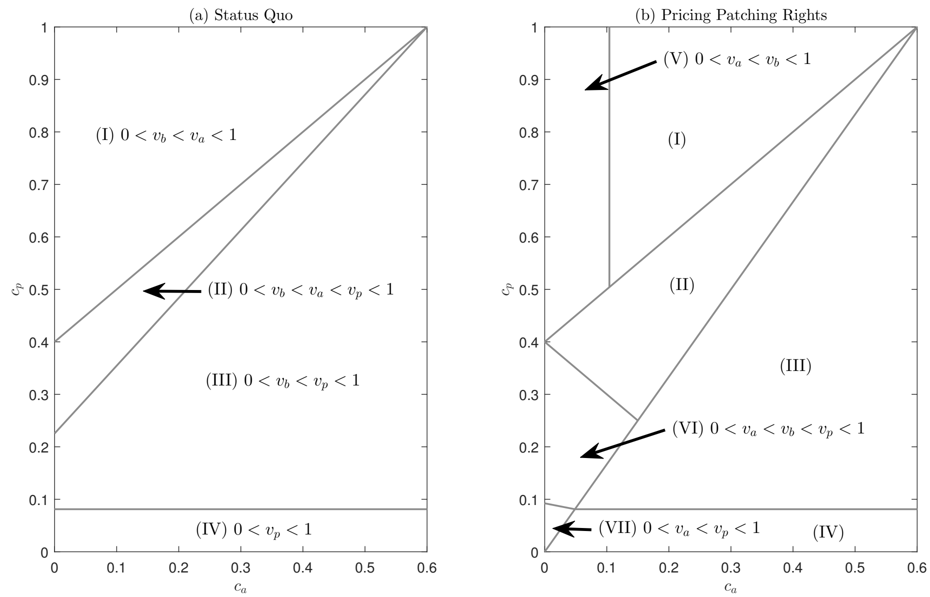


Figure 2.1: Consumer Market Structures Under Equilibrium Pricing. Panel (a) illustrates the market structures that arise in the status quo case, and panel (b) illustrates the market structures that arise under priced patching rights. Other parameter values are $\pi_i\alpha_i = 0.15$ and $\pi_a\alpha_a = 0.4$.

librium price and discount induce a consumer market structure that remains consistent with what unfolds under the status quo, but contrasts with the high independent risk case. Part (ii) of Lemma 5 shows that this structure is characterized by the threshold ordering $0 < v_b < v_a < v_p < 1$, matching the threshold ordering in the first part of Lemma 4. Thus, under both the status quo and under priced patching rights, the middle tier is incentivized to select the automated patching option in equilibrium.

Figure 2.1 demonstrates how pricing patching rights significantly affects the consumer market structures that are obtained in equilibrium. In each panel, we illustrate the consumer market structure threshold characterization that obtains in equilibrium as a function of standard patching (c_p) and automated patching (c_a) costs. For the selected parameter set, panel (a) shows that four possible market structures can arise. When c_p is relatively high as in Region (I), it becomes impractical to conduct standard patching processes. In this case, even high valuation consumers are willing to bear the patch instability risk associated with automated patching, and the consumer market structure is characterized by $0 < v_b < v_a < 1$. On the other extreme, in Regions (III) and (IV) where c_p is relatively low, automated patching is not observed in equilibrium. Because of the cost effectiveness of standard patching processes, the consumer market structure becomes either $0 < v_b < v_p < 1$ or $0 < v_p < 1$. Finally, when standard patching and automated patching costs are comparable as seen in Region (II), the vendor's pricing leads to an equilibrium characterized by $0 < v_b < v_a < v_p < 1$ in which both types of patching segments are present.

When patching rights are priced, there are two distinct, induced changes to user behavior that are illustrated in panel (b) of Figure 2.1. First, the region over which automated patching is preferred by some consumers in equilibrium significantly expands under priced patching rights. For this to occur, the region of the parameter space over which automated patching behavior is absent under the status quo shrinks upon pricing patching rights. This is easily visualized by Regions (III) and (IV) decreasing in size when moving from panel (a) to panel (b). The expansion of automated patching behavior is a critical effect of a priced patching rights policy because it goes hand-in-hand with a decreased unpatched population that will help to improve security risk.

Second, priced patching rights can create entirely new market structures that are not observed under the status quo. Region (VI) of panel (b) illustrates a region of the parameter space in which the threshold characterization is given by $0 < v_a < v_b < v_p < 1$. When patching rights are endowed as in the status quo, it is not possible to get low valuation users to switch to the automated patching solution and therefore they remain as unpatched, externality-contributing users. Under priced patching rights and low c_p (hence a high premium as discussed previously), it becomes no longer incentive compatible for these users to remain as unpatched. This behavior however can change as c_a increases. This is illustrated as a shift from Region (VI) to either Region (II) or Region (III) in panel (b) of Figure 2.1.

Building on this understanding of the impact of priced patching rights, we next aim to provide clarity into the strategic behavior underlying the vendor's pricing as well as its impact on the security of software networks and the software industry as a whole.

Proposition 2 *There exists a bound $\tilde{\alpha}_s$ such that when $\alpha_s > \tilde{\alpha}_s$, if $c_p - \pi_a\alpha_a < c_a < 1 - \pi_a\alpha_a - (1-c_p)\sqrt{1-\pi_a\alpha_a}$ and $\pi_i\alpha_i < \min\left[\frac{c_p\pi_a\alpha_a}{1+c_p-c_a}, \frac{c_p}{1+c_p}\right]$, then pricing patching rights can improve profits while reducing the security externality generated by unpatched users as compared to when patching rights are not priced. When $c_a < \min\left[\pi_a\alpha_a - c_p, c_p(1-\pi_a\alpha_a)\right]$, the improvement in profitability is given by*

$$\begin{aligned}\frac{\Pi_P - \Pi_{SQ}}{\Pi_{SQ}} = \frac{(1-\pi_a\alpha_a)(c_a - c_p + \pi_a\alpha_a)^2}{\pi_a\alpha_a(1-c_a-\pi_a\alpha_a)^2} + \Big((\pi_a\alpha_a - c_a + c_p)^2 M - \\ 4\pi_a\alpha_a(1-\pi_a\alpha_a)(1-\pi_a\alpha_a-c_a)(-\pi_a\alpha_a+c_a+c_p)\big(\pi_a\alpha_a((c_a+2)c_p+c_a)+ \\ c_a(1-c_p)(c_p-c_a) - 2c_p(\pi_a\alpha_a)^2\big)(c_a - c_p(1-\pi_a\alpha_a)) + 4\pi_i\alpha_i(\pi_a\alpha_a-1)(-\pi_a\alpha_a+c_a+1) \\ (-\pi_a\alpha_a+c_a-c_p)\big((\pi_a\alpha_a)^3(\pi_a\alpha_a)^2(c_a+c_p+1) - \pi_a\alpha_a(c_a-1)(c_a+c_p) + c_a(c_a-c_p)^2\big) \\ (c_a + c_p(\pi_a\alpha_a-1))\Big)\Big(\pi_a\alpha_a\pi_s\alpha_s(-\pi_a\alpha_a+c_a+1)^2(\pi_a\alpha_a+c_a-1)^3(\pi_a\alpha_a-c_a+c_p)^2\Big)^{-1} + \\ K_h\,, \quad (2.24)\end{aligned}$$

where

$$\begin{aligned}M = 4c_a(\pi_a\alpha_a - 1)(\pi_a\alpha_a + c_a - c_p)\Big((\pi_a\alpha_a + c_a)\big(\pi_a\alpha_a(2-\pi_a\alpha_a) + c_a(\pi_a\alpha_a - 2)+ \\ 2c_p(\pi_a\alpha_a - 1)^2\big) - \pi_a\alpha_a\Big)\,, \quad (2.25)\end{aligned}$$

$$\begin{aligned}\Pi_{SQ} = \frac{(1-c_a-\pi_a\alpha_a)^2}{4(1-\pi_a\alpha_a)} + \\ \frac{c_a(1-c_a-\pi_a\alpha_a)((1-\pi_a\alpha_a)(\pi_a\alpha_a - \pi_i\alpha_i) + c_a(2-\pi_a\alpha_a-\pi_i\alpha_i))}{(1+c_a-\pi_a\alpha_a)^2\pi_s\alpha_s} + K_i\,, \quad (2.26)\end{aligned}$$

and the reduction in the size of the unpatched population is given by

$$u^*_{SQ} - u^*_P = \Big(\big(c_a^2(\pi_a\alpha_a - 2) + c_a(\pi_a\alpha_a + c_p(2 - 3\pi_a\alpha_a)) +$$
$$\pi_a\alpha_a(\pi_a\alpha_a - 1)(c_p - \pi_a\alpha_a)\big)\Big)\Big(\pi_s\alpha_s(-\pi_a\alpha_a + c_a + 1)(\pi_a\alpha_a - c_a + c_p)\Big)^{-1} + K_j\,. \tag{2.27}$$

When $|\pi_a\alpha_a - c_p| < c_a < c_p(1 - \pi_a\alpha_a)$, the improvement in profitability is given by

$$\frac{\Pi_P - \Pi_{SQ}}{\Pi_{SQ}} = \frac{(1 - \pi_a\alpha_a)(c_a - c_p + \pi_a\alpha_a)^2}{\pi_a\alpha_a(1 - c_a - \pi_a\alpha_a)^2} +$$
$$\Big(M + (-\pi_a\alpha_a + c_a + 1)\big(4\pi_i\alpha_i c_a(\pi_a\alpha_a - 1)(\pi_a\alpha_a + c_a - c_p)$$
$$(c_a + c_p(\pi_a\alpha_a - 1))\big)\Big)\Big(\pi_a\alpha_a\pi_s\alpha_s(-\pi_a\alpha_a + c_a + 1)^2(\pi_a\alpha_a + c_a - 1)^3\Big)^{-1} + K_k\,, \tag{2.28}$$

where Π_{SQ} is given by (2.26) *and the reduction in the size of the unpatched population is given by*

$$u^*_{SQ} - u^*_P = \frac{(1 - \pi_a\alpha_a)(\pi_a\alpha_a + c_a - c_p)}{\pi_s\alpha_s(-\pi_a\alpha_a + c_a + 1)} + K_l\,. \tag{2.29}$$

Proposition 2 highlights an important message from our study: software vendors should consider differentiation of their products based on patching rights. Simply providing patches for security vulnerabilities of software to users as a security strategy has not worked well in the past. In many cases, this leads to large unpatched user populations as these users determine its not in their best interest to patch. The externality they cause is detrimental to security and to the vendor's profitability. Proposition 2 formally establishes that the proper pricing of patching rights can increase profits for vendors to an extent characterized in (2.24) and (2.28), and simultaneously reduce the size of the unpatched population in the network. Thus, there are large potential economic and security benefits associated with a priced patching rights strategy, which can be an important pricing paradigm shift for the software industry.[10]

Product differentiation is an important topic studied in economics and marketing, and the versioning of information goods has further nuanced findings (Bhargava and Choudhary 2001, 2008, Johnson and Myatt 2003). In particular, for these goods which have a negligible marginal cost of reproduction, a software vendor finds it optimal to release only one product (no versioning) when consumers heterogeneous taste for quality is uniformly distributed. In such a case, cannibalization losses outweigh differentiation benefits. In the current work, Proposition 2 demonstrates that if the versioning is instead on patching rights, a versioning strategy is once again optimal for the vendor. In this case, the software vendor can profitably benefit by increasing the price of the version with patching rights (p^*) relative to the price point under the status quo. By doing so, while concurrently decreasing the price of the version without patching rights (automated patching only) to ($\delta^* p^*$), there are several effects as consumers strategically respond. First, a higher p^* puts pressure on any user who would be unpatched under the status quo to reconsider the trade-off. Under the status quo equilibrium consumer market structure (i.e., $0 < v_b < v_a < v_p < 1$), the unpatched users form the lower tier of the consumer market (those with valuations between $[v_b, v_a]$). Because patching rights are endowed to all users under the status quo, these users remain unpatched and contribute to a larger security externality on the network. Under priced patching rights, a higher p^* makes it now more expensive to remain in the population as an unpatched user causing this externality. Second, given the new equilibrium prices, it becomes relatively cheaper to opt for automated patching at a discount of $p^*(1 - \delta^*)$. This provides additional incentives to encourage better security behaviors. On the other hand, a higher price can be detrimental to usage and associated revenues, and a reduced unpatched population can create incentives for users who were patching under the status quo to now remain unpatched.

The net impact of these effects depends on which consumer market structure is induced by the vendor's new prices. As Lemma 5 demonstrates, the ven-

[10]In July of 2015, Microsoft made automatic updates mandatory for Windows 10 Home users (Newman 2015). Notably, users who pay for the premium version, Windows 10 Pro, maintain greater ability to control patching. These actions are consistent with the spirit of our results.

dor may induce a segmentation characterization of either $0 < v_a < v_b < v_p < 1$ or $0 < v_b < v_a < v_p < 1$. We begin by discussing the latter structure since it matches the status quo. In equilibrium under patched pricing rights, v_b increases and v_a decreases relative to the status quo. Thus, the size of the unpatched population (i.e., $u = v_a - v_b$) shrinks as it is compressed on both ends. However, v_p increases because of the patching rights premium. In aggregate, the vendor is able to increase profitability by charging a premium to high tier consumers (valuations in $[v_p, 1]$) who are willing to pay the premium to protect their valuations from the incurrence of either unpatched security losses or automated patching instability losses, and low tier consumers (valuations in $[v_b, v_a]$ are willing to pay the premium because of the smaller security externality that is associated with a smaller equilibrium unpatched user population.

For the former case in which the vendor sets prices such that the equilibrium consumer market structure characterization takes the form $0 < v_a < v_b < v_p < 1$, there is a restructuring in the consumer market segments (see the discussion following Lemma 5). It is in the vendor's best interest to have a relatively large patching rights premium in this region which makes the retaining of patching rights only incentive compatible for the higher valuation users. Low valuation users respond to a substantial discount by forgoing patching rights and switching to automated patching. Because low valuation users tend to be the ones with reduced incentives to patch and protect themselves, the market segmentation that occurs also leads to a smaller unpatched population and less resultant security risk. In a similar spirit to the discussion above, this is profitable to the vendor as it is able to raise prices due to greater security and high valuations users' willingness to pay to retain patching rights.

By characterizing the percentage improvement in profitability associated with a priced patching rights strategy in (2.24) and (2.28), we can highlight the type of market characteristics where efforts for a vendor to reexamine patching rights is more fruitful. In particular, the relative improvement in profitability is increasing in c_a and decreasing in c_p. As the cost of automated patching increases through the relevant region (see Proposition 2), under the status quo the vendor necessarily reduces the software's price to make the automated patching option continue to be affordable. This is important because it prevents a significant loss in users resulting from higher security risk that can arise if the automated patching option becomes too costly. On the other hand, under priced patching rights the vendor can achieve a similar effect by strategically adjusting the discount targeted to the users of the automated patching option rather than the entire user population. When the cost of standard patching (c_p) decreases, the vendor achieves a relatively larger increase in profits. In this case, the premium charged to users who elect to retain patching rights can be increased as patching costs become lower.

In Section 2.4.1, we established that high independent risk ($\pi_i \alpha_i$) precludes a segment of unpatched users from forming in equilibrium. In that sense, provided that the level of risk satisfies an explicit lower bound, small changes in $\pi_i \alpha_i$ cannot affect profitability. However, because independent risk is at a low to moderate level in this section, we can examine its impact on the profitability of a priced patching rights policy.

Corollary 1 *There exists a bound $\tilde{\alpha}_s$ such that when $\alpha_s > \tilde{\alpha}_s$, if $\pi_i \alpha_i < \min\left[\frac{c_p \pi_a \alpha_a}{1+c_p-c_a}, \frac{c_p}{1+c_p}\right]$, then the increase in profitability upon pricing patching rights decreases in α_i.*

Corollary 1 shows that the profitability of priced patching rights tends to decrease in the risk stemming from attacks that are independent of unpatched behavior. Said differently, the value of a priced patching rights strategy is higher for the vendor when users have lower inherent incentives to patch. In today's computing environment, we observe that the most commonly exploited vulnerabilities are ones with patches available, some having been available for years. This observed, persistent unpatched usage of software is strongly suggestive that $\pi_i \alpha_i$ itself must be limited in magnitude; this is a necessary condition for unpatched usage to exist. These observations coupled with Corollary 1 imply that real-world parameter sets will tend to be on a portion of the space where a priced patching rights strategy has relatively increased profitability.

In Figure 2.2, we illustrate the value of pricing patching rights in varying security-loss environments. We plot the percentage increase in profitability under

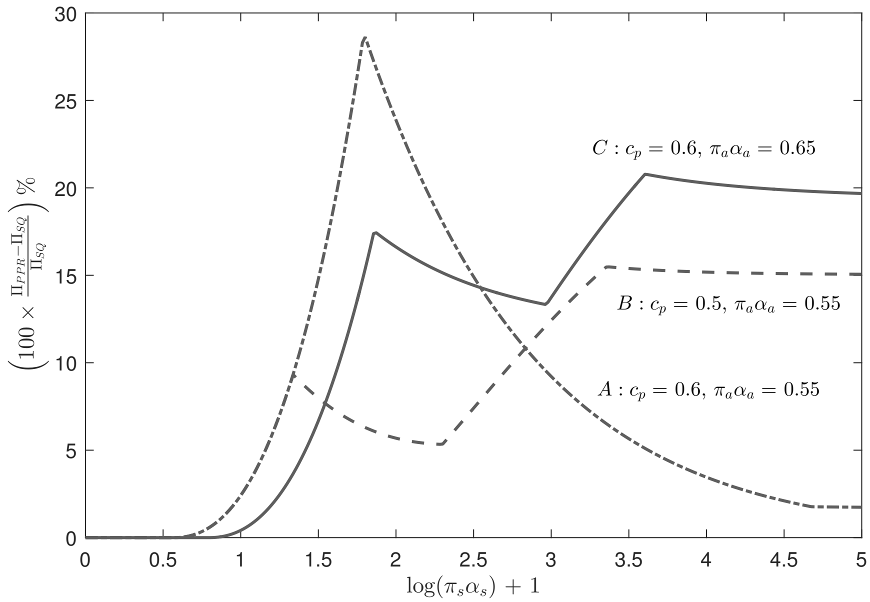


Figure 2.2: Percentage Increase in Profitability as a Function of α_s.
The common parameter values are $\pi_i\alpha_i = 0.05$ and $c_a = 0.1$. The values for the standard patching cost and automated patch instability risk are given in the plot.

three parameter sets A: $c_p = 0.6$, $\pi_a\alpha_a = 0.55$, B: $c_p = 0.5$, $\pi_a\alpha_a = 0.55$, and C: $c_p = 0.6$, $\pi_a\alpha_a = 0.65$, while maintaining one parameter constant across any two cases. Both status quo pricing and priced patching rights induce the consumer market structure characterization $0 < v_b < v_a < v_p < 1$ under high dependent risk (starting near $\log(\pi_s\alpha_s) = 3.5$) for all curves. As was discussed after Proposition 1, under these market characteristics, the percentage increase in profitability is decreasing in c_p; this can be seen by comparing curve A to B. On the other hand, the relative profit improvement increases in $\pi_a\alpha_a$ which can be observed by comparing curve A to C. Figure 2.2 demonstrates that a priced patching rights policy can easily boost profits by 10% even under relatively moderated security risk, and in fact can feasibly lead to an increase in profits that exceeds 20% under particular conditions.

Contrasting with Section 2.4.1 where unpatched usage does not arise in equilibrium, we next examine the value of priced patching rights as it relates to the equilibrium unpatched usage that obtains in this section under low $\pi_i\alpha_i$.

Corollary 2 *There exists a bound $\tilde{\alpha}_s$ such that when $\alpha_s > \tilde{\alpha}_s$, if $\pi_i\alpha_i < \min\left[\frac{c_p\pi_a\alpha_a}{1+c_p-c_a}, \frac{c_p}{1+c_p}\right]$ and $c_p - \pi_a\alpha_a < c_a \leq 1 - \pi_a\alpha_a - (1-c_p)\sqrt{1-\pi_a\alpha_a}$, then the reduction in the size of the unpatched population, $u^*_{SQ} - u^*_P$, is decreasing in c_p and increasing in $\pi_a\alpha_a$.*

In Section 2.4.1, we discussed how a vendor's patch release frequency can impact the value of priced patching rights in terms of relative profitability for the vendor. Some of the insights from that discussion carry over to this section, where some consumers remain unpatched in equilibrium even under high effective security risk. However, in addition, we can also study how a vendor's patch release frequency impacts the value of priced patching rights in terms of reducing risk due to unpatched usage. Very frequent patch release policies impose a burden for companies who have standard patching policies. Corollary 2 establishes that, for software vendors whose market outcomes currently have all segments represented under status quo pricing (i.e., $0 < v_b < v_a < v_p < 1$), the impact on security would be greater for software with bundled patch releases than for more frequent patch releases. Similar to Section 2.4.1, software with bundled patch releases have higher expected automated patching losses that push low-valuation consumers toward unpatched usage. In this sense, vendors who currently bundle their patches instead of using a frequent patch release strategy would have the most to gain in terms of improving software security through the pricing of patching rights.

Another interesting implication of our model concerns a comparison of prices under the status quo and under optimally priced patching rights. One might expect that if p^*_{SQ} is the price under the status quo, then $\delta^* p^* < p^*_{SQ} < p^*$ is satisfied under equilibrium when patching rights are priced. That is, users who want to retain patching rights pay a premium and users who opt for automated patching receive a discount relative to the status quo. However, in the following proposition we demonstrate that the vendor may strategically raise both prices in equilibrium, in comparison to status quo pricing.

Proposition 3 *There exists a bound $\tilde{\alpha}_s$ such that when $\alpha_s > \tilde{\alpha}_s$, if $c_p - \pi_a\alpha_a < c_a < 1 - \pi_a\alpha_a - (1-c_p)\sqrt{1-\pi_a\alpha_a}$ and $\pi_i\alpha_i < \min\left[\frac{c_p\pi_a\alpha_a}{1+c_p-c_a}, \frac{c_p}{1+c_p}\right]$, when either* *(i)* $c_a < \min\left[\pi_a\alpha_a - c_p, c_p(1-\pi_a\alpha_a), \frac{\pi_i\alpha_i(2\pi_a\alpha_a-1)-1}{2\pi_a\alpha_a+\pi_i\alpha_i-3} - \pi_a\alpha_a\right]$, *or*

(ii) $|\pi_a\alpha_a - c_p| < c_a < \min\left[c_p(1-\pi_a\alpha_a), \frac{(1-\pi_a\alpha_a)(\pi_i\alpha_i-2c_p+1)}{-4\pi_a\alpha_a-\pi_i\alpha_i+5}\right]$,

the vendor prices patching rights such that both p^* and $\delta^* p^*$ are higher than the common price, p^*_{SQ}, when patching rights are endowed to all users.

Not only does the endowment of patching rights lead to excessive security risk due to poor patching behavior, it also fails to reflect the value of security provision being offered by the vendor. Vendors who create better, more secure solutions for their customers should be able to harvest some of that value creation via increased prices. Proposition 3 highlights this important point by characterizing broad regions where the vendor increases the price of both options above the single price offered in the case of the status quo. This occurs for a lower level of automated patching costs (c_a), and the reason both prices increase is twofold. First, users who prefer to retain patching rights are willing to pay more for smaller unpatched populations (i.e., reduced security risk) and control over their own patching process. Second, the value associated with cost-efficient and more secure, automated patching options is more readily harvested when users of this option are ungrouped from users who choose not to patch under the status quo. The pricing of patching rights helps to enable this separation. Thus, when a vendor differentiates in this manner based on "rights," he can simultaneously increase prices, encourage more secure behaviors, and generate higher profits. The outcome under this business strategy is noteworthy because it is starkly different than one in which security protections are sold and those who opt out are both unprotected and cause a larger security externality.

Proposition 3 suggests that usage may become more restricted with priced patching rights. Moreover, it is unclear how specific costs associated with security would be affected as consumers strategically adapt their usage and protection decisions. Proposition 2 demonstrates that priced patching rights can reduce the size of the unpatched population relative to the status quo which in turn implies the risk associated with security attacks decreases. However, the magnitude of losses associated with these attacks critically depends on who actually bears them as they are valuation-dependent and consumers' equilibrium strategies will shift when patching rights are priced. We denote the expected losses associated with security attacks stemming from the unpatched population $u(\sigma^*)$ with

$$SL \triangleq \int_{\mathcal{V}} \mathbb{1}_{\{\sigma^*(v)=(B,NP)\}} \left(\pi_s \alpha_s u(\sigma^*) + \pi_i \alpha_i\right) v dv \,. \tag{2.30}$$

In a similar fashion, we denote the expected losses associated with configuration and instability of automated patching with

$$AL \triangleq \int_{\mathcal{V}} \mathbb{1}_{\{\sigma^*(v)=(B,AP)\}} c_a + \pi_a \alpha_a v dv \,, \tag{2.31}$$

and the total costs associated with standard patching with

$$PL \triangleq \int_{\mathcal{V}} \mathbb{1}_{\{\sigma^*(v)=(B,P)\}} c_p dv \,. \tag{2.32}$$

The net impact of consumers changing their patching strategies (standard patching, remaining unpatched, electing for automated patching) on these security-related costs is unclear. In order to examine these concerns in aggregate, we also define total security-related costs as the sum of these three components:

$$L \triangleq SL + AL + PL \,, \tag{2.33}$$

in which case social welfare can be expressed as

$$W \triangleq \int_{\mathcal{V}} \mathbb{1}_{\{\sigma^*(v)\in\{(B,NP),(B,AP),(B,P)\}\}} v dv - L. \tag{2.34}$$

In the following proposition, we establish that when automated patching costs are not too large, the pricing of patching rights can in totality have a negative effect on social welfare. This result is interesting in that both losses associated with security attacks and total costs associated with standard patching can be shown to decrease when patching rights are priced, and yet priced patching rights can still be detrimental from a welfare perspective.

Proposition 4 *There exists a bound $\tilde{\alpha}_s$ such that when $\alpha_s > \tilde{\alpha}_s$, if*

$c_p - \pi_a\alpha_a < c_a < 1 - \pi_a\alpha_a - (1-c_p)\sqrt{1-\pi_a\alpha_a}$ *and* $\pi_i\alpha_i < \min\left[\frac{c_p\pi_a\alpha_a}{1+c_p-c_a}, \frac{c_p}{1+c_p}\right]$, *then priced patching rights can either decrease or increase security attack losses, but leads to a small decrease in social welfare. Technically,* $PL_P < PL_{SQ}$, $AL_P > AL_{SQ}$, $W_P < W_{SQ}$ *and*

(i) if $c_a < \min\left[\pi_a\alpha_a - c_p, c_p(1-\pi_a\alpha_a)\right]$, *and*

$$\frac{4c_p^2\pi_a\alpha_a}{-c_a + c_p + \pi_a\alpha_a} - \frac{(c_a(2-\pi_a\alpha_a) + \pi_a\alpha_a(1-\pi_a\alpha_a))^2}{(1+c_a-\pi_a\alpha_a)(1-\pi_a\alpha_a)} > 0\,, \tag{2.35}$$

then $SL_P > SL_{SQ}$ *;*

(ii) otherwise, $SL_P \leq SL_{SQ}$ *.*

The parameter region in Proposition 4 corresponds to c_a being relatively lower and satisfying the conditions of Lemmas 4 (first part) and 5. Recalling that under priced patching rights, the consumer market structure can be characterized by either $0 < v_b < v_a < v_p < 1$ or $0 < v_a < v_b < v_p < 1$ in equilibrium, we first examine the former case where the consumer market structure matches the characterization under the status quo. By pricing patching rights, the vendor will induce an expansion of the consumer segment that elects for automated patching on both sides. That is, some unpatched users as well as some standard patching users under the status quo will now choose automated patching. Additionally, some unpatched users are now out of the market due to the increase in the price p^* associated with retained patching rights (technically, v_b increases). Therefore, losses associated with unpatched security attacks and costs associated with standard patching are both lower in comparison to the status quo, i.e., $SL_P < SL_{SQ}$ and $PL_P < PL_{SQ}$.

However, the expansion of the consumer segment choosing (B, AP) turns out to be costly. In particular, because consumers have the opportunity to relinquish patching rights to save the premium $(1-\delta^*)p^*$, the consumers that make up the expansion of this segment may incur greater security investments and system instability losses in order to avoid paying this premium. For example, at the higher end of the valuation space, a consumer may have incurred only c_p under status quo pricing but when incentivized to shift to automated patching because of the discount, she now incurs a security cost of $c_a + \pi_a \alpha_a v$ which is valuation-dependent and can exceed c_p. A similar increase in costs can arise at the lower end as consumers shift from losses associated with security attacks to investments and instability losses associated with automated patching. Proposition 4 formally establishes that the decrease in usage and increased aggregate costs incurred related to automated patching ultimately outweigh the reduction in security attack losses and standard patching costs from a welfare perspective. From a software vendor's perspective, the ability to market their product offerings as geared to reduce security risk and attack losses while increasing profits is enticing, and having awareness of the impact on welfare can help shape these initiatives. Encouragingly, we also characterize several regions where social welfare is positively impacted by priced patching rights as well in Proposition 5 and the discussion of Figure 2.3.

But first turning attention to the case where the vendor's pricing behavior restructures the consumer market to $0 < v_a < v_b < v_p < 1$, we find the outcome is similar but has some nuanced differences. In this case, the consumers whose equilibrium strategy is to retain patching rights but not patch (users with valuations between v_b and v_p) have higher valuations than those preferring this strategy under the status quo case. Thus, even though the size of the unpatched population, $u(\sigma^*)$, decreases under priced patching rights, the higher valuations of the consumers exhibiting the risky, unpatched behavior can result in them incurring higher losses when bearing security attacks. It hinges on whether $u(\sigma^*)$ decreases sufficiently to offset the higher valuations of the risky population. In part (i) of Proposition 4, the conditions required for the restructured consumer market as laid out in Lemma 5 appear. Further, (2.35) provides the condition whereupon security attack losses are, in fact, higher under priced patching rights, despite the reduction in unpatched usage. One can think of this outcome as characterized by fewer attacks but on higher value targets leading to greater losses in equilibrium. This condition tends to be satisfied as the likelihood of automated patch instability increases which provides more incentive for consumers to remain unpatched instead. With the potential of security attack losses to also increase, welfare is even further suppressed compared to the status quo.

Next, we study the case where automated patching costs are at a level large enough that an automated patching segment is absent under the status quo but small enough that this segment arises when patching rights are priced (see the second part of Lemma 4 and Lemma 5).

Proposition 5 *There exists a bound $\tilde{\alpha}_s$ such that when $\alpha_s > \tilde{\alpha}_s$, if $1-\pi_a\alpha_a-(1-c_p)\sqrt{1-\pi_a\alpha_a} \leq c_a < c_p(1-\pi_a\alpha_a)$ and $\pi_i\alpha_i < \min\left[\frac{c_p\pi_a\alpha_a}{1+c_p-c_a}, \frac{c_p}{1+c_p}\right]$, then priced patching rights leads to decreased security attack losses and an increase in social welfare. Technically, $SL_P < SL_{SQ}$, $PL_P < PL_{SQ}$, $AL_P > AL_{SQ}$, and $W_P > W_{SQ}$.*

Proposition 5 examines a higher cost of automated patching in which case, under the status quo, the consumer market is characterized by $0 < v_b < v_p < 1$ (Lemma 4). One can think of this as a context where automated patching technology is inferior and users elect not to use it in equilibrium. This behavior can result in a large unpatched population and substantial security risk, causing many potential consumers to prefer not to be users of the product. Thus, the value of a priced patching rights policy can be lucrative if it provides sufficient incentives to reduce unpatched behavior and expand usage. Under optimally priced patching rights, users who were unpatched under the status quo are incentivized by a discount to use the automated patching option. In that automated patching is an inferior technology in this context, these users may bear greater costs and instability losses associated with automated patching in exchange for receiving this discount. These greater costs are detrimental to welfare.

On the other hand, because the unpatched population is significantly reduced, losses associated with security attacks are lower ($SL_P < SL_{SQ}$). Moreover, because the vendor makes the automated patching available at a discount, usage in the market for the software expands. In fact, when the loss factor on automated patching technology ($\pi_a\alpha_a$) is at the high end of the focal region, both the price of the product with patching rights (p^*) and without ($\delta^* p^*$) can be lower than the price under the status quo (p^*_{SQ}). Thus, usage in the market can expand substantially, and the additional surplus generated from these consumers who were non-users under the status quo helps to benefit welfare. Proposition 5 establishes

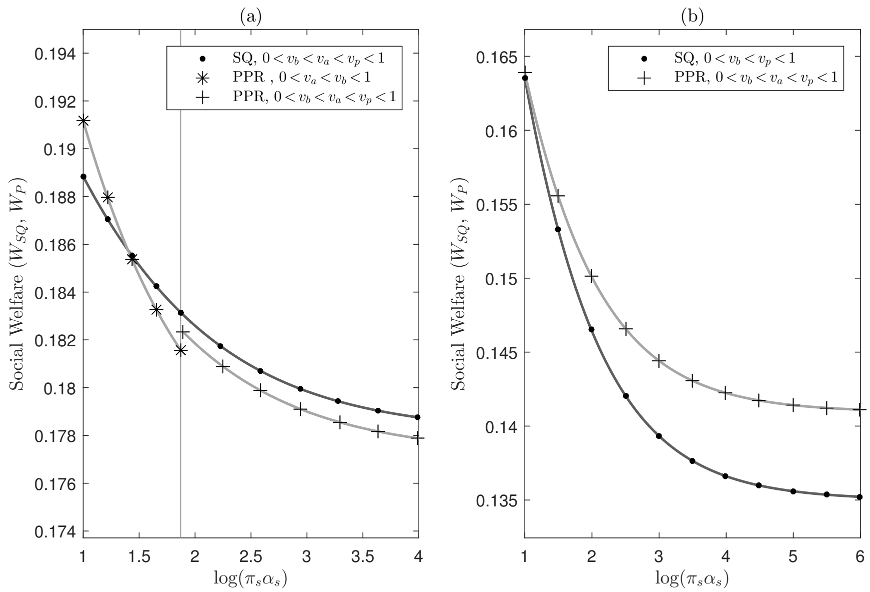


Figure 2.3: Impact of a Priced Patching Rights (PPR) Strategy on Social Welfare for Varying Effective Security Risk, Compared to the Status Quo (SQ).
In panel (a), $c_a = 0.1$, and in panel (b), $c_a = 0.2$. The common parameter values are $\alpha_i = 0.2$, $\pi_i = 0.1$, $\alpha_a = 3.5$, $\pi_a = 0.1$, and $c_p = 0.4$.

that the net effect of these factors is positive, and priced patching rights can have a positive influence on social welfare.

While the pricing of patching rights is quite effective at reducing unpatched populations and losses associated with security attacks, Propositions 4 and 5 demonstrate that its impact of welfare can be mixed when security risk is becomes large. In Figure 2.3, we examine the impact on welfare as the security loss factor becomes lower under smaller automated patching costs. As can be seen in panel (a) of Figure 2.3, a priced patching rights strategy can also be beneficial to both vendor profits and social welfare relative to the status quo strategy as $\pi_s\alpha_s$ decreases. Under the status quo, the consumer market equilibrium is characterized by $0 < v_b < v_a < v_p < 1$ throughout panel (a). However, two different consumer market structures are represented under priced patching rights. To the right of the discontinuity, the characterization is the same, while to the left of the discontinuity (hence lower $\pi_s\alpha_s$), the structure becomes $0 < v_a < v_b < 1$. In other words, as $\pi_s\alpha_s$ decreases, patching rights are priced in a way that significantly restructures the equilibrium consumer strategies in comparison to the status quo; consumers with high valuations retain patching rights but choose to remain unpatched and consumers with lower valuations forgo rights and either shift to automated patching or exit the market.

What is most interesting about this reshuffling is that the consumers who were causing the security risk no longer do so and, as a result, the consumers who were incurring standard patching costs to shield themselves from the security risk also no longer need to do so. Organizations have long carried a large financial burden associated with the rigorous patching processes that they are forced to employ to limit risk. If the ecosystem becomes safer, these organizations could reduce these investments while keeping that risk exposure limited. This is the generalizable insight brought to light here – a priced patching rights strategy not only reduces security risk, it enables high valuation users to avoid incurring typically large patching costs. The net result of priced patching rights is that total costs related to automated patching increase (the automated patching population expands), costs associated with standard patching disappear (patching burden is relieved), and security attack losses stemming from unpatched usage is reduced (significant reduction in the size of the unpatched population). As a result, social welfare increases under priced patching rights in comparison to the status quo as the security loss factor decreases out of region covered by Proposition 4.

Panel (b) of Figure 2.3 illustrates the finding from Proposition 5 that social welfare increases under priced patching rights for a high security loss factor. Moreover, the benefits to welfare also extend to a lower range of security losses which is depicted as well. In summary, the pricing of patching rights presents an opportunity for vendors of proprietary software to not only improve profits, but also to improve welfare by decreasing the magnitude of the externality generated by unpatched usage, even to the degree that the patching burden can be relieved.

2.5 Discussion and Concluding Remarks

We find that a PPR strategy is a very effective way to improve software security. We characterize the optimal manner in which to price these rights and demonstrate that it leads to a significant increase in the profits of firms producing proprietary software. Moreover, the size of unpatched users decreases in equilibrium under PPR which, in turn, tends to lead to a decrease in losses associated with security attacks on systems running the software. In some cases, the firm may raise both the price of the premium version with patching rights and the price of the "discounted" version without patching rights relative to the optimal price offered under the status quo. This demonstrates how a software firm can extract value from a combination of improved automated patching technology and a pricing strategy that incentivizes better security outcomes that are valued by consumers. We establish that the pricing of patching rights can negatively affect social welfare when usage in the market significantly contracts, but in many cases welfare improves as a result of the lower resulting security risk.

Our study is a simplification of a software vendor's versioning strategy which, more practically, can involve managing numerous quality-differentiated versions, bundled security/feature updates, interim update versus major release cycles, and planned obsolescence. The intent of our simplification is to put a spotlight on the value of PPR as a tool that provides incentives for lower valuation users to engage in better behaviors that help reduce the effective security externality. Even in the complex settings that exist in the software industry, the ideas and insights stemming from our work can improve software outcomes if, however implemented, lower valuation users ultimately select versions that remain up to date with security patches. In the software industry, vendors have taken greater actions to ensure that versions of their products remain up to date. We have observed the offering of automatic updates, configuring of default options to turn automatic updates on, and even versions of software which forcibly update. Our work helps clarify the impact of actions that aim to incentivize lower valuation users to forgo patching rights.

Despite the benefits we characterize, there are many frictions associated with the essence of a PPR strategy which can cause vendors to resist adoption. First, by forcibly updating users who choose to forgo patching rights for discounted versions of the software, the vendor is exposing these users to system instability risk. While we capture these losses in the trade-off evaluated by users in our model, vendors may shy away from additional exposure to liability. Second, whenever there is a patch that leads to severe instability, vendors will likely receive backlash from consumers and incur damage to their reputation. In this regard, any vendor strategy that eliminates patching rights from a market segment may go hand-in-hand with investments in patch quality. In that many software vendors have been providing the option of automatic updates for years now, they have made significant progress on patch stability. In order for more vendors to consider adoption of a PPR policy (or practical variants), assurance of patch stability will be critical since the policy leaves the targeted consumer tier with little recourse.

2.5.1 Alternative Models and Application Domains of PPR

The manner in which software is licensed and updated is constantly changing due to technology disruptions and the developing needs of consumers. While there is some diversity in observed licensing strategies in the software industry, in the current work we focus on an important and large class of software products where patches are predominantly made freely available to users who then make decisions whether or not to install them. This class includes versions of software products from an extensive list of top vendors such as Microsoft, Oracle, Symantec, VMware, and IBM. Examples of specific software products that can be installed on-premises and maintained by users include Microsoft Windows, Microsoft Office, Microsoft SQL Server, Oracle Database, Oracle WebLogic Server, Norton Antivirus, Symantec Endpoint Protection, VMware Workstation, and IBM WebSphere. In this section, we discuss both alternative models to PPR and other potential application domains of PPR, highlighting in the latter whether PPR-type policies are likely to be effective in the domain of interest.

Some producers make software source code available for free and build revenue models around service and support (e.g., Red Hat Enterprise Linux, Elasticsearch, and Oracle Java). In this open-source domain, when software is made available for free, developers have no obligation to provide patches. One approach that has been observed is the offering of patches only to paying customers who have contracted for support. The first alternative model that we discuss is charging for patches. Oracle has done this by offering the patches only for paying customers of its freely available open-source Java software (Krill 2015). We examine how this strategy compares to PPR. First, it is important to understand what happens when an OSS developer does not provide patches to its users. In such a setting, a consumer's options can include: (i) use the software for free and be unpatched, risking security attacks, (ii) not use the software, and (iii) leverage the open-source nature of the software to self-patch.[11] Because the software has zero price, if the effective security risk is low, *all* consumers will use the software and obtain positive surplus in equilibrium. On the other hand, if the effective security risk exceeds a certain bound, then users can only enter until the size of the unpatched population reaches a critical size upon which dependent risk wipes out all surplus derived from any unpatched use of the software.

In such a context, the offering of a patch alone will significantly improve welfare. Patches enable more high valuation users to shield themselves from risk, which in turn can reduce the security externality that decays the surplus of users who choose to remain unpatched. Suppose a vendor offers patches only to those who agree to pay for them as part of a support package. This provides users the additional usage option of pay for patches. Examining conditions analogous to Section 2.4.2 (i.e., low $\pi_i\alpha_i$ and high $\pi_s\alpha_s$), the equilibrium consumer market structure will be characterized by both patched and unpatched usage. Because more high valuation consumers can now patch and protect themselves from attack, social surplus will increase in equilibrium in comparison to the case where it is mostly wiped out in the presence of high risk and patch unavailability.

One can compare our PPR policy to this alternative model where an OSS developer prices the patch itself. A PPR policy will tend to dominate this alternative model both in terms of profitability and in terms of welfare. It is straight-

[11] Only a select group of users might have the capability to self-patch.

forward to see why it is more profitable. Because a PPR policy charges premia for both unpatched and standard patched usage and even a positive price for automatic patched usage, it generates a lot of direct revenues. The alternative model requires all the revenues to come from higher valuation users who need to be willing to pay for access to patches. By giving the software away for free, one's revenue model typically hinges on the creation of positive network effects and leveraging those to create higher willingness-to-pay within the higher value segment. However, this alternative model has payment tied to security patches, which means that the higher value segment will only be willing to pay for patches if the security risk is large. This suggests that risk stemming from expanded usage is necessarily required, but that lies in stark contrast to any sort of positive network effects strategy. These opposing effects handicap this model in comparison to the PPR policy we advocate.

In terms of social welfare, this alternative model suffers relative disadvantages dependent upon the region of concern. If automated costs are comparable to standard patching costs, a PPR strategy expands overall usage relative to the status quo by incentivizing more to select automated patching. This usage expansion is also much larger than can be expected under the alternative model because in that case users have fewer options and cannot readily shield risk without paying a premium – this lies in contrast to the PPR strategy where the unpatched users are provided discounts instead to reduce risk. If automated costs are large and it becomes difficult to incentivize automated patching, then the alternative model can in fact generate greater usage than under our policy. However, expanded usage of the free product increases the externality to the point where unpatched users gain no surplus. Therefore, even in this case, the PPR strategy achieves greater welfare with a smaller user segment, each of whom achieves positive surplus.

Even for OSS providers who are not charging for patches, a PPR policy can be a valuable option. In particular, OSS providers who are already employing other monetization strategies can easily implement the essence of a PPR policy and benefit via improved security and revenues. On the other hand, OSS providers who are not currently differentiating their offerings and whose mission is perhaps more squarely centered on public good are not likely to be good candidates for PPR. Such providers may find both the concept of giving up rights and charging for software to be inconsistent with their mission.

Next, there are many producers who offer subscriptions for access to SaaS (e.g., Salesforce CRM, Workday, and NetSuite). In this domain, a vendor manages the patching process directly so that consumers no longer make a decision about whether or not to patch. However, this certainly comes at a cost to the provider, which ultimately gets reflected in the price consumers now pay. SaaS software tends to run on a fairly limited set of servers under the control of the vendor, whereas in the on-premises model the number of systems running software tends to be much larger with decision-making rights being distributed. In the case of on-premises software studied in our model, it is often the case that patchable vulnerabilities are exploited. This occurs because malicious individuals study the vulnerability that has been addressed in the patch, and then exploit the same vulnerability hoping to successfully attack some subset of users who remain in an unpatched state. In the case of SaaS, the approach to finding vulnerabilities differs significantly as both the software as well as patches to the software remain internal to the vendor. Malicious folks must find vulnerabilities in the interface to the SaaS product itself. On the consumer side, it can be the case that companies (particularly SMB ones) fully commit to the cloud as part of their IT strategies, which may introduce correlated risks on cloud platforms making it difficult to analyze a PPR strategy in an context where a vendor additionally offers both on-premises and SaaS alternatives. Because of the many differences in characteristics that surface when comparing SaaS to on-premises software, a PPR policy does not fit the SaaS domain very well. However, studying the role of a PPR strategy under the additional complexity of mixed offerings (SaaS and on-premises) is potentially a fruitful direction for future research because of the distinct security risks associated with the different licensing models (August et al. 2014)

Software producers have long shielded themselves from liability using well-crafted license agreements. However, the increasing breadth of software use in riskier operating environments including the critical infrastructure, biomedical products, and automobiles, comes with increased exposure to strict liability. In particular, liability concerns come to the forefront in software contexts where system failure can lead to safety or health hazards. For example, researchers have demonstrated the ability to hack into automotive systems and hijack control over brakes and steering from drivers; vulnerabilities such as these can be utilized to cause physical harm to citizens (Greenberg 2015). In application domains with this property, a PPR policy may be inappropriate. In our model, these settings would be characterized as having a large $\pi_a \alpha_a$ parameter, in which case a PPR strategy is shown to be weakly dominated. Intuitively, manufacturers would likely not agree to have embedded software on products such as pacemakers automatically updated. In the presence of strict liability, assurance of quality takes precendence.

Many issues tackled in this paper will become increasingly salient as the Internet of things (IoT) comes of age. With IoT, the explosion of interconnected devices will be accompanied by both increased vulnerabilities and increased threats as malicious actors evolve to exploit new possibilities. The standard patching processes employed by organizations will face significant challenges with the scale of device growth and their complex interactions. IoT's scale will require a greater level of automation in patch management. The extent and type of human interaction taking place with software that drives servers, laptops, tablets and phones is fundamentally different than that with the embedded software that drives IoT devices. This, in turn, may necessitate greater sophistication with the automated management of IoT devices. In this landscape, IoT presents as both a challenge and opportunity for security interventions like a PPR policy. For PPR to be effective in this domain, automated patching technology must first improve and attain a service level where minimal failure rates are observed. Until then, it may be premature to consider a PPR policy for IoT devices.

2.5.2 Concluding Remarks

Standard patching costs (c_p) may increase with factors such as the relative size of a user firm, for instance, depending on the number of servers on which they run the software product. A primary driver is that a larger number of users require a higher number of servers likely in different environments and configurations on which to patch the software, driving up the patching costs proportional to the number of servers. Our results however continue to hold with consideration of this dimension: In our model, each host can be considered to be a single server and a large corporation can be thought to have a number of different servers. Our analysis is unaffected provided each decision maker owns at most countably many hosts. It may also be the case that two organizations conducting the same tasks associated with a rigorous, standard patching process would have some variation in the costs incurred. For example, there may be some variation around one thousand dollars per server. Having some variation in this regard will not qualitatively change any of our results for two reasons. First, given the nature of the tasks being performed, there will always be a primary, valuation-independent component of these costs which is what we currently capture. Second, inclusion of a valuation-dependent component will ultimately be swamped by the valuation-dependent losses that are currently present in the model. In particular, in our model, both losses due to system instability and losses due to security attacks are valuation-dependent. Importantly, both of these losses are much more strongly associated with valuation than the residual component of patching costs. This limits the impact that this residual can have on the equilibrium thresholds that we characterize.

We focus on the case where a software vendor releases and immediately applies automatic updates to all systems that have opted in to this form of patching. One interesting extension would be to explore the dynamics of automated patch deployment. In particular, some vendors may prefer to slowly roll out an automated update to subsets of systems that have opted-in and subsequently ramp up roll out as the vendor builds confidence in the quality of its patch (as measured by monitoring of systems that have been updated). Another related extension would be to explore how a vendor's investment in patch quality interacts with this dynamic deployment strategy.

Our study constructs a model of a monopolist software vendor to study PPR. In practice, essentially no sector is monopolistic and there is some competition in any industry and market. However, as long as the firms in the market have some market power (which is true in almost every industry and market, and certainly for the software industry), a monopolistic model captures most of the insights that would come out from the intended arguments in an oligopoly model of the same situation in a much clearer and transparent way. In general, increased competition will negatively affect firm profits and likewise the profitability of PPR, and studying competition-related questions would be an interesting direction for future research.

Under the framework of a PPR policy, the consumers who pay a premium to retain patching rights need no additional monitoring by the proprietary vendor or OSS provider to examine patching status. However, consumers who benefit financially in exchange for these rights must have their systems automatically updated per the contract. This requires a careful implementation that entails the monitoring of systems. First, the updating of systems need not be in full control by the vendor nor instantaneous to derive the benefits of this policy. For example, consumers can be given a 24-hour window to apply patches before they are forcefully installed. This gives users some leeway operationally. Second, vendors can modify the implementation of PPR to the environment in which they reside. For example, mobile application developers could give a discount to users willing to automatically update their software, and even give a slightly greater discount to those who are willing to receive those updates immediately through their mobile data plan. Lastly, a user can always disconnect her system from the Internet to avoid the deployment of automated security updates. In this case, although the patches are not installed, the externality imposed by this unpatched system would also be partially reduced.

2.6 Appendix

2.6.1 High $\pi_i\alpha_i$

Status Quo

Lemma A.1 *When $\pi_i\alpha_i > 1$, under the status quo, i.e., $\delta = 1$, the complete threshold characterization of the consumer market equilibrium is as follows:*

(I) $(0 < v_a < 1)$*, where* $v_a = \frac{p+c_a}{1-\pi_a\alpha_a}$*:*

- *(A)* $p + c_a + \pi_a\alpha_a < 1$
- *(B)* $c_p \geq c_a + \pi_a\alpha_a$
- *(C)* $c_p \geq \frac{c_a+p\pi_a\alpha_a}{1-\pi_a\alpha_a}$

(II) $(0 < v_a < v_p < 1)$*, where* $v_a = \frac{p+c_a}{1-\pi_a\alpha_a}$ *and* $v_p = \frac{c_p-c_a}{\pi_a\alpha_a}$*:*

- *(A)* $c_p < c_a + \pi_a\alpha_a$
- *(B)* $c_p > \frac{c_a+p\pi_a\alpha_a}{1-\pi_a\alpha_a}$

(III) $(0 < v_p < 1)$*, where* $v_p = p + c_p$*:*

- *(A)* $c_p + p < 1$
- *(B)* $c_p \leq \frac{c_a+p\pi_a\alpha_a}{1-\pi_a\alpha_a}$

Proof of Lemma A.1: This is a sub-case in the proof of Lemma A.2, by setting $\delta = 1$. $\square$

Proof of Lemma 2: We prove that if $\pi_i\alpha_i > 1$ and $\delta = 1$ (i.e., when patching rights are not priced), then if $c_p - \pi_a\alpha_a < c_a \leq 1 - \pi_a\alpha_a - (1-c_p)\sqrt{1-\pi_a\alpha_a}$, we have that

$$p^* = \frac{1 - \pi_a\alpha_a - c_a}{2}\,, \tag{2.36}$$

and σ^* is characterized by $0 < v_a < v_p < 1$ such that the lower tier of users prefers automated patching. On the other hand, if $c_a > 1 - \pi_a\alpha_a - (1-c_p)\sqrt{1-\pi_a\alpha_a}$, then

$$p^* = \frac{1-c_p}{2}\,, \tag{2.37}$$

and σ^* is characterized by $0 < v_p < 1$ such that there is no user of automated patching in equilibrium.

Suppose $0 < v_a < 1$ is induced. Then the profit function is $\Pi_I(p) = p(1-v_a)$. Using Lemma A.1, we have that $v_a = \frac{p+c_a}{1-\pi_a\alpha_a}$. The optimal price is found to be $p_1^* = \frac{1}{2}(1-c_a-\pi_a\alpha_a)$ with the corresponding profit $\Pi_I^* = \frac{(1-c_a-\pi_a\alpha_a)^2}{4(1-\pi_a\alpha_a)}$.

Similarly, suppose that $0 < v_a < v_p < 1$ is induced. Then the profit function is $\Pi_{II}(p) = p(1-v_a)$. Using Lemma A.1, we have that $v_a = \frac{p+c_a}{1-\pi_a\alpha_a}$. Again, the optimal price is found to be $p_2^* = \frac{1}{2}(1-c_a-\pi_a\alpha_a)$ with the corresponding profit $\Pi_{II}^* = \frac{(1-c_a-\pi_a\alpha_a)^2}{4(1-\pi_a\alpha_a)}$.

Lastly, suppose that $0 < v_p < 1$ is induced. Then the profit function is $\Pi_{III}(p) = p(1-v_p)$. Using Lemma A.1, we have that $v_p = c_p + p$. Now, the optimal price is found to be $p_3^* = \frac{1-c_p}{2}$ with the corresponding profit $\Pi_{III}^* = \frac{(1-c_p)^2}{4}$.

We next find conditions under which the maximizing price for each case indeed induces that market structure. For $0 < v_a < 1$, we need the set of conditions for Case (I) in Lemma A.1 to hold for p_1^*. To satisfy the first condition, we need $p + c_a + \pi_a\alpha_a < 1$ for $p = p_1^* = \frac{1}{2}(1-c_a-\pi_a\alpha_a)$. This simplifies to $c_a + \pi_a\alpha_a < 1$, which is a preliminary model assumption. Then we need $c_p \geq \frac{c_a+p\pi_a\alpha_a}{1-\pi_a\alpha_a}$ to hold for $p = p_1^*$ as well, which simplifies to $c_a \leq \frac{(2c_p-\pi_a\alpha_a)(1-\pi_a\alpha_a)}{2-\pi_a\alpha_a}$. We also needed the condition $c_a \leq c_p - \pi_a\alpha_a$ for this case to hold. Since $\frac{(2c_p-\pi_a\alpha_a)(1-\pi_a\alpha_a)}{2-\pi_a\alpha_a} > c_p - \pi_a\alpha_a$ follows from $0 < c_p < 1$ and $0 < \pi_a\alpha_a < 1$, then the condition under which p_1^* would induce $0 < v_a < 1$ is $c_a \leq c_p - \pi_a\alpha_a$.

Similarly, for Case (II), the condition under which p_2^* would induce $0 < v_a < v_p < 1$ is $c_p - \pi_a\alpha_a < c_a < \frac{(2c_p-\pi_a\alpha_a)(1-\pi_a\alpha_a)}{2-\pi_a\alpha_a}$. And lastly, for Case (III), the condition under which p_3^* would induce $0 < v_p < 1$ is $c_a \geq c_p - \frac{1}{2}(1+c_p)\pi_a\alpha_a$. Note that $c_p - \frac{1}{2}(1+c_p)\pi_a\alpha_a < \frac{(2c_p-\pi_a\alpha_a)(1-\pi_a\alpha_a)}{2-\pi_a\alpha_a}$, so that in $c_p - \frac{1}{2}(1+c_p)\pi_a\alpha_a < c_a < \frac{(2c_p-\pi_a\alpha_a)(1-\pi_a\alpha_a)}{2-\pi_a\alpha_a}$, we'll need to compare Π_{II}^* and Π_{III}^*.

Next, we find the conditions under which the maximal profits of each case dominate the other cases. In particular, when $c_a < c_p - \frac{1}{2}(1+c_p)\pi_a\alpha_a$, then $\Pi_I^* = \Pi_{II}^* > \Pi_{III}^*$. Since $c_p - \pi_a\alpha_a < c_p - \frac{1}{2}(1+c_p)\pi_a\alpha_a$ from $c_p < 1$, this implies that $0 < v_a < 1$ will be the resulting consumer market structure for $c_a \leq c_p - \pi_a\alpha_a$. Also, $0 < v_a < v_p < 1$ will be the resulting consumer market structure for $c_p - \pi_a\alpha_a < c_a < c_p - \frac{1}{2}(1+c_p)\pi_a\alpha_a$.

Also, for $c_a > \frac{(2c_p - \pi_a\alpha_a)(1-\pi_a\alpha_a)}{2-\pi_a\alpha_a}$, we have that $\Pi_{III}^* > \Pi_{II}^*$, so that $0 < v_p < 1$ will be the resulting market structure when $c_a > \frac{(2c_p - \pi_a\alpha_a)(1-\pi_a\alpha_a)}{2-\pi_a\alpha_a}$.

In between $c_p - \frac{1}{2}(1+c_p)\pi_a\alpha_a < c_a < \frac{(2c_p - \pi_a\alpha_a)(1-\pi_a\alpha_a)}{2-\pi_a\alpha_a}$, both Π_{II}^* and Π_{III}^* can be induced. In this region of the parameter space, we find the conditions under which one profit dominates the other. Comparing the profits, we find that $\Pi_{II}^* \geq \Pi_{III}^*$ when $c_a \leq 1 - \pi_a\alpha_a - (1-c_p)\sqrt{(1-\pi_a\alpha_a)}$. Then we note that $c_p - \frac{1}{2}(1+c_p)\pi_a\alpha_a < 1 - \pi_a\alpha_a - (1-c_p)\sqrt{(1-\pi_a\alpha_a)} < \frac{(2c_p - \pi_a\alpha_a)(1-\pi_a\alpha_a)}{2-\pi_a\alpha_a}$ always holds, so that the resulting market structure when $c_p - \pi_a\alpha_a < c_a \leq 1 - \pi_a\alpha_a - (1-c_p)\sqrt{1-\pi_a\alpha_a}$ is $0 < v_a < v_p < 1$, and the resulting market structure when $c_a > 1 - \pi_a\alpha_a - (1-c_p)\sqrt{1-\pi_a\alpha_a}$ is $0 < v_p < 1$. □

Pricing Patching Rights

Lemma A.2 *When $\pi_i\alpha_i > 1$, under priced patching rights, the complete threshold characterization of the consumer market equilibrium is as follows:*

(I) $(0 < v_a < 1)$, *where* $v_a = \frac{\delta p + c_a}{1-\pi_a\alpha_a}$*:*

(A) $\delta p + c_a + \pi_a\alpha_a < 1$

(B) $c_p + (1-\delta)p \geq c_a + \pi_a\alpha_a$

(C) $c_p \geq \frac{c_a + p(\pi_a\alpha_a - (1-\delta))}{1-\pi_a\alpha_a}$

(II) $(0 < v_a < v_p < 1)$, *where* $v_a = \frac{\delta p + c_a}{1-\pi_a\alpha_a}$ *and* $v_p = \frac{(1-\delta)p + c_p - c_a}{\pi_a\alpha_a}$*:*

(A) $c_p + (1-\delta)p < c_a + \pi_a\alpha_a$

(B) $c_p > \frac{c_a + p(\pi_a\alpha_a - (1-\delta))}{1-\pi_a\alpha_a}$

(III) $(0 < v_p < 1)$, *where* $v_p = p + c_p$*:*

(A) $c_p + p < 1$

(B) $c_p \leq \frac{c_a + p(\pi_a\alpha_a - (1-\delta))}{1-\pi_a\alpha_a}$

Proof of Lemma A.2: First, we establish the general threshold-type equilibrium structure. The proof of this is a sub-case of the argument in Lemma A.4, with the size of the unpatched user population $u = 0$ since $\pi_i\alpha_i > 1$. This establishes the threshold-type consumer market equilibrium structure.

Next, we characterize in more detail each outcome that can arise in equilibrium, as well as the corresponding parameter regions. For Case (I), in which all consumers who purchase choose the automated patching option, i.e., $0 < v_a < 1$, based on the threshold-type equilibrium structure, we have $u = 0$. For this market structure to be an equilibrium, we need $v_a > 0$, $v_a < 1$, the consumer $v = 1$ weakly preferring (B, AP) over (B, P), and the consumer $v = v_a$ weakly preferring (NB, NP) over (B, P).

The condition $v_a > 0$ is satisfied by our assumption that $\pi_a\alpha_a < 1$, since $v_a = \frac{\delta p + c_a}{1-\pi_a\alpha_a}$. Then for $v_a < 1$, we need $\delta p + c_a + \pi_a\alpha_a < 1$. For $v = 1$ to weakly prefer (B, AP) over (B, P), it needs to be the case that $v - \delta p - c_a - \pi_a\alpha_a v \geq v - p - c_p$ for $v = 1$. This simplifies to $c_p + (1-\delta)p \geq c_a + \pi_a\alpha_a$. For $v = v_a$ to weakly prefer (NB, NP) over (B, P), it needs to be the case that $0 \geq v - p - c_p$ for $v = v_a = \frac{\delta p + c_a}{1-\pi_a\alpha_a}$. This simplifies to $c_p \geq \frac{c_a + p(\pi_a\alpha_a - (1-\delta))}{1-\pi_a\alpha_a}$.

Next, for case (II), in which the top tier purchases (B, P) but the lower tier of consumers purchase (B, AP), i.e., $0 < v_a < v_p < 1$, we have $v_a = \frac{\delta p + c_a}{1-\pi_a\alpha_a}$ and $v_p = \frac{(1-\delta)p + c_p - c_a}{\pi_a\alpha_a}$. Following the same steps as before, we find the corresponding conditions for which case (II) arises. For this case to arise, we need $v_a > 0$, $v_p > v_a$, and $v_p < 1$. Again, $v_a > 0$ is satisfied under $\pi_a\alpha_a < 1$, one of the preliminary assumptions of the model. To have $v_p > v_a$, we need $\frac{(1-\delta)p + c_p - c_a}{\pi_a\alpha_a} > \frac{\delta p + c_a}{1-\pi_a\alpha_a}$. This simplifies to $c_p > \frac{c_a + p(\pi_a\alpha_a - (1-\delta))}{1-\pi_a\alpha_a}$. Lastly, to have $v_p < 1$, we need $\frac{(1-\delta)p + c_p - c_a}{\pi_a\alpha_a} < 1$. This simplifies to $c_p + (1-\delta)p < c_a + \pi_a\alpha_a$.

Lastly, for case (III), in which consumers who purchase are all standard patching, choosing (B, P), $v_p = p + c_p$. For this case to be an equilibrium, we need

$v_p > 0$, $v_p < 1$, $v = v_p$ preferring (NB, NP) over (B, AP), and $v = v_p$ preferring (B, P) over (B, AP). The condition $v_p > 0$ is satisfied. For $v_p < 1$, we need the condition $c_p + p < 1$. For $v = v_p$ to prefer (NB, NP) over (B, AP), we need $0 \geq v - \delta p - c_a - \pi_a\alpha_a v$ for $v = v_p = c_p + p$. This becomes $c_p + (1-\delta)p \leq c_a + (c_p + p)\pi_a\alpha_a$. Lastly, for $v = v_p$ to prefer (B, P) over (B, AP), we need $v - p - c_p \geq v - \delta p - c_a - \pi_a\alpha_a v$ for $v = v_p = c_p + p$. This also simplifies to $c_p \leq \frac{c_a + p(\pi_a\alpha_a - (1-\delta))}{1-\pi_a\alpha_a}$. This concludes the proof of the consumer market equilibrium for the priced patching rights case when $\pi_i\alpha_i > 1$. $\square$

Proof of Lemma 3: We prove that if $\pi_i\alpha_i > 1$ and patching rights are priced by the vendor, then if $c_p - \pi_a\alpha_a < c_a \leq c_p(1 - \pi_a\alpha_a)$, we have

$$p^* = \frac{1-c_p}{2}, \tag{2.38}$$

$$\delta^* = \frac{1 - c_a - \pi_a\alpha_a}{1-c_p}, \tag{2.39}$$

and σ^* is characterized by $0 < v_a < v_p < 1$ such that the lower tier of users prefers automated patching. On the other hand, if $c_a > c_p(1 - \pi_a\alpha_a)$, then

$$p^* = \frac{1-c_p}{2}, \tag{2.40}$$

and σ^* is characterized by $0 < v_p < 1$ such that there is no user of automated patching in equilibrium.

Suppose $0 < v_a < 1$ is induced. Then the profit function is $\Pi_I(p, \delta) = \delta p(1 - v_a)$. Using Lemma A.2, we have that $v_a = \frac{\delta p + c_a}{1-\pi_a\alpha_a}$. Similar to the status quo case, the optimal price and discount satisfies $\delta_1^* p_1^* = \frac{1-c_a-\pi_a\alpha_a}{2}$ with the corresponding profit $\Pi_I^* = \frac{(1-c_a-\pi_a\alpha_a)^2}{4(1-\pi_a\alpha_a)}$. Notice in this case that there is not a unique maximizer, and in fact, the optimal (p, δ) traces out an isoprofit curve.

Next, suppose that $0 < v_a < v_p < 1$ is induced. Then the profit function is $\Pi_{II}(p, \delta) = p(1 - v_p) + \delta p(v_p - v_a)$. Using Lemma A.2, we have that $v_a = \frac{\delta p + c_a}{1-\pi_a\alpha_a}$ and $v_p = \frac{(1-\delta)p + c_p - c_a}{\pi_a\alpha_a}$. From the first-order condition for p, we have $p_2^*(\delta) = \frac{(1-\pi_a\alpha_a)(-c_p(1-\delta)+\pi_a\alpha_a) - c_a(1-\delta-\pi_a\alpha_a))}{2((1-\delta)^2 - \pi_a\alpha_a(1-2\delta))}$ with the second-order condition satisfied. Then maximizing $\Pi_{II}(p_2^*(\delta),\delta)$ with respect to δ, we find that $\delta_2^* = \frac{1-c_a-\pi_a\alpha_a}{1-c_p}$, and so $p_2^* = p_2^*(\delta_2^*) = \frac{1-c_p}{2}$. The corresponding profit is $\Pi_{II} = \frac{1}{4}\left(1 - 2c_p + \frac{(c_a-c_p)^2}{\pi_a\alpha_a} + \frac{c_a^2}{1-\pi_a\alpha_a}\right)$.

Lastly, suppose that $0 < v_p < 1$ is induced. Then the profit function is $\Pi_{III}(p,\delta) = p(1-v_p)$. Using Lemma A.2, we have that $v_p = c_p + p$. As in the case when patching rights aren't priced, the optimal price is found to be $p_3^* = \frac{1-c_p}{2}$ with the corresponding profit $\Pi_{III}^* = \frac{(1-c_p)^2}{4}$.

We next find conditions under which the maximizing price for each case indeed induces that market structure. For $0 < v_a < 1$, we need the set of conditions for Case (I) in Lemma A.2 to hold for p_1^*. For $\delta p + c_a + \pi_a\alpha_a < 1$ to hold for the p_1^* and δ_1^*, we need $c_a + \pi_a\alpha_a < 1$, which is one of the preliminary assumptions of the model to not rule out automated patching for every consumer. Secondly, for $c_p + (1-\delta)p \geq c_a + \pi_a\alpha_a$, we need $1 - c_a - \pi_a\alpha_a \geq \delta(1 + c_a + \pi_a\alpha_a - 2c_p)$. If $1 + c_a - 2c_p + \pi_a\alpha_a \leq 0$, then any δ satisfies this condition. Otherwise, we need $\delta \leq \frac{1-c_a-\pi_a\alpha_a}{1+c_a+\pi_a\alpha_a-2c_p}$. Note that in this case, $\frac{1-c_a-\pi_a\alpha_a}{1+c_a+\pi_a\alpha_a-2c_p} > 0$ so that such a δ (the corresponding $p_1^*(\delta)$) can be found. Last, we need $c_p + (1-\delta)p \geq c_a + (c_p+p)\pi_a\alpha_a$ to hold for the profit-maximizing p_1^* and δ_1^*. This simplifies to $\delta(c_a + (1-2c_p)(1-\pi_a\alpha_a)) \leq (1-\pi_a\alpha_a)(1-c_a-\pi_a\alpha_a)$. Then if $c_a + (1-2c_p)(1-\pi_a\alpha_a) \leq 0$, any δ satisfies this condition. Otherwise, we'll need $\delta \leq \frac{(1-\pi_a\alpha_a)(1-c_a-\pi_a\alpha_a)}{c_a+(1-2c_p)(1-\pi_a\alpha_a)}$. Note in this case that $\frac{(1-\pi_a\alpha_a)(1-c_a-\pi_a\alpha_a)}{c_a+(1-2c_p)(1-\pi_a\alpha_a)} > 0$ so that such a δ (and the corresponding $p_1^*(\delta)$) can be found. In summary, $0 < v_a < 1$ can always be induced in equilibrium by some p and δ, given a set of parameters $c_a, \pi_a\alpha_a$, and c_p that satisfy the preliminary model assumptions.

Similarly, for Case (II), the condition under which p_2^* would induce $0 < v_a < v_p < 1$ is $c_p - \pi_a\alpha_a < c_a < c_p(1-\pi_a\alpha_a)$. And lastly, for Case (III), we need $\delta \geq \frac{2c_a+(1+c_p)(1-\pi_a\alpha_a)}{1-c_p}$ for $c_p \leq c_a + (1-\delta)p + (c_p+p)\pi_a\alpha_a$ to hold for $p = p_3^*$. This means that $0 < v_p < 1$ can always be induced in equilibrium using p_3^* by setting a high enough δ, given a set of parameters $c_a, \pi_a\alpha_a$, and c_p that satisfy the preliminary model assumptions.

Next, we find the conditions under which the maximal profits of each case dominate each other. First, note that $\Pi_I^* \geq \Pi_{III}^*$ iff $c_a \leq 1 - \pi_a\alpha_a - (1-c_p)\sqrt{1-\pi_a\alpha_a}$.

Next, note that $\Pi^*_{II} - \Pi^*_I = \frac{(c_a+\pi_a\alpha_a-c_p)^2}{4\pi_a\alpha_a}$, so that if $0 < v_a < v_p < 1$ can be induced, then it will dominate $0 < v_a < 1$. Also, $\Pi^*_{II} - \Pi^*_{III} = \frac{(c_a-c_p(1-\pi_a\alpha_a))^2}{4\pi_a\alpha_a(1-\pi_a\alpha_a)}$, so that if $0 < v_a < v_p < 1$ can be induced, then it will dominate $0 < v_p < 1$ as well. Therefore, when $c_p - \pi_a\alpha_a < c_a < c_p(1 - \pi_a\alpha_a)$, then $0 < v_a < v_p < 1$ will be the equilibrium market structure.

Furthermore, consider the boundaries of this region. When $c_a = c_p - \pi_a\alpha_a$, then the profit of the adjacent region is $\Pi^*_I = \frac{(1-c_p)^2}{4(1-\pi_a\alpha_a)}$ while $\Pi^*_{II} = \frac{(1-c_p)^2}{4(1-\pi_a\alpha_a)}$ as well. Similarly, at the other end, when $c_a = c_p(1-\pi_a\alpha_a)$, then $\Pi^*_{III} = \frac{1}{4}\left(1 - c_p\right)^2 = \Pi^*_{II}$. This means that $0 < v_a < 1$ will be the equilibrium market structure for $c_a \leq c_p - \pi_a\alpha_a$ and $0 < v_p < 1$ will be the equilibrium market structure for $c_a > c_p(1 - \pi_a\alpha_a)$. □

Proof of Proposition 1: We show that for $\pi_i\alpha_i > 1$, if $c_p - \pi_a\alpha_a < c_a < 1 - \pi_a\alpha_a - (1 - c_p)\sqrt{1 - \pi_a\alpha_a}$, the increase in profitability of pricing patching rights is given by

$$\frac{\Pi_P - \Pi_{SQ}}{\Pi_{SQ}} = \frac{(1 - \pi_a\alpha_a)(c_a - c_p + \pi_a\alpha_a)^2}{\pi_a\alpha_a(1 - c_a - \pi_a\alpha_a)^2}\,. \tag{2.41}$$

First, note that $1-\pi_a\alpha_a-(1-c_p)\sqrt{1 - \pi_a\alpha_a} < c_p(1-\pi_a\alpha_a)$, since $0 < c_p < 1$ and $0 < \pi_a\alpha_a < 1$. Hence, when $c_p - \pi_a\alpha_a < c_a < 1 - \pi_a\alpha_a - (1 - c_p)\sqrt{1 - \pi_a\alpha_a}$, in both the status quo case and when patching rights are priced, the equilibrium consumer market structure is $0 < v_a < v_p < 1$. Then from the proof of Lemma 3 above, the profit under priced patching rights $\Pi_P = \frac{1}{4}\left(1 - 2c_p + \frac{(c_a-c_p)^2}{\pi_a\alpha_a} + \frac{c_a^2}{1-\pi_a\alpha_a}\right)$ and from the proof of Lemma 2, the status quo case has $\Pi_{SQ} = \frac{(1-c_a-\pi_a\alpha_a)^2}{4(1-\pi_a\alpha_a)}$. Simplifying, we have $\frac{\Pi_P-\Pi_{SQ}}{\Pi_{SQ}} = \frac{(1-\pi_a\alpha_a)(c_a-c_p+\pi_a\alpha_a)^2}{\pi_a\alpha_a(1-c_a-\pi_a\alpha_a)^2}$. □

2.6.2 Low $\pi_i\alpha_i$

Status Quo

Lemma A.3 *Under the status quo, i.e., $\delta = 1$, the complete threshold characterization of the consumer market equilibrium is as follows:*

(I) $(0 < v_a < 1)$*, where* $v_a = \frac{p+c_a}{1-\pi_a\alpha_a}$*:*

- *(A)* $p + c_a + \pi_a\alpha_a < 1$
- *(B)* $c_p \geq c_a + \pi_a\alpha_a$
- *(C)* $(c_a + p)\pi_i\alpha_i \geq c_a + p(\pi_a\alpha_a)$
- *(D)* $c_p \geq c_a + (c_p + p)\pi_a\alpha_a$
- *(E)* $\pi_i\alpha_i > \pi_a\alpha_a$ *and* $c_a + p(\pi_a\alpha_a) \leq (c_a + p)\pi_i\alpha_i$

(II) $(0 < v_b < 1)$*, where* $v_b = \frac{1}{2} + \frac{-1+\pi_i\alpha_i+\sqrt{(1-\pi_s\alpha_s-\pi_i\alpha_i)^2+4p\pi_s\alpha_s}}{2\pi_s\alpha_s}$*:*

- *(A)* $1 - 2c_p + \pi_i\alpha_i + \pi_s\alpha_s \leq \sqrt{4p\pi_s\alpha_s + (1 - \pi_s\alpha_s - \pi_i\alpha_i)^2}$
- *(B)* $\pi_i\alpha_i < 1 - p$
- *(C)* $(1 - \pi_a\alpha_a)(1 - \pi_i\alpha_i) + (-1 + 2c_a + 2p + \pi_a\alpha_a)\pi_s\alpha_s \geq (1 - \pi_a\alpha_a)\sqrt{4p\pi_s\alpha_s + (1 - \pi_s\alpha_s - \pi_i\alpha_i)^2}$
- *(D) Either* $1 + \pi_i\alpha_i + \pi_s\alpha_s - 2\pi_a\alpha_a > 0$ *and* $p < \frac{(1-\pi_a\alpha_a)(-\pi_a\alpha_a+\pi_i\alpha_i+\pi_s\alpha_s)}{\pi_s\alpha_s}$ *and* $2(\pi_a\alpha_a + c_a) + \sqrt{(\pi_i\alpha_i + \pi_s\alpha_s - 1)^2 + 4p\pi_s\alpha_s} \geq \pi_i\alpha_i + \pi_s\alpha_s + 1$*, or* $\left(-2\pi_a\alpha_a + \pi_i\alpha_i + \pi_s\alpha_s + 1 < 0 \text{ or } p > \frac{(1-\pi_a\alpha_a)(-\pi_a\alpha_a+\pi_i\alpha_i+\pi_s\alpha_s)}{\pi_s\alpha_s}\right)$ *and* $\pi_a\alpha_a(\pi_i\alpha_i + \pi_s\alpha_s) + 2\pi_s\alpha_s(c_a + p) + 1 \geq \pi_a\alpha_a + \pi_i\alpha_i + \pi_s\alpha_s + (1 - \pi_a\alpha_a)\sqrt{(\pi_i\alpha_i + \pi_s\alpha_s - 1)^2 + 4p\pi_s\alpha_s}$*, or* $p = \frac{(1-\pi_a\alpha_a)(-\pi_a\alpha_a+\pi_i\alpha_i+\pi_s\alpha_s)}{\pi_s\alpha_s}$

(III) $(0 < v_b < v_a < 1)$*, where* v_b *is the most positive root of the cubic* $f_1(x) \triangleq (1 - \pi_a\alpha_a)\pi_s\alpha_s x^3 + ((1 - \pi_a\alpha_a)(1 - \pi_i\alpha_i) - c_a\pi_s\alpha_s - p\pi_s\alpha_s)x^2 + (p(-1 + \pi_a\alpha_a) + p(-1 + \pi_i\alpha_i))x + p^2$ *and* $v_a = \frac{c_a v_b}{v_b(1-\pi_a\alpha_a)-p}$*:*

- *(A)* $(\pi_a\alpha_a + c_a - 1)(\pi_a\alpha_a - \pi_i\alpha_i + c_a) > \pi_s\alpha_s(\pi_a\alpha_a + c_a + p - 1)$

(B) $\pi_i\alpha_i(c_a + p) < c_a + p(\pi_a\alpha_a)$

(C) $\pi_a\alpha_a \leq c_p - c_a$

(IV) $(0 < v_b < v_a < v_p < 1)$, *where* v_b *is the most positive root of* $f_1(x)$ *and* $v_a = \frac{c_a v_b}{v_b(1-\pi_a\alpha_a)-p}$ *and* $v_p = \frac{c_p - c_a}{\pi_a\alpha_a}$*:*

(A) $c_p < c_a + \pi_a\alpha_a$

(B) $c_p(1 - \pi_a\alpha_a) > c_a$

(C) $c_p(\pi_a\alpha_a)^2(-c_a + c_p(1 - \pi_a\alpha_a)) + \pi_s\alpha_s(-c_a + c_p)^2(c_a - c_p(1 - \pi_a\alpha_a) + \pi_a\alpha_a p) - (c_a - c_p)(c_a + c_p(-1 + \pi_a\alpha_a))\pi_a\alpha_a\pi_i\alpha_i < 0$

(D) $c_a + p(\pi_a\alpha_a) > (c_a + p)\pi_i\alpha_i$

(V) $(0 < v_b < v_p < 1)$, *where* v_b *is the most positive root of* $f_2(x) \triangleq \pi_s\alpha_s x^3 + (1 - \pi_i\alpha_i - (c_p + p)\pi_s\alpha_s)x^2 - p(2 - \pi_i\alpha_i)x + p^2$ *and* $v_p = \frac{c_p v_b}{v_b - p}$*:*

(A) $(-1 + c_p + p)\pi_s\alpha_s < (1 - c_p)(-c_p + \pi_i\alpha_i)$

(B) $\pi_i\alpha_i < \frac{c_p}{c_p + p}$

(C) $(1 - \pi_a\alpha_a)(c_a + p\pi_a\alpha_a)(c_a + p\pi_a\alpha_a - (c_a + p)\pi_i\alpha_i) + (c_a + p)^2(c_a - c_p + (c_p + p)\pi_a\alpha_a)\pi_s\alpha_s \geq 0$

(D) $c_a + p(\pi_a\alpha_a) > 0$

(E) *Either* $c_a + c_p(-1 + \pi_a\alpha_a) \geq 0$, *or*

$c_a + c_p(-1 + \pi_a\alpha_a) < 0$ *and* $\pi_a\alpha_a(c_a + c_p(-1 + \pi_a\alpha_a))(c_p\pi_a\alpha_a + (c_a - c_p)\pi_i\alpha_i) \leq (-c_a + c_p)^2(c_a - c_p + (c_p + p)\pi_a\alpha_a)\pi_s\alpha_s$

(VI) $(0 < v_a < v_p < 1)$, *where* $v_a = \frac{p + c_a}{1 - \pi_a\alpha_a}$ *and* $v_p = \frac{c_p - c_a}{\pi_a\alpha_a}$*:*

(A) $c_p < c_a + \pi_a\alpha_a$

(B) $c_p > c_a + (c_p + p)\pi_a\alpha_a$

(C) $(c_p - c_a)\pi_i\alpha_i \geq c_p\pi_a\alpha_a$

(D) $\pi_i\alpha_i > \pi_a\alpha_a$ *and* $c_a + p\pi_a\alpha_a \leq (c_a + p)\pi_i\alpha_i$

(VII) $(0 < v_p < 1)$, *where* $v_p = p + c_p$*:*

(A) $c_p + p < 1$

(B) $c_a + (c_p + p)\pi_a\alpha_a \geq c_p$

(C) $(c_p + p)\pi_i\alpha_i \geq c_p$

Proof of Lemma A.3: This is proven as a sub-case in the proof of Lemma A.4. $\square$

Proof of Lemma 4: Technically, we prove the existence of $\tilde{\alpha}_1$ such that if $\pi_i\alpha_i < \min\left[\pi_a\alpha_a, \frac{c_p}{1+c_p}\right]$, then for $\alpha_s > \tilde{\alpha}_1$, p^* is set so that

1. if $c_p - \pi_a\alpha_a < c_a \leq 1 - \pi_a\alpha_a - (1-c_p)\sqrt{1-\pi_a\alpha_a}$, then $\sigma^*(v)$ is characterized by $0 < v_b < v_a < v_p < 1$ under optimal pricing, and
2. if $c_a > 1 - \pi_a\alpha_a - (1-c_p)\sqrt{1-\pi_a\alpha_a}$, then $\sigma^*(v)$ is characterized by $0 < v_b < v_p < 1$ under optimal pricing.

The sketch of the proof is as follows. From Lemma A.3, a unique consumer market equilibrium arises, given a price p. Within each region of the parameter space defined by Lemma A.3, the thresholds v_a, v_b, and v_p are smooth functions of the parameters, including p. In the cases where the thresholds are given in closed-form, this is clear. In the cases where these thresholds are implicitly defined as the root of some cubic equation, then the smoothness of the thresholds in the parameters follows from the Implicit Function Theorem. Specifically, for each of those cases, the threshold defined was the most positive root v_b^* of a cubic function of v_b, $f(v_b, p) = 0$. Moreover, the cubic $f(v_b, p)$ has two local extrema in v_b and is negative to the left of v_b^* and positive to the right of it ($f(v_b^* - \epsilon, p) < 0$ and $f(v_b^* + \epsilon, p) > 0$ for arbitrarily small $\epsilon > 0$). Therefore, $\frac{\partial f}{\partial v_b}(v_b, p) \neq 0$ so that the Implicit Function Theorem applies. The thresholds being smooth in p implies that the profit function for each case of the parameter space defined by Lemma A.3 is smooth in p. We find the profit-maximizing price within the compact closure of each case, so that the price that induces the largest profit among the cases will be the equilibrium price set by the vendor.

Having given the sketch of the proof, we now proceed with the proof. The conditions of this lemma precludes candidate market structures from arising in equilibrium. Specifically, $c_p - \pi_a\alpha_a < c_a$ rules out Cases (I) and (III) of Lemma A.3, and $\pi_i\alpha_i < \min\left[\frac{c_p\pi_a\alpha_a}{1+c_p-c_a}, \frac{c_p}{1+c_p}\right]$ rules out Cases (VI) and (VII). We consider the remaining possible consumer equilibria that can be induced when the vendor sets prices optimally. Suppose $0 < v_b < 1$ is induced. By part (II) of Lemma A.3, we obtain $v_b = \frac{1}{2} + \frac{-1+\pi_i\alpha_i+\sqrt{(1-\pi_s\alpha_s-\pi_i\alpha_i)^2+4p\pi_s\alpha_s}}{2\pi_s\alpha_s}$. The profit function in this case is $\Pi(p) = p(1 - v_b(p))$. Let C_{II} be the compact closure of the region of the parameter space defining $0 < v_b < 1$, given in part (II) of Lemma A.3. By the Weierstrass extreme value theorem, there exists a p in C_{II} that maximizes $\Pi(p)$. This p may be on the boundary, and we show that the vendor's profit function is continuous across region boundaries later. Otherwise, if this p is interior, the unconstrained maximizer satisfies the first-order condition $\Pi'(p) = 0$.

Using the first-order condition and letting

$$Q_1 \triangleq \sqrt{(-\pi_i\alpha_i + \pi_s\alpha_s + 1)^2 \left(\pi_s\alpha_s(\pi_i\alpha_i - 1) + (\pi_i\alpha_i - 1)^2 + (\pi_s\alpha_s)^2\right)},$$

the roots of $\Pi'(p) = 0$ are $\frac{-(\pi_i\alpha_i)^2+2\pi_i\alpha_i(1-2\pi_s\alpha_s)+\pi_s\alpha_s(4-\pi_s\alpha_s)-1}{9\pi_s\alpha_s} \mp \frac{Q_1}{9\pi_s\alpha_s}$.

However, $\frac{-(\pi_i\alpha_i)^2+2\pi_i\alpha_i(1-2\pi_s\alpha_s)+\pi_s\alpha_s(4-\pi_s\alpha_s)-1-Q_1}{9\pi_s\alpha_s} < 0$ for $\pi_s\alpha_s > 1 - \pi_i\alpha_i$, so for $\pi_s\alpha_s > 1 - \pi_i\alpha_i$, the unconstrained maximizer is given by

$$p_{II} = \frac{-\left(\pi_i\alpha_i\right)^2 + 2\pi_i\alpha_i(1 - 2\pi_s\alpha_s) + \pi_s\alpha_s(4 - \pi_s\alpha_s) - 1 + Q_1}{9\pi_s\alpha_s}. \tag{2.42}$$

The SOC is satisfied if $Q_1 + 2\left(\pi_s\alpha_s(\pi_i\alpha_i - 1) + (\pi_i\alpha_i - 1)^2 + (\pi_s\alpha_s)^2\right) > 0$, which holds when $\pi_s\alpha_s > 1 - \pi_i\alpha_i$. Substituting (2.42) into the profit function, we obtain

$$\Pi_{II} = \frac{1}{54\left(\pi_s\alpha_s\right)^2}\Bigg(\Big(\left(\pi_i\alpha_i\right)^2 - Q_1 + 2\pi_i\alpha_i(2\pi_s\alpha_s - 1) + \pi_s\alpha_s(\pi_s\alpha_s - 4) + 1\Big)$$
$$\Big(\sqrt{5\left(\pi_i\alpha_i\right)^2 + 4Q_1 + 2\pi_i\alpha_i(\pi_s\alpha_s - 5) + \pi_s\alpha_s(5\pi_s\alpha_s - 2) + 5}+3\pi_i\alpha_i-3\pi_s\alpha_s-3\Big)\Bigg). \tag{2.43}$$

On the other hand, suppose $0 < v_b < v_a < v_p < 1$ is induced. By part (IV) of Lemma A.3, we obtain that v_b is the most positive root of the cubic

$$f_1(x) \triangleq (1-\pi_a\alpha_a)\pi_s\alpha_s x^3 + ((1-\pi_a\alpha_a)(1-\pi_i\alpha_i) - c_a\pi_s\alpha_s - p\pi_s\alpha_s)x^2 + (p(-1+\pi_a\alpha_a) + p(-1+\pi_i\alpha_i))x + p^2. \quad (2.44)$$

The profit function is $\Pi_{IV}(p) = p(1-v_b(p))$. Let C_{IV} be the compact closure of the region of the parameter space defining $0 < v_b < v_a < v_p < 1$, given in part (IV) of Lemma A.3. Again, by the Weierstrass extreme value theorem, there exists a p in C_{IV} that maximizes $\Pi(p)$. This p may be on the boundary, and we show that the vendor's profit function is continuous across region boundaries later. Otherwise, if this p is interior, the unconstrained maximizer satisfies the first-order condition $\Pi'(p) = 0$. The first-order condition is given by

$$\Pi'_{IV}(p) = (1-v_b(p)) - pv'_b(p) = 0. \quad (2.45)$$

By equating (2.44) to 0 and implicitly differentiating, we have that

$$v'_b(p) = (v_b(p)(\pi_a\alpha_a + \pi_i\alpha_i - \pi_s\alpha_s v_b(p) - 2) + 2p)\big(v_b(p)(-2(\pi_a\alpha_a - 1)(\pi_i\alpha_i - 1) + 2\pi_s\alpha_s(c_a+p) + 3\pi_s\alpha_s(\pi_a\alpha_a - 1)v_b(p)) - p(\pi_a\alpha_a + \pi_i\alpha_i - 2)\big)^{-1} \quad (2.46)$$

Substituting this into (2.45) and re-writing (2.44), we have that $v_b(p^*)$ and p^* simultaneously need to solve

$$(1-\pi_a\alpha_a)\pi_s\alpha_s v_b^3 + ((1-\pi_a\alpha_a)(1-\pi_i\alpha_i) - c_a\pi_s\alpha_s - p\pi_s\alpha_s)v_b^2 + (p(-1+\pi_a\alpha_a) + p(-1+\pi_i\alpha_i))v_b + p^2 = 0, \text{and} \quad (2.47)$$

$$1 - v_b - (p(2p + v_b(\pi_a\alpha_a + \pi_i\alpha_i - \pi_s\alpha_s v_b - 2)))\big(v_b(-2(\pi_a\alpha_a - 1)(\pi_i\alpha_i - 1) + 2\pi_s\alpha_s(c_a+p) + 3\pi_s\alpha_s v_b(\pi_a\alpha_a - 1)) - p(\pi_a\alpha_a + \pi_i\alpha_i - 2)\big)^{-1} = 0. \quad (2.48)$$

Letting

$$Q_2 \triangleq \Big((\pi_a\alpha_a+\pi_i\alpha_i+\pi_s\alpha_s(v_b-2)v_b-2)^2-8(v_b-1)v_b(\pi_s\alpha_s(2c_a+3v_b(\pi_a\alpha_a-1))- 2(\pi_a\alpha_a-1)(\pi_i\alpha_i-1))\Big)^{\frac{1}{2}} \quad (2.49)$$

and solving (2.48) for p, we have that p is either

$$p = \frac{1}{4}\left(-\pi_a\alpha_a-\pi_i\alpha_i-\pi_s\alpha_s v_b^2+2\pi_s\alpha_s v_b+2-Q_2\right)$$

or

$$p = \frac{1}{4}\left(-\pi_a\alpha_a-\pi_i\alpha_i-\pi_s\alpha_s v_b^2+2\pi_s\alpha_s v_b+2+Q_2\right).$$

We can rule out the larger root when $\pi_s\alpha_s > 2+\pi_a\alpha_a+\pi_i\alpha_i$ since when $\pi_s\alpha_s > 2+\pi_a\alpha_a+\pi_i\alpha_i$, then $Q_2 > -2+4v_b+\pi_a\alpha_a+\pi_i\alpha_i-(2-v_b)v_b\pi_s\alpha_s$. This is equivalent to the larger root for p^* being greater than v_b, which can't happen in equilibrium. Therefore,

$$p(v_b) = \frac{1}{4}\left(-\pi_a\alpha_a-\pi_i\alpha_i-\pi_s\alpha_s v_b^2+2\pi_s\alpha_s v_b+2-Q_2\right) \quad (2.50)$$

Substituting this into (2.47), we have that $v_b(p^*)$ solves

$$\big((\pi_a\alpha_a+\pi_i\alpha_i-2)^2+4v_b^2(\pi_a\alpha_a(4\pi_i\alpha_i+5\pi_s\alpha_s-4)+2\pi_i\alpha_i(\pi_s\alpha_s-2)+ \pi_s\alpha_s(\pi_s\alpha_s-4c_a-7)+4)-4\pi_s\alpha_s v_b(\pi_a\alpha_a+\pi_i\alpha_i-2c_a-2)+3(\pi_s\alpha_s)^2v_b^4- 2\pi_s\alpha_s v_b^3(11\pi_a\alpha_a+\pi_i\alpha_i+4\pi_s\alpha_s-12)-16v_b-2v_b\big((\pi_a\alpha_a)^2+2\pi_a\alpha_a(3\pi_i\alpha_i-4)+ \pi_i\alpha_i(\pi_i\alpha_i-8)\big)\big)+(\pi_s\alpha_s v_b(3v_b-2)-(2v_b-1)(\pi_a\alpha_a+\pi_i\alpha_i-2))Q_2 = 0 \quad (2.51)$$

A generalization of the Implicit Function Theorem gives that v_b is not only a smooth function of the parameters, but it's also an analytic function of the parameters so that it can be represented locally as a Taylor series of its parameters (Brillinger 1966). More specifically, since $f_1'(x) \neq 0$ at the root for which v_b is defined, there exists an $\alpha_1 > 0$ such that for $\alpha_s > \alpha_1$, $v_b = \sum_{k=0}^{\infty} \frac{a_k}{(\pi_s\alpha_s)^k}$ for some coefficients a_k .

Substituting this into (2.51), we have $\pi_s\alpha_s\left(\frac{a_0^3((2a_0-1)\pi_a\alpha_a-2a_0+c_a+1)}{a_0-2}\right)+\sum_{k=0}^{\infty}\frac{A_k}{(\pi_s\alpha_s)^k}=0$. Then $a_0=0$ or $a_0=\frac{1+c_a-\pi_a\alpha_a}{2(1-\pi_a\alpha_a)}$ are the only solutions for a_0 that make the first term 0. However, if $a_0=0$, then $\pi_s\alpha_s\left(\frac{a_0^3((2a_0-1)\pi_a\alpha_a-2a_0+c_a+1)}{a_0-2}\right)+\sum_{k=0}^{\infty}\frac{A_k}{(\pi_s\alpha_s)^k}$ becomes $c_a^2+\sum_{k=1}^{\infty}\frac{A_k}{(\pi_s\alpha_s)^k}$. Note that $c_a^2+\sum_{k=1}^{\infty}\frac{A_k}{(\pi_s\alpha_s)^k}>0$ for large enough $\pi_s\alpha_s$ so that there exists no coefficients A_k such that $c_a^2+\sum_{k=1}^{\infty}\frac{A_k}{(\pi_s\alpha_s)^k}=0$ for $\alpha_s>\alpha_1$, a contradiction.

Thus $a_0=\frac{1+c_a-\pi_a\alpha_a}{2(1-\pi_a\alpha_a)}$. Substituting $v_b=\frac{1+c_a-\pi_a\alpha_a}{2(1-\pi_a\alpha_a)}+\sum_{k=1}^{\infty}a_k\xi^k$ into (2.51), we have that $\frac{a_1(-\pi_a\alpha_a+c_a+1)^3+2c_a(\pi_a\alpha_a-1)\left((\pi_a\alpha_a-1)(\pi_a\alpha_a-\pi_i\alpha_i)+c_a^2+c_a(2\pi_a\alpha_a+\pi_i\alpha_i-3)\right)}{2(\pi_a\alpha_a-1)(3\pi_a\alpha_a+c_a-3)}+\sum_{k=1}^{\infty}\frac{A_k}{(\pi_s\alpha_s)^k}=0$. Solving for a_1 to make this first term zero, we have

$$a_1=\frac{2c_a(1-\pi_a\alpha_a)\left((1-\pi_a\alpha_a)(-\pi_a\alpha_a+\pi_i\alpha_i)+c_a^2+c_a(2\pi_a\alpha_a+\pi_i\alpha_i-3)\right)}{(-\pi_a\alpha_a+c_a+1)^3}.$$

Using $v_b=\frac{1+c_a-\pi_a\alpha_a}{2(1-\pi_a\alpha_a)}+\frac{2c_a(1-\pi_a\alpha_a)\left((1-\pi_a\alpha_a)(-\pi_a\alpha_a+\pi_i\alpha_i)+c_a^2+c_a(2\pi_a\alpha_a+\pi_i\alpha_i-3)\right)}{(-\pi_a\alpha_a+c_a+1)^3\pi_s\alpha_s}+\sum_{k=2}^{\infty}\frac{a_k}{(\pi_s\alpha_s)^k}$, we can solve for a_2, a_3, and so on recursively by repeatedly substituting this expression for v_b into (2.51) and solving for the coefficients to make the expression zero. Doing this, we find that the threshold v_b is

$$\begin{aligned}v_b=&\frac{1+c_a-\pi_a\alpha_a}{2(1-\pi_a\alpha_a)}+\\&\frac{2c_a(1-\pi_a\alpha_a)\left((1-\pi_a\alpha_a)(-\pi_a\alpha_a+\pi_i\alpha_i)+c_a^2+c_a(2\pi_a\alpha_a+\pi_i\alpha_i-3)\right)}{(-\pi_a\alpha_a+c_a+1)^3\pi_s\alpha_s}+\\&\sum_{k=2}^{\infty}a_k\left(\frac{1}{\pi_s\alpha_s}\right)^k,\end{aligned}\tag{2.52}$$

and substituting this into (2.50), we have that the optimal price set by the vendor is

$$\begin{aligned}p_{IV}^*=&\frac{1}{2}(1-\pi_a\alpha_a-c_a)+\\&\frac{2c_a^2(\pi_a\alpha_a-1)((\pi_a\alpha_a-1)(2\pi_a\alpha_a-\pi_i\alpha_i-1)+c_a(2\pi_a\alpha_a+\pi_i\alpha_i-3))}{(-\pi_a\alpha_a+c_a+1)^3\pi_s\alpha_s}+\\&\sum_{k=2}^{\infty}b_k\left(\frac{1}{\pi_s\alpha_s}\right)^k.\end{aligned}\tag{2.53}$$

The corresponding profit is given as

$$\Pi^*_{IV} = \frac{(\pi_a\alpha_a + c_a - 1)^2}{4(1-\pi_a\alpha_a)} + \frac{c_a(\pi_a\alpha_a + c_a - 1)((\pi_a\alpha_a - 1)(\pi_a\alpha_a - \pi_i\alpha_i) + c_a(\pi_a\alpha_a + \pi_i\alpha_i - 2))}{(-\pi_a\alpha_a + c_a + 1)^2\pi_s\alpha_s} + \sum_{k=2}^{\infty} c_k \left(\frac{1}{\pi_s\alpha_s}\right)^k. \quad (2.54)$$

As a matter of notation, we will use a_k, b_k, and c_k to denote coefficients in the Taylor expansions without referring to specific expressions throughout the appendix. These will be used across different cases, and they don't refer to the same quantities or expressions across cases.

Lastly, suppose $0 < v_b < v_p < 1$ is induced. The profit function in this case is $\Pi_V(p) = p(1 - v_b(p))$, where v_b is the most positive root of $f_2(x) \triangleq \pi_s\alpha_s x^3 + (1 - \pi_i\alpha_i - (c_p + p)\pi_s\alpha_s)x^2 - p(2 - \pi_i\alpha_i)x + p^2$ by part (V) of Lemma A.3. Omitting the algebra (similar to the previous case), there exists an $\alpha_2 > 0$ such that for $\alpha_s > \alpha_2$, the unconstrained maximizer is given as

$$p^*_V = \frac{1-c_p}{2} - \frac{2c_p^2(1 - 3c_p + \pi_i\alpha_i(1+c_p))}{(1+c_p)^3\pi_s\alpha_s} + \sum_{k=2}^{\infty} b_k \left(\frac{1}{\pi_s\alpha_s}\right)^k, \quad (2.55)$$

and the profit induced is given as

$$\Pi_V = \frac{1}{4}(1-c_p)^2 + \frac{(1-c_p)c_p\,(2c_p - \pi_i\alpha_i(1+c_p))}{(1+c_p)^2\pi_s\alpha_s} + \sum_{k=2}^{\infty} c_k \left(\frac{1}{\pi_s\alpha_s}\right)^k. \quad (2.56)$$

To compare profits, we note that there exists $\alpha_3 > 0$ such that for $\alpha_s > \alpha_3$, (2.43) can be expressed as the Taylor series

$$\Pi_{II} = \frac{(1-\pi_i\alpha_i)^2}{4\pi_s\alpha_s} + \sum_{k=2}^{\infty} c_k \left(\frac{1}{\pi_s\alpha_s}\right)^k. \quad (2.57)$$

Then comparing (2.57) with either (2.54) or (2.56), we find that (2.43) is dominated by the other two profits when α_s exceeds an implicit bound (say,

$\alpha_s > \hat{\alpha}_1$, for some $\hat{\alpha}_1 > 0$).

Next, using (2.53) with Lemma A.3, we find the conditions under which the interior optimal price of $0 < v_b < v_a < v_p < 1$ would indeed induce this market structure. First, we still need the conditions $c_p < c_a + \pi_a\alpha_a$ and $c_p(1-\pi_a\alpha_a) > c_a$. Next, for $c_p(\pi_a\alpha_a)^2(-c_a + c_p(1-\pi_a\alpha_a)) + \pi_s\alpha_s(-c_a+c_p)^2(c_a - c_p(1-\pi_a\alpha_a) + \pi_a\alpha_a p) - (c_a - c_p)(c_a + c_p(-1+\pi_a\alpha_a))\pi_a\alpha_a\pi_i\alpha_i < 0$ to hold for $p = p^*_{IV}$, we need $-\frac{1}{2}\left((c_p-c_a)^2(-2c_a+2c_p+(-1+c_a-2c_p)\pi_a\alpha_a+(\pi_a\alpha_a)^2)\right)\pi_s\alpha_s + \sum_{k=0}^{\infty} A_k\left(\frac{1}{\pi_s\alpha_s}\right)^k < 0$ for some coefficients A_k. There exists $\alpha_4 > 0$ such that if $\alpha_s > \alpha_4$, then $c_a < \frac{(2c_p-\pi_a\alpha_a)(1-\pi_a\alpha_a)}{2-\pi_a\alpha_a}$ is sufficient for this condition to hold. Note that $\frac{(2c_p-\pi_a\alpha_a)(1-\pi_a\alpha_a)}{2-\pi_a\alpha_a} < c_p(1-\pi_a\alpha_a)$, so $c_a < \frac{(2c_p-\pi_a\alpha_a)(1-\pi_a\alpha_a)}{2-\pi_a\alpha_a}$ is the tighter bound on c_a. Lastly, for $c_a + p(\pi_a\alpha_a) > (c_a+p)\pi_i\alpha_i$ to hold for $p = p^*_{IV}$, we need $c_a - \frac{1}{2}\pi_a\alpha_a(-1+c_a+\pi_a\alpha_a) - \frac{1}{2}\pi_i\alpha_i(1+c_a-\pi_a\alpha_a) + \sum_{k=0}^{\infty} B_k\left(\frac{1}{\pi_s\alpha_s}\right)^k > 0$ for some coefficients $B_k > 0$. It suffices to have $c_a > \frac{(1-\pi_a\alpha_a)(\pi_a\alpha_a-\pi_i\alpha_i)}{(-2+\pi_a\alpha_a+\pi_i\alpha_i)}$. Since $\pi_i\alpha_i < \pi_a\alpha_a$ follows from one of the assumptions of this lemma ($\pi_i\alpha_i < \frac{c_p\pi_a\alpha_a}{1+c_p-c_a}$), this condition is automatically satisfied since $c_a > 0$.

In summary, the optimal price $0 < v_b < v_a < v_p < 1$ indeed induces this market structure when $c_p - \pi_a\alpha_a < c_a < \frac{(2c_p-\pi_a\alpha_a)(1-\pi_a\alpha_a)}{2-\pi_a\alpha_a}$ for $\alpha > \alpha_4$.

Similarly, using (2.55) with Lemma A.3, there exists $\alpha_5 > 0$ such that when $\alpha_s > \alpha_5$, the optimal price of $0 < v_b < v_p < 1$ indeed induces the correct market structure when $c_a > c_p - \frac{1}{2}(1+c_p)\pi_a\alpha_a$ and $\pi_i\alpha_i < \frac{2c_p}{1+c_p}$.

Let $\tilde{\alpha}_1$ be the max of $\hat{\alpha}_1$, α_1, α_2, α_3, α_4, and α_5. Then for $\alpha_s > \tilde{\alpha}_1$, if $c_a > c_p - \pi_a\alpha_a$, we have that (2.43) is dominated. Moreover, since $c_p - \frac{1}{2}(1+c_p)\pi_a\alpha_a < \frac{(2c_p-\pi_a\alpha_a)(1-\pi_a\alpha_a)}{2-\pi_a\alpha_a}$, there will be a region in which the interior maximizers of both $0 < v_b < v_a < v_p < 1$ and $0 < v_b < v_p < 1$ induce their corresponding cases.

Comparing (2.54) with (2.56), we see that (2.54) is greater when $2c_a(1-\pi_a\alpha_a) < c_a^2 + (1-\pi_a\alpha_a)(c_p(2-c_p) - \pi_a\alpha_a)$, which can be written as $c_a < 1 - \pi_a\alpha_a - (1-c_p)\sqrt{1-\pi_a\alpha_a}$. Note that for any $c_p \in [0,1]$, we have that $c_p - \frac{1}{2}(1+c_p)\pi_a\alpha_a < 1-\pi_a\alpha_a-(1-c_p)\sqrt{1-\pi_a\alpha_a} < \frac{(2c_p-\pi_a\alpha_a)(1-\pi_a\alpha_a)}{2-\pi_a\alpha_a}$ from $\pi_a\alpha_a \in (0,1)$. Therefore, for $\alpha_s > \tilde{\alpha}_1$, if $c_p - \pi_a\alpha_a < c_a \leq 1-\pi_a\alpha_a-(1-c_p)\sqrt{1-\pi_a\alpha_a}$, then $\sigma^*(v)$ is characterized by $0 < v_b < v_a < v_a < 1$ under optimal pricing, and if $c_a > 1 - \pi_a\alpha_a - (1-c_p)\sqrt{1-\pi_a\alpha_a}$, then $\sigma^*(v)$ is characterized by $0 < v_b < v_p < 1$ under optimal pricing. $\square$

Pricing Patching Rights

Lemma A.4 *Under priced patching rights, the complete threshold characterization of the consumer market equilibrium is as follows:*

(I) $(0 < v_a < 1)$, *where* $v_a = \frac{\delta p + c_a}{1-\pi_a\alpha_a}$*:*

(A) $\delta p + c_a + \pi_a\alpha_a < 1$

(B) $(1-\delta)p + c_p \geq c_a + \pi_a\alpha_a$

(C) $(c_a + \delta p)\pi_i\alpha_i \geq c_a + p(-1+\delta+\pi_a\alpha_a)$

(D) $c_p + (1-\delta)p \geq c_a + (c_p + p)\pi_a\alpha_a$

(E) Either $\pi_i\alpha_i < \pi_a\alpha_a$ *and* $\pi_a\alpha_a - \pi_i\alpha_i + c_a \leq (1-\delta)p$*, or*
$\pi_i\alpha_i > \pi_a\alpha_a$ *and* $c_a + p(-1+\delta+\pi_a\alpha_a) \leq (c_a+\delta p)\pi_i\alpha_i$*, or*
$\pi_i\alpha_i = \pi_a\alpha_a$ *and* $(1-\delta)p - c_a \geq 0$

(II) $(0 < v_a < v_b < 1)$, *where* $v_a = \frac{\delta p + c_a}{1-\pi_a\alpha_a}$ *and*
$v_b = \frac{-\pi_a\alpha_a + \pi_i\alpha_i + \pi_s\alpha_s + \sqrt{(-\pi_a\alpha_a+\pi_i\alpha_i+\pi_s\alpha_s)^2 - 4\pi_s\alpha_s(c_a+(\delta-1)p)}}{2\pi_s\alpha_s}$*:*

(A) $(1-\delta)p - c_a > 0$

(B) $\frac{(2\pi_s\alpha_s)(c_a+\delta p)}{1-\pi_a\alpha_a} < -\pi_a\alpha_a + \pi_i\alpha_i + \pi_s\alpha_s + \sqrt{(-\pi_a\alpha_a + \pi_i\alpha_i + \pi_s\alpha_s)^2 - 4\pi_s\alpha_s(c_a + (\delta-1)p)}$

(C) $(1-\delta)p < \pi_a\alpha_a - \pi_i\alpha_i + c_a$

(D) $\pi_a\alpha_a - \pi_i\alpha_i + \pi_s\alpha_s > 0$

(E) $\pi_a\alpha_a + \pi_i\alpha_i + \pi_s\alpha_s \leq \sqrt{(-\pi_a\alpha_a + \pi_i\alpha_i + \pi_s\alpha_s)^2 - 4\pi_s\alpha_s(c_a + (\delta-1)p)} + 2c_p$

(III) $(0 < v_a < v_b < v_p < 1)$, *where* $v_a = \frac{\delta p + c_a}{1-\pi_a\alpha_a}$, v_b *is the most positive root of the cubic* $f_3(x) \triangleq \pi_s\alpha_s x^2(c_a - c_p + (\delta-1)p + \pi_a\alpha_a x) + (c_a + (\delta-1)p + \pi_a\alpha_a x)(c_a + (\delta-1)p + x(\pi_a\alpha_a - \pi_i\alpha_i))$*, and* $v_p = \frac{c_p v_b}{c_a - (1-\delta)p + \pi_a\alpha_a v_b}$*:*

(A) $(1-\delta)p - c_a \geq 0$

(B) $\pi_a\alpha_a > c_p$

(C) $\pi_s\alpha_s(-\pi_a\alpha_a - c_a + c_p - \delta p + p) < (c_p - \pi_a\alpha_a)(c_p - \pi_i\alpha_i)$

(D) $\pi_i\alpha_i(c_a + \delta p) + \pi_a\alpha_a c_p > \pi_i\alpha_i(c_p + p)$

(E) *Either* $c_a + p(\pi_a\alpha_a + \delta) \leq p$, *or*

$c_a + p(\pi_a\alpha_a + \delta) > p$ *and* $\pi_s\alpha_s(c_a + \delta p)^2(c_a + \pi_a\alpha_a(c_p + p) - c_p + (\delta - 1)p) + (\pi_a\alpha_a - 1)(c_a + p(\pi_a\alpha_a + \delta - 1))(\pi_i\alpha_i(c_a + \delta p) - c_a - p(\pi_a\alpha_a + \delta - 1)) > 0$

(IV) $(0 < v_b < 1)$, *where* $v_b = \frac{1}{2} + \frac{-1+\pi_i\alpha_i+\sqrt{(1-\pi_s\alpha_s-\pi_i\alpha_i)^2+4p\pi_s\alpha_s}}{2\pi_s\alpha_s}$:

(A) $1 - 2c_p + \pi_i\alpha_i + \pi_s\alpha_s \leq \sqrt{4p\pi_s\alpha_s + (1 - \pi_s\alpha_s - \pi_i\alpha_i)^2}$

(B) $\pi_i\alpha_i < 1 - p$

(C) $(1 - \pi_a\alpha_a)(1 - \pi_i\alpha_i) + (-1 + 2c_a + 2\delta p + \pi_a\alpha_a)\pi_s\alpha_s \geq (1 - \pi_a\alpha_a)\sqrt{4p\pi_s\alpha_s + (1 - \pi_s\alpha_s - \pi_i\alpha_i)^2}$

(D) *Either* $1 + \pi_i\alpha_i + \pi_s\alpha_s - 2\pi_a\alpha_a > 0$ *and* $p < \frac{(1-\pi_a\alpha_a)(-\pi_a\alpha_a+\pi_i\alpha_i+\pi_s\alpha_s)}{\pi_s\alpha_s}$ *and* $2(\pi_a\alpha_a + c_a - (1-\delta)p) + \sqrt{(\pi_i\alpha_i + \pi_s\alpha_s - 1)^2 + 4p\pi_s\alpha_s} \geq \pi_i\alpha_i + \pi_s\alpha_s + 1$, *or*

$\left(-2\pi_a\alpha_a + \pi_i\alpha_i + \pi_s\alpha_s + 1 < 0 \text{ or } p > \frac{(1-\pi_a\alpha_a)(-\pi_a\alpha_a+\pi_i\alpha_i+\pi_s\alpha_s)}{\pi_s\alpha_s} \right)$ *and* $\pi_a\alpha_a(\pi_i\alpha_i + \pi_s\alpha_s) + 2\pi_s\alpha_s(c_a + \delta p) + 1 \geq \pi_a\alpha_a + \pi_i\alpha_i + \pi_s\alpha_s + (1 - \pi_a\alpha_a)\sqrt{(\pi_i\alpha_i + \pi_s\alpha_s - 1)^2 + 4p\pi_s\alpha_s}$, *or*

$p = \frac{(1-\pi_a\alpha_a)(-\pi_a\alpha_a+\pi_i\alpha_i+\pi_s\alpha_s)}{\pi_s\alpha_s}$ *and* $(1-\delta)p - c_a \leq 0$

(V) $(0 < v_b < v_a < 1)$, *where* v_b *is the most positive root of the cubic* $f_4(x) \triangleq (1 - \pi_a\alpha_a)\pi_s\alpha_s x^3 + ((1 - \pi_a\alpha_a)(1 - \pi_i\alpha_i) - c_a\pi_s\alpha_s - \delta p\pi_s\alpha_s)x^2 + (p(-1 + \pi_a\alpha_a) + p(-1 + \pi_i\alpha_i))x + p^2$ *and* $v_a = \frac{(c_a-(1-\delta)p)v_b}{v_b(1-\pi_a\alpha_a)-p}$:

(A) $(1-\delta)p - c_a < 0$

(B) $(\pi_a\alpha_a + c_a - 1 - (1-\delta)p)(\pi_a\alpha_a - \pi_i\alpha_i + c_a - (1-\delta)p) > \pi_s\alpha_s(\pi_a\alpha_a + c_a + \delta p - 1)$

(C) $\pi_i\alpha_i(c_a + \delta p) < c_a + p(-1 + \delta + \pi_a\alpha_a)$

(D) $\pi_a\alpha_a \le (1-\delta)p + c_p - c_a$

(VI) $(0 < v_b < v_a < v_p < 1)$, *where* v_b *is the most positive root of* $f_4(x)$ *and* $v_a = \frac{(c_a-(1-\delta)p)v_b}{v_b(1-\pi_a\alpha_a)-p}$ *and* $v_p = \frac{(1-\delta)p+c_p-c_a}{\pi_a\alpha_a}$*:*

(A) $(1-\delta)p - c_a < 0$

(B) $c_p + (1-\delta)p < c_a + \pi_a\alpha_a$

(C) $c_p(1-\pi_a\alpha_a) + (1-\delta)p > c_a$

(D) $c_p(\pi_a\alpha_a)^2(-c_a+(1-\delta)p+c_p(1-\pi_a\alpha_a))+\pi_s\alpha_s(-c_a+(1-\delta)p+c_p)^2(c_a - c_p(1-\pi_a\alpha_a) + (\pi_a\alpha_a - 1 + \delta)p) - (c_a - c_p - (1-\delta)p)(c_a - (1-\delta)p - c_p(1-\pi_a\alpha_a))\pi_a\alpha_a\pi_i\alpha_i < 0$

(E) $c_a + p(-1+\delta+\pi_a\alpha_a) > (c_a+\delta p)\pi_i\alpha_i$

(VII) $(0 < v_b < v_p < 1)$, *where* v_b *is the most positive root of* $f_5(x) \triangleq \pi_s\alpha_s x^3 + (1-\pi_i\alpha_i - (c_p+p)\pi_s\alpha_s)x^2 - p(2-\pi_i\alpha_i)x + p^2$ *and* $v_p = \frac{c_p v_b}{v_b - p}$*:*

(A) $(-1+c_p+p)\pi_s\alpha_s < (1-c_p)(-c_p+\pi_i\alpha_i)$

(B) $\pi_i\alpha_i < \frac{c_p}{c_p+p}$

(C) $(1-\pi_a\alpha_a)(c_a+(\pi_a\alpha_a-(1-\delta))p)(c_a+p(\pi_a\alpha_a-(1-\delta))-(c_a+\delta p)\pi_i\alpha_i) + (c_a+\delta p)^2(c_a - c_p - (1-\delta)p + (c_p+p)\pi_a\alpha_a)\pi_s\alpha_s \ge 0$

(D) $c_a + p(\delta + \pi_a\alpha_a) > p$

(E) *Either* $c_a - (1-\delta)p + c_p(-1+\pi_a\alpha_a) \ge 0$, *or*

$c_a+c_p(-1+\pi_a\alpha_a) < 0$ *and* $\pi_a\alpha_a(c_a-(1-\delta)p+c_p(-1+\pi_a\alpha_a))(c_p\pi_a\alpha_a + (c_a - c_p - (1-\delta)p)\pi_i\alpha_i) \le (-c_a + c_p + (1-\delta)p)^2(c_a - c_p - (1-\delta)p + (c_p+p)\pi_a\alpha_a)\pi_s\alpha_s$

(VIII) $(0 < v_a < v_p < 1)$, *where* $v_a = \frac{\delta p + c_a}{1-\pi_a\alpha_a}$ *and* $v_p = \frac{(1-\delta)p+c_p-c_a}{\pi_a\alpha_a}$*:*

(A) $c_p + (1-\delta)p < c_a + \pi_a\alpha_a$

(B) $c_p + (1-\delta)p > c_a + (c_p+p)\pi_a\alpha_a$

(C) $(c_p - c_a + (1-\delta)p)\pi_i\alpha_i \ge c_p\pi_a\alpha_a$

(D) Either $\pi_i\alpha_i < \pi_a\alpha_a$ *and* $c_p\pi_a\alpha_a + \pi_i\alpha_i(c_a - c_p - (1-\delta)p) \leq 0$*, or*

$\pi_i\alpha_i > \pi_a\alpha_a$ *and* $c_a + p(-1+\delta+\pi_a\alpha_a) \leq (c_a+\delta p)\pi_i\alpha_i$*, or*

$\pi_i\alpha_i = \pi_i\alpha_i$ *and* $(1-\delta)p - c_a \geq 0$

(IX) $(0 < v_p < 1)$*, where* $v_p = p + c_p$*:*

(A) $c_p + p < 1$

(B) $c_a + (c_p+p)\pi_a\alpha_a \geq c_p + (1-\delta)p$

(C) $(c_p+p)\pi_i\alpha_i \geq c_p$

Proof of Lemma A.4: First, we establish the general threshold-type equilibrium structure. Given the size of unpatched user population u, the net payoff of the consumer with type v for strategy profile σ is written as

$$U(v,\sigma) \triangleq \begin{cases} v - p - c_p & if \quad \sigma(v) = (B,P)\,; \\ v - p - \pi_s\alpha_s uv - \pi_i\alpha_i v & if \quad \sigma(v) = (B,NP)\,; \\ v - \delta p - c_a - \pi_a\alpha_a v & if \quad \sigma(v) = (B,AP)\,; \\ 0 & if \quad \sigma(v) = (NB,NP)\,. \end{cases} \tag{2.58}$$

Note $\sigma(v) = (B,P)$ if and only if

$$v - p - c_p \geq v - p - \pi_s\alpha_s uv - \pi_i\alpha_i v \Leftrightarrow v \geq \frac{c_p}{\pi_s\alpha_s u + \pi_i\alpha_i}, \text{ and}$$

$$v - p - c_p \geq v - \delta p - c_a - \pi_a\alpha_a v \Leftrightarrow v \geq \frac{(1-\delta)p + c_p - c_a}{\pi_a\alpha_a}, \text{ and}$$

$$v - p - c_p \geq 0 \Leftrightarrow v \geq c_p + p,$$

which can be summarized as

$$v \geq \max\left(\frac{c_p}{\pi_s\alpha_s u + \pi_i\alpha_i}, \frac{(1-\delta)p + c_p - c_a}{\pi_a\alpha_a}, c_p + p\right). \tag{2.59}$$

By (2.59), if a consumer with valuation v_0 buys and patches the software, then every consumer with valuation $v > v_0$ will also do so. Hence, there exists a threshold $v_p \in (0, 1]$ such that for all $v \in \mathcal{V}$, $\sigma^*(v) = (B, P)$ if and only if $v \geq v_p$. Similarly, $\sigma(v) \in \{(B, P), (B, NP), (B, AP)\}$, i.e., the consumer purchases one of the alternatives, if and only if

$$v - p - c_p \geq 0 \Leftrightarrow v \geq c_p + p, \text{ or}$$

$$v - p - \pi_s\alpha_s uv - \pi_i\alpha_i v \geq 0 \Leftrightarrow v \geq \frac{p}{1 - \pi_s\alpha_s u - \pi_i\alpha_i}, \text{ or}$$

$$v - \delta p - c_a - \pi_a\alpha_a v \geq 0 \Leftrightarrow v \geq \frac{\delta p + c_a}{1 - \pi_a\alpha_a},$$

which can be summarized as

$$v \geq \min\left(c_p + p, \frac{p}{1 - \pi_s\alpha_s u - \pi_i\alpha_i}, \frac{\delta p + c_a}{1 - \pi_a\alpha_a}\right). \tag{2.60}$$

Let $0 < v_1 \leq 1$ and $\sigma^*(v) \in \{(B, P), (B, NP), (B, AP)\}$, then by (2.60), for all $v > v_1$, $\sigma^*(v) \in \{(B, P), (B, NP), (B, AP)\}$, and hence there exists a $\underline{v} \in (0, 1]$ such that a consumer with valuation $v \in \mathcal{V}$ will purchase if and only if $v \geq \underline{v}$.

By (2.59) and (2.60), $\underline{v} \leq v_p$ holds. Moreover, if $\underline{v} < v_p$, consumers with types in $[\underline{v}, v_p]$ choose either (B, NP) or (B, AP). A purchasing consumer with valuation v will prefer (B, NP) over (B, AP) if and only if

$$v - p - \pi_s\alpha_s uv - \pi_i\alpha_i v > v - \delta p - c_a - \pi_a\alpha_a v \Leftrightarrow v\left(\pi_a\alpha_a - \pi_s\alpha_s u - \pi_i\alpha_i\right) > (1-\delta)p - c_a. \tag{2.61}$$

This inequality can be either $v > \frac{(1-\delta)p - c_a}{\pi_a\alpha_a - \pi_s\alpha_s u - \pi_i\alpha_i}$ or $v < \frac{(1-\delta)p - c_a}{\pi_a\alpha_a - \pi_s\alpha_s u - \pi_i\alpha_i}$, depending on the sign of $\pi_a\alpha_a - \pi_s\alpha_s u - \pi_i\alpha_i$. Consequently, there can be two cases for (B, NP) and (B, AP) in equilibrium: first, there exists $v_u \in [\underline{v}, v_p]$ such that $\sigma(v) = (B, NP)$ for all $v \in [v_u, v_p)$, and $\sigma(v) = (B, AP)$ for all $v \in [v_d, v_u)$ where $v_d = \underline{v}$. In the second case, there exists $v_d \in [\underline{v}, v_p]$ such that $\sigma(v) = (B, AP)$ for all $v \in [v_d, v_p)$, and $\sigma(v) = (B, NP)$ for all $v \in [v_u, v_d)$, where $v_u = \underline{v}$. If $\pi_a\alpha_a - \pi_s\alpha_s u - \pi_i\alpha_i = 0$, then depending on the sign of $(1-\delta)p - c_a$, all consumers unilaterally prefer either (B, NP) or (B, AP); e.g., if $(1-\delta)p > c_a$, all consumers prefer (B, AP), and if $(1-\delta)p < c_a$, then all consumers prefer (B, NP). Finally, if $(1-\delta)p = c_a$, then all consumers are indifferent between (B, NP) and (B, AP), in which case only the size of the consumer population u matters in equilibrium, i.e., $\pi_a\alpha_a - \pi_s\alpha_s u - \pi_i\alpha_i = 0$ in equilibrium. Technically, there are multiple equilibria in this case; however, utility of each consumer and the vendor's profit are the same in all equilibria. So, without loss of generality, we focus on the threshold-type equilibrium in this case. In summary, we have established the threshold-type consumer market equilibrium structure.

Next, we characterize in more detail each outcome that can arise in equilibrium, as well as the corresponding parameter regions. For Case (I), in which all consumers who purchase choose the automated patching option, i.e., $0 < v_a < 1$, based on the threshold-type equilibrium structure, we have $u = 0$. We prove the following claim related to the corresponding parameter region in which Case (I) arises.

Claim 3 *The equilibrium that corresponds to case (I) arises if and only if the following conditions are satisfied:*

$$\begin{gathered}\delta p + c_a + \pi_a\alpha_a < 1 \ \textit{and}\ (1-\delta)p + c_p \geq c_a + \pi_a\alpha_a \ \textit{and} \\ (c_a + \delta p)\pi_i\alpha_i \geq c_a + p(-1+\delta+\pi_a\alpha_a) \ \textit{and}\ c_p + (1-\delta)p \geq c_a + (c_p + p)\pi_a\alpha_a \ \textit{and} \\ \left\{ \begin{array}{l} \pi_i\alpha_i < \pi_a\alpha_a \ \textit{and}\ \pi_a\alpha_a - \pi_i\alpha_i + c_a \leq (1-\delta)p, \ \textit{or} \\ \pi_i\alpha_i > \pi_a\alpha_a \ \textit{and}\ c_a + p(-1+\delta+\pi_a\alpha_a) \leq (c_a+\delta p)\pi_i\alpha_i, \ \textit{or} \\ \pi_i\alpha_i = \pi_a\alpha_a \ \textit{and}\ (1-\delta)p - c_a \geq 0 \end{array} \right\}. \end{gathered} \tag{2.62}$$

In this case, the threshold for the consumer indifferent between purchasing the automated patching option and not purchasing at all, v_a, satisfies

$$v_a = \frac{\delta p + c_a}{1 - \pi_a\alpha_a}. \tag{2.63}$$

For this to be an equilibrium, we it is necessary and sufficient to have

$0 < v_a < 1$, type $v = 1$ prefers (B, AP) over (B, P), and all v prefer (B, AP) over (B, NP). Note that type $v = 1$ preferring (B, AP) over (B, P) implies that $v < 1$ does the same, by (2.59). Also, the type $v = v_a$ preferring (NB, NP) over both (B, NP) and (B, P) implies that all $v < v_a$ do the same, by (2.60).

We always have $v_a > 0$ from our model assumptions, namely $\pi_a\alpha_a < 1$. To have $v_a < 1$, a necessary and sufficient condition is $\delta p + c_a + \pi_a\alpha_a < 1$.

For $v = 1$ to weakly prefer (B, AP) over (B, P), a necessary and sufficient condition is $1-\delta p-c_a-\pi_a\alpha_a \geq 1-p-c_p$, which reduces to $(1-\delta)p \geq \pi_a\alpha_a+c_a-c_p$.

The condition for all v to prefer (B, AP) over (B, NP) depends on the magnitude of $\pi_i\alpha_i$. If $\pi_i\alpha_i < \pi_a\alpha_a$, then if $v = 1$ prefers (B, AP) over (B, NP), then all $v < 1$ do too. Therefore, a necessary and sufficient condition is $v - \delta p - c_a - \pi_a\alpha_a v \geq v - p - (\pi_s\alpha_s u(\sigma) + \pi_i\alpha_i)v$ for $v = 1$. With $u(\sigma) = 0$, this becomes $c_a + p\delta + \pi_a\alpha_a \leq p + \pi_i\alpha_i$. So if $\pi_i\alpha_i < \pi_a\alpha_a$, then we need the condition $c_a + p\delta + \pi_a\alpha_a \leq p + \pi_i\alpha_i$.

On the other hand, if $\pi_i\alpha_i > \pi_a\alpha_a$, then $v = v_a$ preferring (B, AP) over (B, NP) implies that all $v > v_a$ do too. Hence, a necessary and sufficient condition is for $v - \delta p - c_a - \pi_a\alpha_a v \geq v - p - (\pi_s\alpha_s u(\sigma) + \pi_i\alpha_i)v$ for $v = v_a$. This simplifies to $c_a + p(-1 + \delta + \pi_a\alpha_a) \leq (c_a + \delta p)\pi_i\alpha_i$.

In the case of $\pi_i\alpha_i = \pi_a\alpha_a$, we need $(1-\delta)p - c_a \geq 0$ for everyone to weakly prefer (B, AP) over (B, NP).

We also need $v = v_a$ to weakly prefer (NB, NP) over both (B, NP) and (B, P), so that all $v < v_a$ do too. We need $0 \geq v - p - \pi_i\alpha_i v$ and $0 \geq v - p - c_p$ for $v = v_a$. These simplify to $(c_a + \delta p)\pi_i\alpha_i \geq c_a + p(-1 + \delta + \pi_a\alpha_a)$ and $c_p + (1-\delta)p \geq c_a + (c_p + p)\pi_a\alpha_a$. Altogether, Case (I) arises if and only if the condition in (2.62) occurs. □

Next, for Case (II), in which there are no consumers choosing (B, P) but the upper tier of consumers is unpatched, i.e., $0 < v_a < v_b < 1$, we have $u = 1 - v_b$. Following the same steps as before, we prove the following claim related to the corresponding conditions for which Case (II) arises.

Claim 4 *The equilibrium that corresponds to Case (II) arises if and only if the*

following conditions are satisfied:

$$(1-\delta)p-c_a > 0 \text{ and } (1-\delta)p < \pi_a\alpha_a-\pi_i\alpha_i+c_a \text{ and } \pi_a\alpha_a-\pi_i\alpha_i+\pi_s\alpha_s > 0 \text{ and }$$
$$\frac{(2\pi_s\alpha_s)(c_a+\delta p)}{1-\pi_a\alpha_a} < -\pi_a\alpha_a+\pi_i\alpha_i+\pi_s\alpha_s+$$
$$\sqrt{(-\pi_a\alpha_a+\pi_i\alpha_i+\pi_s\alpha_s)^2-4\pi_s\alpha_s(c_a+(\delta-1)p)} \text{ and }$$
$$\pi_a\alpha_a+\pi_i\alpha_i+\pi_s\alpha_s \leq \sqrt{(-\pi_a\alpha_a+\pi_i\alpha_i+\pi_s\alpha_s)^2-4\pi_s\alpha_s(c_a+(\delta-1)p)}+2c_p. \tag{2.64}$$

In this case, the threshold for the consumer indifferent between purchasing the automated patching option and not purchasing at all, v_a, again satisfies

$$v_a = \frac{\delta p + c_a}{1-\pi_a\alpha_a}. \tag{2.65}$$

The threshold for the consumer indifferent between being unpatched an purchasing the automated patching option, v_b, satisfies

$$v_b = \frac{(1-\delta)p-c_a}{\pi_a\alpha_a-\pi_s\alpha_s u-\pi_i\alpha_i}. \tag{2.66}$$

Using $u = 1-v_b$, we find that v_b solves a quadratic equation. To find which root is the solution, we note that v_b must satisfy $\pi_a\alpha_a-\pi_s\alpha_s u-\pi_i\alpha_i > 0$ since higher types choose (B, NP) over (B, AP) by (2.61). This implies $(1-\delta)p-c_a > 0$ in order for $v_b > 0$. Using this, we find that the root of the quadratic which specifies v_b is given by

$$v_b = \frac{\pi_s\alpha_s+\pi_i\alpha_i-\pi_a\alpha_a+\sqrt{(\pi_a\alpha_a-\pi_s\alpha_s-\pi_i\alpha_i)^2+4\pi_s\alpha_s((1-\delta)p-c_a)}}{2\pi_s\alpha_s}. \tag{2.67}$$

For this to be an equilibrium, we need $0 < v_a < v_b < 1$ and type $v = 1$ prefers (B, NP) over (B, P). Note that type $v = 1$ preferring (B, NP) over (B, P) implies that $v < 1$ does the same, by (2.59). This also implies that v_b prefers (B, AP) over (B, P), so that $v < v_b$ also prefer (B, AP) over (B, P), again by (2.59). Moreover, type $v = v_a$ preferring (NB, NP) over both (B, NP) and (B, P) implies that all $v < v_a$ do the same, by (2.60).

Again, we always have $v_a > 0$ from our model assumptions, namely $\pi_a\alpha_a < 1$. For $v_b < 1$, an equivalent condition is $(1-\delta)p < c_a + \pi_a\alpha_a - \pi_i\alpha_i$ and $\pi_i\alpha_i < \pi_s\alpha_s + \pi_a\alpha_a$.

For $v_a < v_b$, $\sqrt{(\pi_a\alpha_a - \pi_s\alpha_s - \pi_i\alpha_i)^2 + 4\pi_s\alpha_s((1-\delta)p - c_a)} > (2\pi_s\alpha_s)\left(\frac{c_a+\delta p}{1-\pi_a\alpha_a}\right) + \pi_a\alpha_a - \pi_s\alpha_s - \pi_i\alpha_i$ is necessary and sufficient.

For type $v = 1$ to weakly prefer (B, NP) over (B, P), we equivalently have that $\pi_a\alpha_a + \pi_s\alpha_s + \pi_i\alpha_i - 2c_p \leq \sqrt{(\pi_a\alpha_a - \pi_s\alpha_s - \pi_i\alpha_i)^2 + 4\pi_s\alpha_s((1-\delta)p - c_a)}$. These conditions can all be found in (2.64). □

Next, for case (III), in which all segments are represented and the middle tier is unpatched, i.e., $0 < v_a < v_b < v_p < 1$, we have $u = v_p - v_b$. Following the same steps as before, we prove the following claim related to the corresponding parameter region in which case (III) arises.

Claim 5 *The equilibrium that corresponds to case (III) arises if and only if the following conditions are satisfied:*

$$
\begin{gathered}
(1-\delta)p - c_a \geq 0 \textit{ and } \pi_a\alpha_a > c_p \textit{ and } \pi_s\alpha_s(-\pi_a\alpha_a - c_a + c_p - \delta p + p) < \\
(c_p - \pi_a\alpha_a)(c_p - \pi_i\alpha_i) \textit{ and } \pi_i\alpha_i(c_a + \delta p) + \pi_a\alpha_a c_p > \pi_i\alpha_i(c_p + p) \textit{ and} \\
\Bigg\{\Bigg(c_a + p(\pi_a\alpha_a + \delta) \leq p\Bigg) \textit{ or } \Bigg(c_a + p(\pi_a\alpha_a + \delta) > p \textit{ and} \\
\pi_s\alpha_s(c_a + \delta p)^2(c_a + \pi_a\alpha_a(c_p + p) - c_p + (\delta - 1)p) + \\
(\pi_a\alpha_a - 1)(c_a + p(\pi_a\alpha_a + \delta - 1))(\pi_i\alpha_i(c_a + \delta p) - c_a - p(\pi_a\alpha_a + \delta - 1)) > 0\Bigg)\Bigg\}.
\end{gathered} \tag{2.68}
$$

In this case, the threshold for the consumer indifferent between purchasing the automated patching option and not purchasing at all, v_a, again satisfies

$$v_a = \frac{\delta p + c_a}{1 - \pi_a\alpha_a}. \tag{2.69}$$

To solve for the thresholds v_b and v_p, using $u = v_p - v_b$, note that they solve

$$v_b = \frac{(1-\delta)p - c_a}{\pi_a\alpha_a - \pi_s\alpha_s(v_p - v_b) - \pi_i\alpha_i}, \text{ and} \tag{2.70}$$

$$v_p = \frac{c_p}{\pi_s\alpha_s(v_p - v_b) + \pi_i\alpha_i}. \tag{2.71}$$

From (2.70), we have $\pi_s\alpha_s(v_p - v_b) + \pi_i\alpha_i = \pi_a\alpha_a - \frac{(1-\delta)p-c_a}{v_b}$, while from (2.71), we have $\pi_s\alpha_s(v_p - v_b) + \pi_i\alpha_i = \frac{c_p}{v_p}$. Equating these two expressions and solving for v_p in terms of v_b, we have

$$v_p = \frac{c_p v_b}{c_a - (1-\delta)p + v_b\pi_a\alpha_a}. \tag{2.72}$$

Plugging this expression for v_p into (2.71) and noting that $c_a - (1-\delta)p + v_b\pi_a\alpha_a > 0$ in order for $v_p > 0$, we find that v_b must be a zero of the cubic equation:

$$\begin{aligned} f_1(x) \triangleq (c_a - (1-\delta)p + x\pi_a\alpha_a)(c_a - (1-\delta)p+ \\ x(\pi_a\alpha_a - \pi_i\alpha_i)) + x^2\pi_s\alpha_s(c_a - c_p - (1-\delta)p + x\pi_a\alpha_a). \end{aligned} \tag{2.73}$$

To find which root of the cubic v_b must be, first note that $\pi_a\alpha_a - \pi_s\alpha_s u - \pi_i\alpha_i > 0$ for consumers of higher valuation to prefer (B, NP) over (B, AP) by (2.61). From that, we have that $(1-\delta)p - c_a > 0$ in order for $v_b > 0$. To pin down the root of the cubic, note that the cubic's highest order term is $\pi_s\alpha_s\pi_a\alpha_a x^3$, so $\lim_{x\to-\infty} f_1(x) = -\infty$ and $\lim_{x\to\infty} f_1(x) = \infty$. We find $f_1(0) = ((1-\delta)p - c_a)^2 > 0$ and $f_1\left(\frac{(1-\delta)p-c_a}{\pi_a\alpha_a}\right) = -\frac{c_p\pi_s\alpha_s((1-\delta)p-c_a)^2}{(\pi_a\alpha_a)^2} < 0$, while $0 < \frac{(1-\delta)p-c_a}{\pi_a\alpha_a}$.

We note that from (2.70), we have that $v_b > \frac{(1-\delta)p-c_a}{\pi_a\alpha_a}$, so it follows that v_b is the largest (i.e., most positive) root of the cubic. Then using (2.72), we solve for v_p.

For this to be an equilibrium, a necessary and sufficient condition is $0 < v_a < v_b < v_p < 1$. This tells us that all $v \in [v_p, 1]$ have the same preferences and will purchase (B, P), all $v \in [v_b, v_p)$ have the same preferences and will purchase (B, NP), and all $v \in [v_a, v_b)$ have the same preferences and will purchase (B, AP). Finally, all $v < v_a$ have the same preferences and will not purchase in equilibrium.

For $v_p < 1$, using (2.72), a necessary and sufficient condition for this to hold is $(1-\delta)p - c_a < v_b(\pi_a\alpha_a - c_p)$. Since $(1-\delta)p - c_a > 0$ (again, from $v_b > 0$), we need $\pi_a\alpha_a > c_p$. To have $v_b > \frac{(1-\delta)p-c_a}{\pi_a\alpha_a-c_p}$, a necessary and sufficient condition is that $f_1\left(\frac{(1-\delta)p-c_a}{\pi_a\alpha_a-c_p}\right) < 0$ so that the third root of $f_1(x)$ is greater than $\frac{(1-\delta)p-c_a}{\pi_a\alpha_a-c_p}$. Omitting the algebra, this simplifies to $\pi_s\alpha_s(-\pi_a\alpha_a-c_a+c_p-\delta p+p) < (c_p-\pi_a\alpha_a)(c_p-\pi_i\alpha_i)$.

For $v_b < v_p$, using (2.72), it is equivalent to have $v_b < \frac{(1-\delta)p-c_a+c_p}{\pi_a\alpha_a}$. A necessary and sufficient condition for this is that $f_1\left(\frac{(1-\delta)p-c_a+c_p}{\pi_a\alpha_a}\right) > 0$ so that the third root of $f_1(x)$ is smaller than $\frac{(1-\delta)p-c_a+c_p}{\pi_a\alpha_a}$. This condition becomes $\pi_i\alpha_i(c_a + \delta p) + \pi_a\alpha_a c_p > \pi_i\alpha_i(c_p + p)$.

For $v_a < v_b$, using (2.63), an equivalent condition is $v_b > \frac{\delta p+c_a}{1-\pi_a\alpha_a}$. Since $v_b > \frac{(1-\delta)p-c_A}{\pi_a\alpha_a}$ (by the construction of v_b above as the largest root of the cubic), it follows that if $\frac{(1-\delta)p-c_A}{\pi_a\alpha_a} \geq \frac{\delta p+c_a}{1-\pi_a\alpha_a}$, then we don't need any extra conditions. The condition $\frac{(1-\delta)p-c_A}{\pi_a\alpha_a} \geq \frac{\delta p+c_a}{1-\pi_a\alpha_a}$ simplifies to $(1-\delta-\pi_a\alpha_a)p \geq c_a$. Otherwise, if $(1-\delta-\pi_a\alpha_a)p < c_a$, then we need $f_1\left(\frac{\delta p+c_a}{1-\pi_a\alpha_a}\right) < 0$ for $v_a < v_b$. This condition is $\pi_s\alpha_s(c_a+\delta p)^2(c_a+\pi_a\alpha_a(c_p+p)-c_p+(\delta-1)p)+(\pi_a\alpha_a-1)(c_a+p(\pi_a\alpha_a+\delta-1))(\pi_i\alpha_i(c_a+\delta p)-c_a-p(\pi_a\alpha_a+\delta-1)) > 0$, which is given in (2.68). $\square$

Next, for case (IV), in which all consumers who purchase are unpatched, i.e., $0 < v_b < 1$, we have $u = 1 - v_b$. Following the same steps as before, we prove the following claim related to the corresponding parameter conditions for which case (IV) arises.

Claim 6 *The equilibrium that corresponds to case (IV) arises if and only if the*

following conditions are satisfied:

$$\begin{gathered}
1-2c_p+\pi_i\alpha_i+\pi_s\alpha_s \le \sqrt{4p\pi_s\alpha_s+(1-\pi_s\alpha_s-\pi_i\alpha_i)^2} \textit{ and } \pi_i\alpha_i < 1-p \textit{ and} \\
(1-\pi_a\alpha_a)(1-\pi_i\alpha_i)+(-1+2c_a+2\delta p+\pi_a\alpha_a)\pi_s\alpha_s \ge \\
(1-\pi_a\alpha_a)\sqrt{4p\pi_s\alpha_s+(1-\pi_s\alpha_s-\pi_i\alpha_i)^2} \textit{ and} \\
\Bigg\{\Bigg(1+\pi_i\alpha_i+\pi_s\alpha_s-2\pi_a\alpha_a>0 \textit{ and } p<\frac{(1-\pi_a\alpha_a)(-\pi_a\alpha_a+\pi_i\alpha_i+\pi_s\alpha_s)}{\pi_s\alpha_s} \textit{ and} \\
2(\pi_a\alpha_a+c_a-(1-\delta)p)+\sqrt{(\pi_i\alpha_i+\pi_s\alpha_s-1)^2+4p\pi_s\alpha_s} \ge \pi_i\alpha_i+\pi_s\alpha_s+1\Bigg) \textit{ or} \\
\Bigg(\Bigg(-2\pi_a\alpha_a+\pi_i\alpha_i+\pi_s\alpha_s+1<0 \textit{ or } p>\frac{(1-\pi_a\alpha_a)(-\pi_a\alpha_a+\pi_i\alpha_i+\pi_s\alpha_s)}{\pi_s\alpha_s}\Bigg) \textit{ and} \\
\pi_a\alpha_a(\pi_i\alpha_i+\pi_s\alpha_s)+2\pi_s\alpha_s(c_a+\delta p)+1 \ge \pi_a\alpha_a+\pi_i\alpha_i+\pi_s\alpha_s+ \\
(1-\pi_a\alpha_a)\sqrt{(\pi_i\alpha_i+\pi_s\alpha_s-1)^2+4p\pi_s\alpha_s}\Bigg) \textit{ or} \\
\Bigg(p=\frac{(1-\pi_a\alpha_a)(-\pi_a\alpha_a+\pi_i\alpha_i+\pi_s\alpha_s)}{\pi_s\alpha_s} \textit{ and } (1-\delta)p-c_a\le 0\Bigg)\Bigg\}. \qquad (2.74)
\end{gathered}$$

To solve for the threshold v_b, using $u = 1 - v_b$, we solve

$$v_b = \frac{p}{1-\pi_s\alpha_s(1-v_b)-\pi_i\alpha_i}. \tag{2.75}$$

For this to be an equilibrium, we have that $1-\pi_s\alpha_s u-\pi_i\alpha_i > 0$, otherwise all consumers would prefer (NB, NP) over (B, NP), which can't happen in equilibrium. Using $1-\pi_s\alpha_s(1-v_b)-\pi_i\alpha_i > 0$, we find the right root of the quadratic for v_b to be

$$v_b = \frac{1}{2} + \frac{-1+\pi_i\alpha_i+\sqrt{(1-\pi_s\alpha_s-\pi_i\alpha_i)^2+4p\pi_s\alpha_s}}{2\pi_s\alpha_s}. \tag{2.76}$$

For this to be an equilibrium, the necessary and sufficient conditions are that $0 < v_b < 1$, type $v = 1$ weakly prefers (B, NP) to both (B, AP) over (B, P), and $v = v_b$ weakly prefers (NB, NP) over (B, AP).

For $0 < v_b < 1$, it is equivalent to have $\pi_i\alpha_i < 1 - p$.

For $v = 1$ to prefer (B, NP) over (B, P), we need $1-2c_p+\pi_i\alpha_i+\pi_s\alpha_s \le \sqrt{4p\pi_s\alpha_s+(1-\pi_s\alpha_s-\pi_i\alpha_i)^2}$.

For $v = v_b$ to weakly prefer (NB, NP) over (B, AP), we need $0 \geq v_b - \delta p - c_a - \pi_a\alpha_a v_b$. This simplifies to $(1-\pi_a\alpha_a)(1-\pi_i\alpha_i)+(-1+2c_a+2\delta p+\pi_a\alpha_a)\pi_s\alpha_s \geq (1-\pi_a\alpha_a)\sqrt{4p\pi_s\alpha_s + (1-\pi_s\alpha_s-\pi_i\alpha_i)^2}$.

For everyone to prefer (B, NP) over (B, AP), the condition needed depends on whether $u(\sigma) > \frac{\pi_a\alpha_a-\pi_i\alpha_i}{\pi_s\alpha_s}$, as seen in (2.61). If $u(\sigma) > \frac{\pi_a\alpha_a-\pi_i\alpha_i}{\pi_s\alpha_s}$, then lower valuation consumers would prefer (B, NP) over (B, AP) so that a sufficient condition for everyone to prefer (B, NP) over (B, AP) is that $v = 1$ weakly prefers (B, NP) over (B, AP). On the other hand, if $u(\sigma) < \frac{\pi_a\alpha_a-\pi_i\alpha_i}{\pi_s\alpha_s}$, then higher valuation consumers prefer (B, NP) over (B, AP) so that the condition would be $v = v_b$ weakly prefers (B, NP) over (B, AP).

The condition $u(\sigma) > \frac{\pi_a\alpha_a-\pi_i\alpha_i}{\pi_s\alpha_s}$ is equivalent to $1+\pi_i\alpha_i+\pi_s\alpha_s-2\pi_a\alpha_a > 0$ and $p < \frac{(1-\pi_a\alpha_a)(-\pi_a\alpha_a+\pi_i\alpha_i+\pi_s\alpha_s)}{\pi_s\alpha_s}$.

The condition that $v = 1$ weakly prefers (B, NP) over (B, AP) is $v - p - ((1-v_b)\pi_s\alpha_s v + \pi_i\alpha_i v) \geq v - \delta p - c_a - \pi_a\alpha_a v$ for $v = v_b$. This simplifies to $2(\pi_a\alpha_a + c_a - (1-\delta)p) + \sqrt{(\pi_i\alpha_i + \pi_s\alpha_s - 1)^2 + 4p\pi_s\alpha_s} \geq \pi_i\alpha_i + \pi_s\alpha_s + 1$.

The condition that $v = v_b$ weakly prefers (B, NP) over (B, AP) is $\pi_a\alpha_a(\pi_i\alpha_i + \pi_s\alpha_s) + 2\pi_s\alpha_s(c_a + \delta p) + 1 \geq \pi_a\alpha_a + \pi_i\alpha_i + \pi_s\alpha_s + (1-\pi_a\alpha_a)\big((\pi_i\alpha_i + \pi_s\alpha_s - 1)^2 + 4p\pi_s\alpha_s\big)^{\frac{1}{2}}$.

Lastly, if $u(\sigma) = \frac{\pi_a\alpha_a-\pi_i\alpha_i}{\pi_s\alpha_s}$, then everyone will prefer (B, NP) over (B, AP) if $(1-\delta)p - c_a \leq 0$. The conditions of these subcases are given in (2.74).

Next, for case (V), in which the lower tier of purchasing consumers is unpatched while the upper tier does automated patching, i.e., $0 < v_b < v_a < 1$, we have $u = v_a - v_b$. Following the same steps as before, we prove the following claim related to the corresponding parameter region in which case (V) arises.

Claim 7 *The equilibrium that corresponds to case (V) arises if and only if the following conditions are satisfied:*

$$(1-\delta)p - c_a < 0 \textit{ and } (\pi_a\alpha_a + c_a - 1 - (1-\delta)p)(\pi_a\alpha_a - \pi_i\alpha_i + c_a - (1-\delta)p) > \pi_s\alpha_s(\pi_a\alpha_a + c_a + \delta p - 1) \textit{ and }$$
$$\pi_i\alpha_i(c_a + \delta p) < c_a + p(-1 + \delta + \pi_a\alpha_a) \textit{ and } \pi_a\alpha_a \leq (1-\delta)p + c_p - c_a. \quad (2.77)$$

To solve for the thresholds v_b and v_a, using $u = v_a - v_b$, note that they solve

$$v_b = \frac{p}{1 - \pi_s\alpha_s(v_a - v_b) - \pi_i\alpha_i}, \text{ and} \tag{2.78}$$

$$v_a = \frac{(1-\delta)p - c_a}{\pi_a\alpha_a - \pi_s\alpha_s(v_a - v_b) - \pi_i\alpha_i}. \tag{2.79}$$

From (2.78), we have $\pi_s\alpha_s(v_a - v_b) + \pi_i\alpha_i = 1 - \frac{p}{v_b}$, while from (2.79), we have $\pi_s\alpha_s(v_a - v_b) + \pi_i\alpha_i = \pi_a\alpha_a - \frac{(1-\delta)p - c_a}{v_a}$. Equating these two expressions and solving for v_a in terms of v_b, we have

$$v_a = \frac{v_b(-c_a + (1-\delta)p)}{p - v_b(1 - \pi_a\alpha_a)}. \tag{2.80}$$

Plugging this expression for v_a into (2.78), we find that v_b must be a zero of the cubic equation:

$$\begin{aligned} f_2(x) \triangleq (1 - \pi_a\alpha_a)\pi_s\alpha_s x^3 + ((1 - \pi_a\alpha_a)(1 - \pi_i\alpha_i) - c_a\pi_s\alpha_s - \delta p\pi_s\alpha_s)x^2 - \\ p(2 - \pi_a\alpha_a - \pi_i\alpha_i)x + p^2. \end{aligned} \tag{2.81}$$

To find which root of the cubic v_b must be, first note that $\pi_a\alpha_a - \pi_s\alpha_s u - \pi_i\alpha_i < 0$ for consumers of higher valuation to prefer (B, NP) over (B, AP) by (2.61). From that, we have that $c_a - (1-\delta)p > 0$ in order for $v_a > 0$. To pin down the root of the cubic, note that the cubic's highest order term is $\pi_s\alpha_s(1 - \pi_a\alpha_a)x^3$, so $\lim_{x\to-\infty} f_2(x) = -\infty$ and $\lim_{x\to\infty} f_2(x) = \infty$. We find $f_2(0) = p^2 > 0$ and $f_2\left(\frac{p}{1-\pi_a\alpha_a}\right) = -\frac{p^2\pi_s\alpha_s(c_a - (1-\delta)p)}{(1-\pi_a\alpha_a)^2} < 0$. Since $\lim_{x\to\infty} f_2(x) = \infty$, there exists a root larger than $\frac{p}{1-\pi_a\alpha_a}$. Note that from (2.78), we have $v_b > \frac{p}{1-\pi_a\alpha_a}$. Therefore, v_b is the largest root of the cubic, lying past $\frac{p}{1-\pi_a\alpha_a}$. Then using (2.80), we solve for v_a.

For this to be an equilibrium, the necessary and sufficient conditions are $0 < v_b < v_a < 1$ and type $v = 1$ prefers (B, AP) over (B, P). Type $v = 1$ preferring (B, AP) over (B, P) ensures $v < 1$ does so too, by (2.59). Moreover, since type $v = v_a$ is indifferent between (B, AP) and (B, NP), and since (B, AP) is preferred over (B, P), by transitivity, it follows that type v_a prefers (B, NP) over (B, P). It follows that $v < v_a$ prefers (B, NP) over (B, P) as well by (2.59).

For $v_a < 1$, using (2.80), an equivalent condition for this to hold is $p + v_b(-1 + c_a - p(1-\delta) + \pi_a\alpha_a) < 0$. If $c_a - p(1-\delta) \geq 1 - \pi_a\alpha_a$, then this case can't happen. Otherwise, if $c_a - p(1-\delta) < 1 - \pi_a\alpha_a$, then this condition becomes $v_b > \frac{p}{1-c_a+p(1-\delta)-\pi_a\alpha_a}$. This is equivalent to $f_2(\frac{p}{1-c_a+p(1-\delta)-\pi_a\alpha_a}) > 0$, which simplifies to $(\pi_a\alpha_a + c_a - 1 - (1-\delta)p)(\pi_a\alpha_a - \pi_i\alpha_i + c_a - (1-\delta)p) > \pi_s\alpha_s(\pi_a\alpha_a + c_a + \delta p - 1)$.

For $v_a > v_b$, using (2.80), it is equivalent to require $v_b < \frac{c_a+p\delta}{1-\pi_a\alpha_a}$. For this to happen, we need the condition $f_2(\frac{c_a+p\delta}{1-\pi_a\alpha_a}) > 0$, which simplifies to $\pi_i\alpha_i(c_a + \delta p) < c_a + p(-1 + \delta + \pi_a\alpha_a)$.

For $v_b > 0$, this holds by construction of v_b as the largest root of $f_2(x)$ (which was shown to be larger than $\frac{p}{1-\pi_a\alpha_a} > 0$), so no additional conditions are needed.

Finally, for type $v = 1$ to prefer (B, AP) over (B, P), a necessary and sufficient condition is $\pi_a\alpha_a \leq (1-\delta)p + c_p - c_a$. The conditions above are summarized in (2.77). $\square$

Next, for case (VI), in which all segments are represented and the middle tier does automated patching, i.e., $0 < v_b < v_a < v_p < 1$, we have $u = v_a - v_b$. Following the same steps as before, we prove the following claim related to the corresponding parameter region in which case (VI) arises.

Claim 8 *The equilibrium that corresponds to case (VI) arises if and only if the following conditions are satisfied:*

$$(1-\delta)p - c_a < 0 \text{ and } c_p + (1-\delta)p < c_a + \pi_a\alpha_a \text{ and } c_p(1-\pi_a\alpha_a) + (1-\delta)p > c_a \text{ and }$$
$$c_p(\pi_a\alpha_a)^2(-c_a + (1-\delta)p + c_p(1-\pi_a\alpha_a)) + \pi_s\alpha_s(-c_a + (1-\delta)p + c_p)^2(c_a - c_p(1-\pi_a\alpha_a) +$$
$$(\pi_a\alpha_a - 1 + \delta)p) - (c_a - c_p - (1-\delta)p)(c_a - (1-\delta)p - c_p(1-\pi_a\alpha_a))\pi_a\alpha_a\pi_i\alpha_i < 0 \text{ and }$$
$$c_a + p(-1 + \delta + \pi_a\alpha_a) > (c_a + \delta p)\pi_i\alpha_i. \quad (2.82)$$

To solve for the thresholds v_b and v_a, using $u = v_a - v_b$, note that they solve

$$v_b = \frac{p}{1 - \pi_s\alpha_s(v_a - v_b) - \pi_i\alpha_i}, \text{ and} \quad (2.83)$$

$$v_a = \frac{(1-\delta)p - c_a}{\pi_a\alpha_a - \pi_s\alpha_s(v_a - v_b) - \pi_i\alpha_i}. \tag{2.84}$$

These are the same as (2.78) and (2.79). Using the exact same argument, it follows that v_b is the largest root of the cubic $f_2(x)$, lying past $\frac{c_a-(1-\delta)p}{1-\pi_a\alpha_a}$. Note that the largest root v_b is indeed larger $\frac{p}{1-\pi_a\alpha_a}$ in this case as well since $v_b = \frac{pv_a}{-c_a+p(1-\delta)+v_a(1-\pi_a\alpha_a)}$ and $(1-\delta)p - c_a < 0$. Then using (2.80), we solve for v_a.

In this case, however, we also have a standard patching population, with the standard patching threshold given by $v_p = \frac{c_p-(c_a-(1-\delta)p)}{\pi_a\alpha_a}$.

For this to be an equilibrium, the necessary and sufficient conditions are $0 < v_b < v_a < v_p < 1$. This tells us that all $v \in [v_p, 1]$ have the same preferences and will purchase (B, P), all $v \in [v_a, v_p)$ have the same preferences and will purchase (B, AP), and all $v \in [v_b, v_a)$ have the same preferences and will purchase (B, NP). Finally, all $v < v_b$ have the same preferences and will not purchase in equilibrium.

For $v_p < 1$, the necessary and sufficient condition is $c_p+(1-\delta)p < c_a+\pi_a\alpha_a$.

For $v_a < v_p$, using (2.80) to write v_a in terms of v_b, it is equivalent to write $v_b((1-\delta)p - c_a + c_p(1-\pi_a\alpha_a)) > p((1-\delta)p - c_a + c_p)$.

If $-c_a + c_p(1-\pi_a\alpha_a) + p(1-\delta) > 0$, then we can rewrite this as $v_b > \frac{p((1-\delta)p-c_a+c_p)}{(1-\delta)p-c_a+c_p(1-\pi_a\alpha_a)}$. This is equivalent to $f_2(\frac{p((1-\delta)p-c_a+c_p)}{(1-\delta)p-c_a+c_p(1-\pi_a\alpha_a)}) < 0$, which simplifies to $c_p(\pi_a\alpha_a)^2(-c_a+(1-\delta)p+c_p(1-\pi_a\alpha_a))+\pi_s\alpha_s(-c_a+(1-\delta)p+c_p)^2(c_a-c_p(1-\pi_a\alpha_a)+(\pi_a\alpha_a-1+\delta)p)-(c_a-c_p-(1-\delta)p)(c_a-(1-\delta)p-c_p(1-\pi_a\alpha_a))\pi_a\alpha_a\pi_i\alpha_i < 0$.

On the other hand, if $-c_a + c_p(1-\pi_a\alpha_a) + p(1-\delta) < 0$, then we need $v_b < \frac{p((1-\delta)p-c_a+c_p)}{(1-\delta)p-c_a+c_p(1-\pi_a\alpha_a)}$. If $p((1-\delta)p - c_a + c_p) \geq 0$, then this can't happen since the denominator is negative and $v_b > 0$. If $p((1-\delta)p - c_a + c_p) < 0$, then $\frac{p((1-\delta)p-c_a+c_p)}{(1-\delta)p-c_a+c_p(1-\pi_a\alpha_a)} < p$. However, $v_b > \frac{p}{1-\pi_a\alpha_a}$, so this can't happen either. Therefore, $-c_a + c_p(1-\pi_a\alpha_a) + p(1-\delta) < 0$ rules out this case.

Lastly, if $-c_a+c_p(1-\pi_a\alpha_a)+p(1-\delta) = 0$, then this $v_b((1-\delta)p-c_a+c_p(1-\pi_a\alpha_a)) > p((1-\delta)p - c_a + c_p)$ becomes $0 > ((1-\delta)p - c_a + c_p)$. This simplifies to $0 < c_p\pi_a\alpha_a$, which can't happen.

Therefore, the conditions for $v_a < v_p$ are $c_p(1-\pi_a\alpha_a) + (1-\delta)p > c_a$ and $c_p(\pi_a\alpha_a)^2(-c_a + (1-\delta)p + c_p(1-\pi_a\alpha_a)) + \pi_s\alpha_s(-c_a + (1-\delta)p + c_p)^2(c_a - c_p(1-\pi_a\alpha_a)+(\pi_a\alpha_a-1+\delta)p)-(c_a-c_p-(1-\delta)p)(c_a-(1-\delta)p-c_p(1-\pi_a\alpha_a))\pi_a\alpha_a\pi_i\alpha_i < 0$.

For $v_b < v_a$, again using (2.80) to write v_a in terms of v_b, this simplifies to $v_b < \frac{c_a+p\delta}{1-\pi_a\alpha_a}$. This can equivalently be expressed as $f_2\left(\frac{c_a+p\delta}{1-\pi_a\alpha_a}\right) > 0$, which simplifies to $c_a + p(-1+\delta+\pi_a\alpha_a) > (c_a+\delta p)\pi_i\alpha_i$.

For $0 < v_b$, no condition is needed since v_b is defined to be the largest root of the cubic, which was shown to be larger than $\frac{p}{1-\pi_a\alpha_a}$.

A summary of the above necessary and sufficient conditions is given in (2.82). □

Next, for case (VII), in which there are no automated patching users while the lower tier is unpatched and the upper tier is patched, i.e., $0 < v_b < v_p < 1$, we have $u = v_p - v_b$. Following the same steps as before, we prove the following claim related to the corresponding parameter region in which case (VII) arises.

Claim 9 *The equilibrium that corresponds to case (VII) arises if and only if the following conditions are satisfied:*

$$\begin{gathered}(-1+c_p+p)\pi_s\alpha_s < (1-c_p)(-c_p+\pi_i\alpha_i) \text{ and } \pi_i\alpha_i < \frac{c_p}{c_p+p} \text{ and} \\ (1-\pi_a\alpha_a)(c_a+(\pi_a\alpha_a-(1-\delta))p)(c_a+p(\pi_a\alpha_a-(1-\delta))-(c_a+\delta p)\pi_i\alpha_i)+ \\ (c_a+\delta p)^2(c_a-c_p-(1-\delta)p+(c_p+p)\pi_a\alpha_a)\pi_s\alpha_s \geq 0 \text{ and } c_a+p(\delta+\pi_a\alpha_a) > p \text{ and} \\ \left\{\left(c_a-(1-\delta)p+c_p(-1+\pi_a\alpha_a)\geq 0\right) \text{ or}\right. \\ \Big(c_a+c_p(-1+\pi_a\alpha_a) < 0 \text{ and } \pi_a\alpha_a(c_a-(1-\delta)p+c_p(-1+\pi_a\alpha_a))(c_p\pi_a\alpha_a+ \\ (c_a-c_p-(1-\delta)p)\pi_i\alpha_i) \leq (-c_a+c_p+(1-\delta)p)^2(c_a-c_p- \\ \left.(1-\delta)p+(c_p+p)\pi_a\alpha_a)\pi_s\alpha_s\Big)\right\}. \end{gathered} \tag{2.85}$$

To solve for the thresholds v_b and v_p, using $u = v_p - v_b$, note that they solve

$$v_b = \frac{p}{1-\pi_s\alpha_s(v_p-v_b)-\pi_i\alpha_i}, \text{ and} \tag{2.86}$$

$$v_p = \frac{c_p}{\pi_s\alpha_s(v_p-v_b)+\pi_i\alpha_i}. \tag{2.87}$$

From (2.86), we have $\pi_s\alpha_s(v_p-v_b)+\pi_i\alpha_i = 1-\frac{p}{v_b}$, while from (2.87), we have $\pi_s\alpha_s(v_p - v_b) + \pi_i\alpha_i = \frac{c_p}{v_p}$. Equating these two expressions and solving for v_p in terms of v_b, we have

$$v_p = \frac{c_p v_b}{v_b - p}. \tag{2.88}$$

Plugging this expression for v_p into (2.86), we find that v_b must be a zero of the cubic equation:

$$f_3(x) \triangleq \pi_s\alpha_s x^3 + (1 - \pi_i\alpha_i - \pi_s\alpha_s(c_p + p))x^2 - (2 - \pi_i\alpha_i)px + p^2. \tag{2.89}$$

To find which root of the cubic v_b must be, note that the cubic's highest order term is $\pi_s\alpha_s x^3$, so $\lim_{x\to-\infty} f_3(x) = -\infty$ and $\lim_{x\to\infty} f_3(x) = \infty$. We find $f_3(0) = p^2 > 0$, and $f_3(p) = -c_p\pi_s\alpha_s p^2 < 0$. Since $v_b - p > 0$ in equilibrium, we have that v_b is the largest root of the cubic, lying past p. Then using (2.88), we solve for v_p.

For this to be an equilibrium, the necessary and sufficient conditions are $0 < v_b < v_p < 1$, type $v = v_p$ prefers (B, P) over (B, AP), and type $v = v_b$ prefers (NB, NP) to (B, AP). Type $v = v_p$ preferring (B, P) over (B, AP) ensures $v > v_p$ also prefer (B, P) over (B, AP), by (2.59). Moreover, type $v = v_b$ preferring (NB, NP) over (B, AP) ensures $v < v_b$ do so too, by (2.60).

For $v_p < 1$, a necessary and sufficient condition for this to hold is $v_b > \frac{p}{1-c_p}$. This is equivalent to $f_3(\frac{p}{1-c_p}) < 0$. This simplifies to $(-1 + c_p + p)\pi_s\alpha_s < (1 - c_p)(-c_p + \pi_i\alpha_i)$.

For $v_p > v_b$, a necessary and sufficient condition is $v_b < c_p + p$. This is equivalent to $f_3(c_p + p) > 0$, or $\pi_i\alpha_i < \frac{c_p}{c_p+p}$.

We don't need any conditions for $v_b > 0$, since by construction, $v_b > p > 0$.

For type $v = v_p$ to prefer (B, P) over (B, AP), a necessary and sufficient condition is $\frac{c_p v_b}{v_b-p} \geq \frac{c_p-(c_a-(1-\delta)p)}{\pi_a\alpha_a}$. This simplifies to $v_b(c_a-(1-\delta)p-c_p(1-\pi_a\alpha_a)) \geq p(c_a - (1 - \delta)p - c_p)$. This can be broken down into three cases, depending on the sign of $c_a-(1-\delta)p-c_p(1-\pi_a\alpha_a)$ (also considering the case when the factor is zero). When $c_a-(1-\delta)p-c_p(1-\pi_a\alpha_a) = 0$, the left side is 0 while the right side is negative, so this inequality holds. If $c_a - (1 - \delta)p - c_p(1 - \pi_a\alpha_a) > 0$, then the inequality becomes $v_b \geq p\frac{c_a-(1-\delta)p-c_p}{c_a-(1-\delta)p-c_p(1-\pi_a\alpha_a)}$. But $\frac{c_a-(1-\delta)p-c_p}{c_a-(1-\delta)p-c_p(1-\pi_a\alpha_a)} < 1$, and since $v_b > p$ by construction, this inequality holds without further conditions. On the other hand, if $c_a - (1-\delta)p - c_p(1-\pi_a\alpha_a) < 0$, then we need $v_b \leq p\frac{c_a-(1-\delta)p-c_p}{c_a-(1-\delta)p-c_p(1-\pi_a\alpha_a)}$. So we need $f_3\left(\frac{p(c_a-(1-\delta)p-c_p)}{c_a-(1-\delta)p-c_p(1-\pi_a\alpha_a)}\right) \geq 0$. Omitting the algebra, this simplifies to $\pi_a\alpha_a(c_a-(1-\delta)p+c_p(-1+\pi_a\alpha_a))(c_p\pi_a\alpha_a+(c_a-c_p-(1-\delta)p)\pi_i\alpha_i) \leq (-c_a+c_p+(1-\delta)p)^2(c_a-c_p-(1-\delta)p+(c_p+p)\pi_a\alpha_a)\pi_s\alpha_s$.

For $v = v_b$ to prefer (NB, NP) to (B, AP), a necessary and sufficient condition is $v_b \leq \frac{\delta p+c_a}{1-\pi_a\alpha_a}$, which becomes $f_3\left(\frac{\delta p+c_a}{1-\pi_a\alpha_a}\right) \geq 0$. This simplifies to $c_a > (1-\delta-\pi_a\alpha_a)p$ and $(1-\pi_a\alpha_a)(c_a+(\pi_a\alpha_a-(1-\delta))p)(c_a+p(\pi_a\alpha_a-(1-\delta))-(c_a+\delta p)\pi_i\alpha_i)+(c_a+\delta p)^2(c_a-c_p-(1-\delta)p+(c_p+p)\pi_a\alpha_a)\pi_s\alpha_s \geq 0$. The conditions are summarized in (2.85). $\square$

Next, for case (VIII), in which there are no unpatched users while the lower tier chooses automated patching and the upper tier chooses standard patching, i.e., $0 < v_a < v_p < 1$, we have $u = 0$. Following the same steps as before, we prove the following claim related to the corresponding parameter region in which case (VIII) arises.

Claim 10 *The equilibrium that corresponds to case (VIII) arises if and only if the following conditions are satisfied:*

$$
\begin{gathered}
c_p+(1-\delta)p < c_a+\pi_a\alpha_a \text{ and } c_p+(1-\delta)p > c_a+(c_p+p)\pi_a\alpha_a \text{ and} \\
(c_p-c_a+(1-\delta)p)\pi_i\alpha_i \geq c_p\pi_a\alpha_a \text{ and} \\
\Big\{\Big(\pi_i\alpha_i < \pi_a\alpha_a \text{ and } c_p\pi_a\alpha_a+\pi_i\alpha_i(c_a-c_p-(1-\delta)p) \leq 0\Big) \text{ or} \\
\Big(\pi_i\alpha_i > \pi_a\alpha_a \text{ and } c_a+p(-1+\delta+\pi_a\alpha_a) \leq (c_a+\delta p)\pi_i\alpha_i\Big) \text{ or} \\
\Big(\pi_i\alpha_i = \pi_i\alpha_i \text{ and } (1-\delta)p-c_a \geq 0\Big)\Big\}.
\end{gathered} \tag{2.90}
$$

In this case, the threshold for the consumer indifferent between purchasing the automated patching option and not purchasing at all, v_a, satisfies

$$v_a = \frac{\delta p+c_a}{1-\pi_a\alpha_a}, \tag{2.91}$$

and the consumer indifferent between choosing automated patching and standard patching is given by

$$v_p = \frac{(1-\delta)p + c_p - c_a}{\pi_a \alpha_a}, \tag{2.92}$$

For this to be an equilibrium, it is necessary and sufficient to have $0 < v_a < v_p < 1$, no one prefers (B, NP) over (B, AP), and no one prefers (B, NP) over (B, P).

For $v_p < 1$, this is equivalently written as $c_p + (1-\delta)p < c_a + \pi_a \alpha_a$.

To have $v_a < v_p$, a necessary and sufficient condition is $c_p + (1-\delta)p > c_a + (c_p + p)\pi_a \alpha_a$.

We always have $v_a > 0$ from our model assumptions, namely $\pi_a \alpha_a < 1$.

To ensure that no one prefers (B, NP) over (B, P), it suffices to make $v = v_p$ weakly prefer (B, P) over (B, NP) so that everyone of higher valuation would also have the same preference (by (2.59)). This condition then becomes $(c_p - c_a + (1-\delta)p)\pi_i \alpha_i \geq c_p \pi_a \alpha_a$.

To ensure that no one prefers (B, NP) over (B, AP), we need $(\pi_a \alpha_a - (\pi_s \alpha_s u(\sigma) + \pi_i \alpha_i))v \leq (1-\delta)p - c_a$. In this case, $u(\sigma) = 0$ so that there are three cases, depending on the sign of $\pi_a \alpha_a - \pi_i \alpha_i$.

If $\pi_a \alpha_a > \pi_i \alpha_i$, then higher valuation consumers prefer (B, NP) over (B, AP), so a necessary and sufficient condition is for $v = v_p$ to weakly prefer (B, AP) over (B, NP). This becomes $c_p \pi_a \alpha_a + \pi_i \alpha_i (c_a - c_p - (1-\delta)p) \leq 0$.

On the other hand, if $\pi_a \alpha_a < \pi_i \alpha_i$, then lower valuation consumers prefer (B, NP) over (B, AP). In this case, a necessary and sufficient for no one to prefer (B, NP) over (B, AP) is for $v = v_a$ to weakly prefer (B, AP) over (B, NP). This simplifies to $c_a + p(-1 + \delta + \pi_a \alpha_a) \leq (c_a + \delta p)\pi_i \alpha_i$.

Lastly, if $\pi_a \alpha_a = \pi_i \alpha_i$, then we need $(1-\delta)p - c_a \geq 0$ for everyone to prefer (B, AP) over (B, NP). Altogether, Case (VIII) arises if and only if the condition in (2.90) occurs. $\square$

Lastly, for case (IX), in which all users choose standard patching, i.e., $0 < v_p < 1$, we have $u = 0$. Following the same steps as before, we prove the following claim related to the corresponding parameter region in which case (IX) arises.

Claim 11 *The equilibrium that corresponds to case (IX) arises if and only if the*

following conditions are satisfied:

$$c_p + p < 1 \textit{ and } c_a + (c_p + p)\pi_a\alpha_a \geq c_p + (1-\delta)p \textit{ and } (c_p + p)\pi_i\alpha_i \geq c_p. \quad (2.93)$$

In this case, the threshold for the consumer indifferent between choosing the standard patching option and not purchasing at all, v_p, satisfies

$$v_p = c_p + p. \quad (2.94)$$

For this to be an equilibrium, it is necessary and sufficient to have $0 < v_p < 1$, $v = v_p$ prefers (NB, NP) over both (B, NP) and (B, AP), and $v = v_p$ prefers (B, P) over both (B, NP) and (B, AP).

For $v_p < 1$, this is equivalently written as $c_p + p < 1$.

For $v = v_p$ to weakly prefer (NB, NP) over (B, NP), we need $0 \geq v_p - p - \pi_i\alpha_i v_p$, which simplifies to $(c_p + p)\pi_i\alpha_i \geq c_p$.

For $v = v_p$ to weakly prefer (NB, NP) over (B, AP), we need $c_a + (c_p + p)\pi_a\alpha_a \geq c_p + (1-\delta)p$.

For $v = v_p$ to weakly prefer (B, P) over (B, NP) and (B, AP), the conditions will be the same as above since $v = v_p$ is indifferent between (NB, NP) and (B, P).

Altogether, Case (IX) arises if and only if the condition in (2.93) occurs.

This completes the proof of the general consumer market equilibrium for the proprietary case. $\square$

Proof of Lemma 5: Technically, we prove that there exists an $\tilde{\alpha}_2$ such that if $\pi_i\alpha_i < \min\left[\frac{c_p\pi_a\alpha_a}{1+c_p-c_a}, \frac{c_p}{1+c_p}\right]$, then for $\alpha_s > \tilde{\alpha}_2$, p^* and δ^* are set so that

1. if $c_a < \min\left[\pi_a\alpha_a - c_p, c_p(1-\pi_a\alpha_a)\right]$, then $\sigma^*(v)$ is characterized by $0 < v_a < v_b < v_p < 1$ under optimal pricing,

2. if $|\pi_a\alpha_a - c_p| < c_a < c_p(1-\pi_a\alpha_a)$, then $\sigma^*(v)$ is characterized by $0 < v_b < v_a < v_p < 1$ under optimal pricing, and if

3. if $c_a > c_p(1-\pi_a\alpha_a)$, then $\sigma^*(v)$ is characterized by $0 < v_b < v_p < 1$ under optimal pricing.

The sketch of the proof is as follows. From Lemma A.4, a unique consumer market equilibrium arises, given a price p and discount δ. Within each region of the parameter space defined by Lemma A.4, the thresholds v_a, v_b, and v_p are smooth functions of the parameters, including p and δ. In the cases where the thresholds are given in closed-form, this is clear. In the cases where these thresholds are implicitly defined as the root of some cubic equation, then the smoothness of the thresholds in the parameters follows from the Implicit Function Theorem. Specifically, for each of those cases, the threshold defined was the most positive root v_b^* of a cubic function of v_b, $f(v_b, p, \delta) = 0$. Moreover, the cubic $f(v_b, p, \delta)$ has two local extrema in v_b and is negative to the left of v_b^* and positive to the right of it ($f(v_b^* - \epsilon, p, \delta) < 0$ and $f(v_b^* + \epsilon, p, \delta) > 0$ for arbitrarily small $\epsilon > 0$). Therefore, $\frac{\partial f}{\partial v_b}(v_b, p, \delta) \neq 0$ so that the Implicit Function Theorem applies. The thresholds being smooth in p and δ implies that the profit function for each case of the parameter space defined by Lemma A.4 is smooth in p and δ. We find the profit-maximizing price within the compact closure of each case, so that the price that induces the largest profit among the cases will be the equilibrium price set by the vendor.

Having given the sketch of the proof, we now proceed with the proof. The conditions of this lemma precludes candidate market structures from arising in equilibrium. Specifically, using Lemma A.4, $\pi_i\alpha_i < \min\left[\pi_a\alpha_a, \frac{c_p}{1+c_p}\right]$ rules out Cases (VIII) and (IX). We consider the remaining consumer structures that can arise when the vendor sets prices optimally.

Suppose $0 < v_a < 1$ is induced. By part (I) of Lemma A.4, we obtain $v_a = \frac{\delta p + c_a}{1-\pi_a\alpha_a}$. The profit function in this case is $\Pi_I(p,\delta) = \delta p(1 - v_a(p,\delta))$. Let C_I be the compact closure of the region of the parameter space defining $0 < v_a < 1$, given in part (I) of Lemma A.4. By the Weierstrass extreme value theorem, there exists p and δ in C_I that maximizes $\Pi(p,\delta)$. This p and δ combination may be on the boundary, and we show that the vendor's profit function is continuous across region boundaries later. Otherwise, if this p and δ are interior, the unconstrained maximizer satisfies the first-order conditions.

Differentiating the profit function with respect to p, we have that $p_I^*(\delta) = \frac{1-c_a-\pi_a\alpha_a}{2\delta}$. The second-order condition gives $\frac{\partial^2}{\partial p^2}\Pi(p,\delta) = -\frac{2\delta^2}{1-\pi_a\alpha_a}$. We see that

$$\Pi_I \triangleq \Pi(p^*(\delta),\delta) = \frac{(1-c_a-\pi_a\alpha_a)^2}{4(1-\pi_a\alpha_a)} \tag{2.95}$$

for any δ, so this is the maximal profit of this case.

On the other hand, suppose $0 < v_a < v_b < 1$ is induced. By part (II) of Lemma A.4, we obtain that $v_a = \frac{\delta p + c_a}{1-\pi_a\alpha_a}$ and $v_b = \frac{-\pi_a\alpha_a+\pi_i\alpha_i+\pi_s\alpha_s+\sqrt{(-\pi_a\alpha_a+\pi_i\alpha_i+\pi_s\alpha_s)^2-4\pi_s\alpha_s(c_a+(\delta-1)p)}}{2\pi_s\alpha_s}$. The profit function in this case is

$$\Pi_{II}(p,\delta) = p(1-v_b(p,\delta)) + \delta p(v_b(p,\delta) - v_a(p,\delta)).$$

The first-order condition in p yields

$$\begin{gathered}
\frac{(\delta-1)p}{\sqrt{(-\pi_a\alpha_a+\pi_i\alpha_i+\pi_s\alpha_s)^2-4\pi_s\alpha_s(c_a+(\delta-1)p)}} + \\
\delta\Bigg(\frac{-\pi_a\alpha_a+\pi_i\alpha_i+\pi_s\alpha_s+\sqrt{(-\pi_a\alpha_a+\pi_i\alpha_i+\pi_s\alpha_s)^2-4\pi_s\alpha_s(c_a+(\delta-1)p)}}{2\pi_s\alpha_s} - \\
\frac{c_a+\delta p}{1-\pi_a\alpha_a}\Bigg) + \\
\delta p\left(-\frac{\delta}{1-\pi_a\alpha_a} - \frac{\delta-1}{\sqrt{(-\pi_a\alpha_a+\pi_i\alpha_i+\pi_s\alpha_s)^2-4\pi_s\alpha_s(c_a+(\delta-1)p)}}\right) - \\
\frac{-\pi_a\alpha_a+\pi_i\alpha_i+\pi_s\alpha_s+\sqrt{(-\pi_a\alpha_a+\pi_i\alpha_i+\pi_s\alpha_s)^2-4\pi_s\alpha_s(c_a+(\delta-1)p)}}{2\pi_s\alpha_s} + 1 = 0.
\end{gathered} \tag{2.96}$$

Letting $X = \sqrt{(-\pi_a\alpha_a + \pi_i\alpha_i + \pi_s\alpha_s)^2 - 4\pi_s\alpha_s(c_a + (\delta - 1)p)}$, we can rewrite this as

$$X^2 + \left(-\pi_a\alpha_a + \pi_i\alpha_i + \frac{\pi_s\alpha_s(\pi_a\alpha_a + \delta(\pi_a\alpha_a + 2c_a + 4\delta p - 1) - 1)}{(\delta - 1)(\pi_a\alpha_a - 1)}\right) X + 2\pi_s\alpha_s(1 - \delta)p = 0. \quad (2.97)$$

Similarly, the first-order condition in δ can be written as

$$X^2 + \left(-\pi_a\alpha_a + \pi_i\alpha_i + \frac{\pi_s\alpha_s(\pi_a\alpha_a + 2c_a + 4\delta p - 1)}{\pi_a\alpha_a - 1}\right) X + 2\pi_s\alpha_s(1 - \delta)p = 0. \quad (2.98)$$

Using the first-order conditions together, we have that $X(1 - c_a - 2p\delta - \pi_a\alpha_a) = 0$. If $X = 0$, then using the definition of X, we have that $(1 - \delta)p = \frac{4c_a\pi_s\alpha_s - (-\pi_a\alpha_a + \pi_i\alpha_i + \pi_s\alpha_s)^2}{4\pi_s\alpha_s}$.

However, if $\pi_s\alpha_s > \pi_a\alpha_a - \pi_i\alpha_i + 2\sqrt{(c_a + 1)(\pi_a\alpha_a - \pi_i\alpha_i + c_a + 1)} + 2c_a + 2$, then $\delta p > 1 + p$. This can't happen in equilibrium since $\delta p < 1$ for consumers to be willing to pay for the automated patching option. Therefore, it cannot be the case that $X = 0$.

Then from $X(1 - c_a - 2p\delta - \pi_a\alpha_a) = 0$, we have that $\delta^*(p) = \frac{1 - c_a - \pi_a\alpha_a}{2p}$. Plugging this back into the profit function and maximizing over p again, we have two roots for p:

$$p = -\left(\frac{1}{18\pi_s\alpha_s}\right) \Big(\pi_s\alpha_s(\pi_a\alpha_a + 8\pi_i\alpha_i) + 2(\pi_a\alpha_a - \pi_i\alpha_i)^2 + 2(\pi_s\alpha_s)^2 - 3(c_a + 3)\pi_s\alpha_s \mp 2\sqrt{(\pi_a\alpha_a - \pi_i\alpha_i + \pi_s\alpha_s)^2\left((\pi_a\alpha_a - \pi_i\alpha_i)^2 + (\pi_s\alpha_s)^2 + \pi_s\alpha_s(-\pi_a\alpha_a + \pi_i\alpha_i - 3c_a)\right)}\Big) \quad (2.99)$$

However, when

$$\pi_s\alpha_s > \frac{1}{8(1-\pi_i\alpha_i)}\Bigg((\pi_a\alpha_a)^2 - 2\pi_a\alpha_a + 8(\pi_i\alpha_i)^2 + c_a^2 + 2c_a\pi_a\alpha_a - 8c_a\pi_i\alpha_i + 6c_a + 9 - 16\pi_i\alpha_i - (\pi_a\alpha_a - 4\pi_i\alpha_i + c_a + 3)\sqrt{(\pi_a\alpha_a - 1)(\pi_a\alpha_a + 8\pi_i\alpha_i - 9) + c_a^2 + 2c_a(\pi_a\alpha_a - 4\pi_i\alpha_i + 3)},\Bigg) \quad (2.100)$$

then the smaller root will be negative while the larger root is positive.

The equilibrium discount of this case is given as

$$\delta_{II}^* = (9\pi_s\alpha_s(\pi_a\alpha_a + c_a - 1))\Bigg(\pi_s\alpha_s(\pi_a\alpha_a + 8\pi_i\alpha_i) + 2(\pi_a\alpha_a - \pi_i\alpha_i)^2 + 2(\pi_s\alpha_s)^2 - 3\pi_s\alpha_s(c_a + 3) - 2\Big((\pi_a\alpha_a - \pi_i\alpha_i + \pi_s\alpha_s)^2\big((\pi_a\alpha_a - \pi_i\alpha_i)^2 + (\pi_s\alpha_s)^2 + \pi_s\alpha_s(-\pi_a\alpha_a + \pi_i\alpha_i - 3c_a)\big)\Big)^{\frac{1}{2}}\Bigg)^{-1}. \quad (2.101)$$

The equilibrium profit is given as $\Pi_{II}^* = \Pi_{II}(p_{II}^*, \delta_{II}^*)$. As we had done in Lemma 4, we can characterize the profit of this case using Taylor series. In particular, there exists an α_6 such for that $\alpha_s > \alpha_6$, the maximal profit of this case is

$$\Pi_{II}^* = \frac{(1 - c_a - \pi_a\alpha_a)^2}{4(1 - \pi_a\alpha_a)} + \frac{(c_a + \pi_a\alpha_a - \pi_i\alpha_i)^2}{4\pi_s\alpha_s} + \sum_{k=2}^{\infty} c_k\left(\frac{1}{\pi_s\alpha_s}\right)^k \quad (2.102)$$

for some coefficients c_k. As done before in Lemma 4, we will use a_k, b_k, c_k, and d_k to denote coefficients in the Taylor expansions without referring to specific expressions. These will be used across different cases, and they don't refer to the same quantities or expressions across cases.

On the other hand, suppose $0 < v_a < v_b < v_p < 1$ is induced. By part (III) of Lemma A.4, we obtain that $v_a = \frac{\delta p + c_a}{1 - \pi_a\alpha_a}$, v_b is the most positive root of the cubic $f_3(x) \triangleq \pi_s\alpha_s x^2(c_a - c_p + (\delta - 1)p + \pi_a\alpha_a x) + (c_a + (\delta - 1)p + \pi_a\alpha_a x)(c_a + (\delta - 1)p +$

$x(\pi_a\alpha_a - \pi_i\alpha_i))$, and $v_p = \frac{c_p v_b}{c_a-(1-\delta)p+\pi_a\alpha_a v_b}$. The profit function in this case is

$$\Pi_{III}(p,\delta) = p(1 - v_b(p,\delta)) + \delta p(v_b(p,\delta) - v_a(p,\delta)). \tag{2.103}$$

As we had done in Lemma 4, we employ asymptotic analysis to characterize the equilibrium prices and profit of this case. In particular, since v_b is the most positive root of the cubic equation $f_3(x) \triangleq \pi_s\alpha_s x^2(c_a - c_p + (\delta-1)p + \pi_a\alpha_a x) + (c_a + (\delta-1)p + \pi_a\alpha_a x)(c_a + (\delta-1)p + x(\pi_a\alpha_a - \pi_i\alpha_i))$ (and since $f'(x) \neq 0$ at the value of x that defines v_b), it follows that v_b is an analytic function of the parameters. Letting $v_b = A_0 + \sum_{k=1}^{\infty} d_k \left(\frac{1}{\pi_s\alpha_s}\right)^k$, the cubic equation defining v_b becomes $A_0^2(c_a - c_p - p(1-\delta) + A_0\pi_a\alpha_a)\pi_s\alpha_s + \sum_{k=0}^{\infty} e_k \left(\frac{1}{\pi_s\alpha_s}\right)^k = 0$ for some coefficients e_k. For this equation to hold, we must have $A_0 = 0$ or $A_0 = \frac{c_p-c_a+(1-\delta)p}{\pi_a\alpha_a}$. The double root $A_0 = 0$ corresponds to the two solutions of the cubic converging to zero, while $A_0 = \frac{c_p-c_a+(1-\delta)p}{\pi_a\alpha_a} > 0$ corresponds to the largest root of cubic.

Then substituting $v_b = \frac{c_p-c_a+(1-\delta)p}{\pi_a\alpha_a} + \frac{A_1}{\pi_s\alpha_s} + \sum_{k=2}^{\infty} d_k \left(\frac{1}{\pi_s\alpha_s}\right)^k$ into $f_3(x)$, we have $\frac{c_p(c_a+(\delta-1)p)(\pi_i\alpha_i-2A_1)+\mathrm{A1}(c_a+(\delta-1)p)^2+c_p^2(A_1+\pi_a\alpha_a-\pi_i\alpha_i)}{\pi_a\alpha_a} + \sum_{k=1}^{\infty} e_k \left(\frac{1}{\pi_s\alpha_s}\right)^k = 0$ for some coefficients e_k.

Solving for A_1 gives $A_1 = -\frac{c_p(\pi_i\alpha_i(c_a+(\delta-1)p)+c_p(\pi_a\alpha_a-\pi_i\alpha_i))}{(c_a-c_p+(\delta-1)p)^2}$. Successively iterating in this way, we can solve for the coefficients in the Taylor series for v_b, giving

$$\begin{aligned} v_b(p,\delta) = {} & \frac{-c_a + c_p - \delta p + p}{\pi_a\alpha_a} - \frac{(c_p(\pi_i\alpha_i(c_a + (\delta-1)p) + c_p(\pi_a\alpha_a - \pi_i\alpha_i)))}{\pi_s\alpha_s(c_a - c_p + (\delta-1)p)^2} + \\ & \big(c_p\pi_a\alpha_a(c_a + (\delta-1)p)(\pi_i\alpha_i(c_a + (\delta-1)p) + c_p(\pi_a\alpha_a - \pi_i\alpha_i))(\pi_i\alpha_i(c_a + (\delta-1)p) + \\ & c_p(2\pi_a\alpha_a - \pi_i\alpha_i))\big)\left((\pi_s\alpha_s)^2(c_a - c_p + (\delta-1)p)^5\right)^{-1} + \sum_{k=3}^{\infty} d_k \left(\frac{1}{\pi_s\alpha_s}\right)^k \end{aligned} \tag{2.104}$$

Substituting (2.104) into the profit function (2.103), differentiating with respect to p for the first-order condition, and then substituting in $p = \sum_{k=0}^{\infty} a_k \left(\frac{1}{\pi_s\alpha_s}\right)^k$

to iteratively solve for the coefficients a_k as done above for $v_b(p,\delta)$, we get

$$p^*(\delta) = -\frac{c_a(\pi_a\alpha_a+\delta-1)+(\pi_a\alpha_a-1)(\pi_a\alpha_a+c_p(\delta-1))}{2\left(2\delta(\pi_a\alpha_a-1)-\pi_a\alpha_a+\delta^2+1\right)}+ \Bigg(\pi_s\alpha_s\Bigg(-\pi_a\alpha_a(\delta-1)(\pi_a\alpha_a-1)+c_a\left(\delta(3\pi_a\alpha_a-2)-\pi_a\alpha_a+\delta^2+1\right)- c_p\left(\pi_a\alpha_a\delta^2+2\pi_a\alpha_a\delta-\pi_a\alpha_a+\delta^2-2\delta+1\right)\Bigg)^3\Bigg)^{-1}\Bigg(\big(2\pi_a\alpha_a c_p(\delta-1)(\pi_a\alpha_a-1) \left(2\delta(\pi_a\alpha_a-1)-\pi_a\alpha_a+\delta^2+1\right)\left(\pi_i\alpha_i c_a^2\left(\delta(3\pi_a\alpha_a-2)-\pi_a\alpha_a+\delta^2+1\right)+ c_a\Big(c_p\Big((\pi_a\alpha_a)^2(5\delta-3)+\pi_a\alpha_a\Big(\delta^2(3-\pi_i\alpha_i)-\delta(5\pi_i\alpha_i+6)+2\pi_i\alpha_i+3\Big)- 2\pi_i\alpha_i(\delta-1)^2\Big)-\pi_a\alpha_a\pi_i\alpha_i(\delta-1)(\pi_a\alpha_a-1)\Big)+c_p\Big(\pi_a\alpha_a(\delta-1)(\pi_a\alpha_a-1)(\pi_a\alpha_a+\pi_i\alpha_i)+ c_p\Big((\pi_a\alpha_a)^2\left(\delta^2-6\delta+3\right)+\pi_a\alpha_a\left(\delta^2(\pi_i\alpha_i-3)+2\delta(\pi_i\alpha_i+3)-\pi_i\alpha_i-3\right)+ \pi_i\alpha_i(\delta-1)^2\Big)\Big)\Big)\Big)\Bigg)+\sum_{k=2}^{\infty}a_k\left(\frac{1}{\pi_s\alpha_s}\right)^k. \quad (2.105)$$

Substituting (2.105) into the profit function (2.103), differentiating with respect to δ for the first-order condition, and then substituting in $\delta = \sum_{k=0}^{\infty} b_k\left(\frac{1}{\pi_s\alpha_s}\right)^k$ to iteratively solve for the coefficients b_k, we get

$$\delta^*_{III} = \frac{1-\pi_a\alpha_a-c_a}{1-c_p}-\Big(4c_p\pi_a\alpha_a(\pi_a\alpha_a+c_a-1)\big(\pi_i\alpha_i c_a^2+c_a(-\pi_a\alpha_a\pi_i\alpha_i+3\pi_a\alpha_a c_p- 2\pi_i\alpha_i c_p)+c_p(\pi_a\alpha_a(\pi_a\alpha_a+\pi_i\alpha_i)+c_p(\pi_i\alpha_i-3\pi_a\alpha_a))\big)\Big)(\pi_s\alpha_s(c_p-1)^2(\pi_a\alpha_a-c_a+c_p)^3)^{-1}+ \sum_{k=2}^{\infty}b_k\left(\frac{1}{\pi_s\alpha_s}\right)^k \quad (2.106)$$

for some coefficients b_k.

Substituting this into (2.105), we have that

$$p^*_{III} = \frac{1-c_p}{2}+\big(2\pi_a\alpha_a c_p\big(\pi_i\alpha_i c_a^2+c_a(-\pi_a\alpha_a\pi_i\alpha_i+3\pi_a\alpha_a c_p-2c_p\pi_i\alpha_i)+ c_p(\pi_a\alpha_a(\pi_a\alpha_a+\pi_i\alpha_i)+c_p(\pi_i\alpha_i-3\pi_a\alpha_a))\big)\big)\left(\pi_s\alpha_s(-\pi_a\alpha_a+c_a-c_p)^3\right)^{-1}+ \sum_{k=2}^{\infty}a_k\left(\frac{1}{\pi_s\alpha_s}\right)^k \quad (2.107)$$

The second-order conditions are satisfied, and the profit at this maximizer is given as

$$\Pi^*_{III} = \frac{1}{4}\left(\frac{c_a^2}{1-\pi_a\alpha_a} + \frac{(c_a-c_p)^2}{\pi_a\alpha_a} - 2c_p + 1\right) + \frac{c_p(\pi_a\alpha_a + c_a - c_p)(\pi_i\alpha_i(c_a - \pi_a\alpha_a) + c_p(2\pi_a\alpha_a - \pi_i\alpha_i))}{\pi_s\alpha_s(\pi_a\alpha_a - c_a + c_p)^2} + \sum_{k=2}^{\infty} c_k\left(\frac{1}{\pi_s\alpha_s}\right)^k \tag{2.108}$$

Next, suppose $0 < v_b < 1$ is induced. By part (IV) of Lemma A.4, we obtain that $v_b = \frac{1}{2} + \frac{-1+\pi_i\alpha_i+\sqrt{(1-\pi_s\alpha_s-\pi_i\alpha_i)^2+4p\pi_s\alpha_s}}{2\pi_s\alpha_s}$. The profit function in this case is

$$\Pi_{IV}(p,\delta) = p(1 - v_b(p,\delta)). \tag{2.109}$$

For brevity of exposition, we will just quickly give the optimal prices and profits after writing the profit function for the remaining cases. The derivation is the same as these previous cases.

The optimal price is given as

$$p^*_{IV} = \frac{1}{9\pi_s\alpha_s}\Bigg(-(\pi_i\alpha_i)^2 + 2\pi_i\alpha_i(1-2\pi_s\alpha_s) + \pi_s\alpha_s(4-\pi_s\alpha_s) - 1 + \sqrt{(-\pi_i\alpha_i + \pi_s\alpha_s + 1)^2\left(\pi_s\alpha_s(\pi_i\alpha_i - 1) + (\pi_i\alpha_i - 1)^2 + (\pi_s\alpha_s)^2\right)}\Bigg) \tag{2.110}$$

The discount δ can be any δ high enough to satisfy the conditions of part (IV) of Lemma A.4 are met under optimal pricing. The profit induced by the optimal price is given as

$$\Pi_{IV}^* = \frac{1}{54(\pi_s\alpha_s)^2}\Bigg(\Bigg((\pi_i\alpha_i)^2 - \sqrt{(-\pi_i\alpha_i+\pi_s\alpha_s+1)^2\big(\pi_s\alpha_s(\pi_i\alpha_i-1)+(\pi_i\alpha_i-1)^2+(\pi_s\alpha_s)^2\big)+ 2\pi_i\alpha_i(2\pi_s\alpha_s-1)+\pi_s\alpha_s(\pi_s\alpha_s-4)+1}\Bigg)\Bigg(3\pi_i\alpha_i-3\pi_s\alpha_s-3+\Bigg(5(\pi_i\alpha_i)^2+ 4\sqrt{(-\pi_i\alpha_i+\pi_s\alpha_s+1)^2\big(\pi_s\alpha_s(\pi_i\alpha_i-1)+(\pi_i\alpha_i-1)^2+(\pi_s\alpha_s)^2\big)}+ 2\pi_i\alpha_i(\pi_s\alpha_s-5)+\pi_s\alpha_s(5\pi_s\alpha_s-2)+5\Bigg)^{1/2}\Bigg)\Bigg) \quad (2.111)$$

To compare this with the asymptotic profit expressions for the other cases, it will be helpful to also represent the above profit as a Taylor series. There exists $\alpha_7 > 0$ such that for $\alpha_s > \alpha_7$, the profit above can be written as

$$\Pi_{IV}^* = \frac{1}{4\pi_s\alpha_s} - \frac{1}{8(\pi_s\alpha_s)^2} + \sum_{k=2}^{\infty} c_k \left(\frac{1}{\pi_s\alpha_s}\right)^k \quad (2.112)$$

Next, suppose $0 < v_b < v_a < 1$ is induced. By part (V) of Lemma A.4, we obtain that v_b is the most positive root of the cubic $f_4(x) \triangleq (1-\pi_a\alpha_a)\pi_s\alpha_s x^3 + ((1-\pi_a\alpha_a)(1-\pi_i\alpha_i) - c_a\pi_s\alpha_s - \delta p\pi_s\alpha_s)x^2 + (p(-1+\pi_a\alpha_a)+p(-1+\pi_i\alpha_i))x + p^2$ and $v_a = \frac{(c_a-(1-\delta)p)v_b}{v_b(1-\pi_a\alpha_a)-p}$. The profit function in this case is

$$\Pi_V(p,\delta) = \delta p(1-v_a(p,\delta)) + p(v_a(p,\delta)-v_b(p,\delta)). \quad (2.113)$$

The optimal price, if interior, is given as

$$p_V^* = \frac{(1-\pi_i\alpha_i)(-\pi_a\alpha_a+c_a+1)}{4(1-\pi_a\alpha_a)} - \frac{\big((\pi_i\alpha_i-1)^2\big((\pi_a\alpha_a-1)^2+c_a(3\pi_a\alpha_a-2\pi_i\alpha_i-1)\big)\big)}{16(c_a\pi_s\alpha_s(\pi_a\alpha_a-1))} + \sum_{k=2}^{\infty} a_k\left(\frac{1}{\pi_s\alpha_s}\right)^k. \quad (2.114)$$

The optimal discount, if interior, is given as

$$\delta_V^* = \frac{2(1-\pi_a\alpha_a)(1-\pi_a\alpha_a-c_a)}{(1-\pi_i\alpha_i)(1-\pi_a\alpha_a+c_a)} + \frac{(1-\pi_a\alpha_a)(\pi_a\alpha_a+c_a-1)\big((\pi_a\alpha_a-1)^2+c_a(3\pi_a\alpha_a-2\pi_i\alpha_i-1)\big)}{2c_a\pi_s\alpha_s(-\pi_a\alpha_a+c_a+1)^2} + \sum_{k=2}^{\infty} b_k\left(\frac{1}{\pi_s\alpha_s}\right)^k. \quad (2.115)$$

The profit induced by the interior maximizer in this case is given by

$$\Pi_V^* = \frac{(\pi_a\alpha_a+c_a-1)^2}{4(1-\pi_a\alpha_a)} + \frac{\big(c_a(1-\pi_i\alpha_i)^2\big)}{4(\pi_s\alpha_s(1-\pi_a\alpha_a))} + \sum_{k=2}^{\infty} c_k\left(\frac{1}{\pi_s\alpha_s}\right)^k. \quad (2.116)$$

Next, suppose $0<v_b<v_a<v_p<1$ is induced. By part (VI) of Lemma A.4, we obtain that v_b is the most positive root of $f_4(x)$, $v_a = \frac{(c_a-(1-\delta)p)v_b}{v_b(1-\pi_a\alpha_a)-p}$, and $v_p = \frac{(1-\delta)p+c_p-c_a}{\pi_a\alpha_a}$. The profit function in this case is

$$\Pi_{VI}(p,\delta) = \delta p(v_p(p,\delta)-v_a(p,\delta)) + p((1-v_p(p,\delta))+(v_a(p,\delta)-v_b(p,\delta))). \quad (2.117)$$

The optimal price, if interior, is given as

$$p_{VI}^* = \frac{1-c_p}{2} + \Big(c_a(c_a-c_p\pi_a\alpha_a+c_p)(c_a(1-\pi_i\alpha_i)+(1-\pi_a\alpha_a)(-\pi_i\alpha_i+2c_p-1))\Big)\Big(\pi_s\alpha_s(1+c_a-\pi_a\alpha_a)^3\Big)^{-1} + \sum_{k=2}^{\infty} a_k\left(\frac{1}{\pi_s\alpha_s}\right)^k. \quad (2.118)$$

The optimal discount, if interior, is given as

$$\delta_{VI}^* = \frac{1-c_a-\pi_a\alpha_a}{1-c_p} - \Big(2c_a\big(c_a^2+(\pi_a\alpha_a-1)\big(\pi_a\alpha_a+c_p^2-2c_p\pi_a\alpha_a\big)\big)(c_a(\pi_i\alpha_i-1)+(\pi_a\alpha_a-1)(-\pi_i\alpha_i+2c_p-1))\Big)\Big((c_p-1)^2\pi_s\alpha_s(-\pi_a\alpha_a+c_a+1)^3\Big)^{-1} + \sum_{k=2}^{\infty} b_k\left(\frac{1}{\pi_s\alpha_s}\right)^k. \quad (2.119)$$

The profit induced by the interior maximizer in this case is given by

$$\Pi^*_{VI} = \frac{1}{4}\left(\frac{c_a^2}{1-\pi_a\alpha_a} + \frac{(c_a-c_p)^2}{\pi_a\alpha_a} - 2c_p + 1\right) + \frac{c_a(1-c_p)(c_a(1-\pi_i\alpha_i)+(1-\pi_a\alpha_a)(c_p-\pi_i\alpha_i))}{\pi_s\alpha_s(-\pi_a\alpha_a+c_a+1)^2} + \sum_{k=2}^{\infty} c_k\left(\frac{1}{\pi_s\alpha_s}\right)^k. \quad (2.120)$$

Lastly, suppose $0<v_b<v_p<1$ is induced. By part (VII) of Lemma A.4, we obtain that v_b is the most positive root of $f_5(x) \triangleq \pi_s\alpha_s x^3 + (1-\pi_i\alpha_i-(c_p+p)\pi_s\alpha_s)x^2 - p(2-\pi_i\alpha_i)x + p^2$ and $v_p = \frac{c_p v_b}{v_b - p}$. The profit function in this case is

$$\Pi_{VII}(p,\delta) = p(1-v_b(p,\delta)). \quad (2.121)$$

The optimal price, if interior, is given as

$$p^*_{VII} = \frac{1-c_p}{2} - \frac{\left(2c_p^2(\pi_i\alpha_i(c_p+1)-3c_p+1)\right)}{(c_p+1)^3\pi_s\alpha_s} + \sum_{k=2}^{\infty} a_k\left(\frac{1}{\pi_s\alpha_s}\right)^k. \quad (2.122)$$

The discount δ can be any δ high enough to satisfy the conditions of part (VII) of Lemma A.4 are met under optimal pricing. The profit induced by the interior maximizer in this case is given by

$$\Pi^*_{VII} = \frac{1}{4}(1-c_p)^2 - \frac{c_p(1-c_p)(\pi_i\alpha_i(c_p+1)-2c_p)}{(c_p+1)^2\pi_s\alpha_s} + \sum_{k=2}^{\infty} c_k\left(\frac{1}{\pi_s\alpha_s}\right)^k. \quad (2.123)$$

To find the conditions under which the interior maximizer of each case indeed induces the correct market structure, we follow the same steps as in Lemma 4 by finding the conditions under which the interior maximizer satisfies the conditions of the cases in Lemma A.4. For brevity, we omit the algebra.

To have $0<v_b<v_a<v_p<1$ be induced by the maximizing prices given by (2.118) and (2.119), the conditions are $|c_p-\pi_a\alpha_a| < c_a < c_p(1-\pi_a\alpha_a)$ and $\pi_i\alpha_i < \frac{(c_a+c_p(1-\pi_a\alpha_a))}{1+c_a-\pi_a\alpha_a}$. Note that $\frac{(c_a+c_p(1-\pi_a\alpha_a))}{1+c_a-\pi_a\alpha_a} > \frac{c_p\pi_a\alpha_a}{1+c_p-c_a}$ from $0<c_p<1$, $0<c_a<1$, and $0<\pi_a\alpha_a<1$, so that the condition of the lemma $\pi_i\alpha_i < \frac{c_p\pi_a\alpha_a}{1+c_p-c_a}$ implies $\pi_i\alpha_i < \frac{(c_a+c_p(1-\pi_a\alpha_a))}{1+c_a-\pi_a\alpha_a}$.

To have $0<v_a<v_b<v_p<1$ be induced by the maximizing prices given by

(2.107) and (2.106), the conditions are $c_a < \min[c_p(1-\pi_a\alpha_a), \pi_a\alpha_a - c_p]$ and $\pi_i\alpha_i < \frac{2c_p\pi_a\alpha_a}{c_p-c_a+\pi_a\alpha_a}$. Note that $\frac{2c_p\pi_a\alpha_a}{c_p-c_a+\pi_a\alpha_a} > \frac{c_p\pi_a\alpha_a}{1+c_p-c_a}$ from $0 < \pi_a\alpha_a < 1$ and $0 < c_a < c_p(1-\pi_a\alpha_a)$, so that the condition of the lemma $\pi_i\alpha_i < \frac{c_p\pi_a\alpha_a}{1+c_p-c_a}$ implies $\pi_i\alpha_i < \frac{2c_p\pi_a\alpha_a}{c_p-c_a+\pi_a\alpha_a}$.

To have $0 < v_b < v_p < 1$ be induced by the maximizing price given by (2.122), the conditions are $\pi_i\alpha_i < \frac{2c_p}{1+c_p}$ and $\delta \geq \frac{-2c_a+(1+c_p)(1-\pi_a\alpha_a)}{1-c_p}$. Note that the condition $\pi_i\alpha_i < \frac{2c_p}{1+c_p}$ holds since one of the conditions of this lemma is $\pi_i\alpha_i < \frac{c_p}{1+c_p}$. Then given any parameters in the parameter space satisfying the lemma, this case $0 < v_b < v_p < 1$ can always be induced with any δ large enough to satisfy these conditions.

Now we compare the maximizing profits of each case to establish the lemma. By comparing (2.108) and (2.120) with (2.95), (2.102), (2.116), (2.112), and (2.123), it follows that there exists $\alpha_8 > 0$ such that if $\alpha_s > \alpha_8$, if either $0 < v_b < v_a < v_p < 1$ or $0 < v_a < v_b < v_p < 1$ can be induced by their maximizing prices, then they will because they dominate the profits of the other cases. Furthermore, since (2.108) can only be achieved when $c_a < \min[c_p(1-\pi_a\alpha_a), \pi_a\alpha_a - c_p]$ and (2.120) can only be achieved when $|c_p - \pi_a\alpha_a| < c_a < c_p(1-\pi_a\alpha_a)$ (which doesn't overlap with the region over which (2.108) can be achieved), it follows that p^* and δ^* are set so that

1. if $c_a < \min[\pi_a\alpha_a - c_p, c_p(1-\pi_a\alpha_a)]$, then $\sigma^*(v)$ is characterized by $0 < v_a < v_b < v_p < 1$ under optimal pricing,

2. if $|\pi_a\alpha_a - c_p| < c_a < c_p(1-\pi_a\alpha_a)$, then $\sigma^*(v)$ is characterized by $0 < v_b < v_a < v_p < 1$ under optimal pricing.

When $c_a = c_p(1-\pi_a\alpha_a)$, the maximal profit when inducing $0 < v_b < v_p < 1$ equals the maximal profits when inducing either $0 < v_a < v_b < v_p < 1$ or $0 < v_b < v_a < v_p < 1$. Furthermore, for $c_a \geq c_p(1-\pi_a\alpha_a)$, by comparing (2.123) with (2.95), (2.102), (2.116), and (2.112), it follows that there exists $\alpha_9 > 0$ such that if $\alpha_s > \alpha_9$, then the profit of $0 < v_b < v_p < 1$ will dominate the other cases. Also, $\delta = 1$ can be set to induce this case since $\delta = 1$ satisfies $\delta \geq \frac{-2c_a+(1+c_p)(1-\pi_a\alpha_a)}{1-c_p}$ when $c_a \geq c_p(1-\pi_a\alpha_a)$. Altogether, when $\alpha_s > \tilde{\alpha}_2 \triangleq \max[\alpha_8, \alpha_9]$ and if $\pi_i\alpha_i < \min\left[\frac{c_p\pi_a\alpha_a}{1+c_p-c_a}, \frac{c_p}{1+c_p}\right]$, then p^* and δ^* are set so that

1. if $c_a < \min\left[\pi_a\alpha_a - c_p, c_p(1-\pi_a\alpha_a)\right]$, then $\sigma^*(v)$ is characterized by $0 < v_a < v_b < v_p < 1$ under optimal pricing,

2. if $|\pi_a\alpha_a - c_p| < c_a < c_p(1-\pi_a\alpha_a)$, then $\sigma^*(v)$ is characterized by $0 < v_b < v_a < v_p < 1$ under optimal pricing, and if

3. if $c_a > c_p(1-\pi_a\alpha_a)$, then $\sigma^*(v)$ is characterized by $0 < v_b < v_p < 1$ under optimal pricing.

$\square$

Proofs of Propositions

Proof of Proposition 2: We focus on the region in which all segments are represented under optimal pricing in the base case. Specifically, for $\alpha_s > \tilde{\alpha}_1$, by Lemma 4, we have that p^* is set so that if $c_p - \pi_a\alpha_a < c_a < 1 - \pi_a\alpha_a - (1 - c_p)\sqrt{1 - \pi_a\alpha_a}$, then $\sigma^*(v)$ is characterized by $0 < v_b < v_a < v_p < 1$ under optimal pricing. By Lemma 5, for $\alpha_s > \tilde{\alpha}_2$, when patching rights are priced under the same parameter region, there are two cases: either $0 < v_a < v_b < v_p < 1$ is induced or $0 < v_b < v_a < v_p < 1$ is induced. Specifically, p^* and δ^* are set so that

(i) if $c_a < \min[\pi_a\alpha_a - c_p, c_p(1 - \pi_a\alpha_a)]$, then $\sigma^*(v)$ is characterized by $0 < v_a < v_b < v_p < 1$ under optimal pricing, and

(ii) if $|\pi_a\alpha_a - c_p| < c_a < c_p(1 - \pi_a\alpha_a)$, then $\sigma^*(v)$ is characterized by $0 < v_b < v_a < v_p < 1$ under optimal pricing.

In either case, since $c_p(1 - \pi_a\alpha_a) > 1 - \pi_a\alpha_a - (1 - c_p)\sqrt{1 - \pi_a\alpha_a}$ using the assumptions that $0 < c_p < 1$ and $0 < \pi_a\alpha_a < 1$, we have that $c_p - \pi_a\alpha_a < c_a < 1 - \pi_a\alpha_a - (1 - c_p)\sqrt{1 - \pi_a\alpha_a}$ is a subset of the union of the regions $c_a < \min[\pi_a\alpha_a - c_p, c_p(1 - \pi_a\alpha_a)]$ and $|\pi_a\alpha_a - c_p| < c_a < c_p(1 - \pi_a\alpha_a)$. Moreover, the intersection of $c_p - \pi_a\alpha_a < c_a < 1 - \pi_a\alpha_a - (1 - c_p)\sqrt{1 - \pi_a\alpha_a}$ with either $c_a < \min[\pi_a\alpha_a - c_p, c_p(1 - \pi_a\alpha_a)]$ or $|\pi_a\alpha_a - c_p| < c_a < c_p(1 - \pi_a\alpha_a)$ is non-empty.

In the first case, the induced profit under optimal pricing in the status quo case when patching rights aren't priced is given by (2.54), and induced profit when patching rights are priced is given by (2.108). The percentage increase in profit is given by

$$\frac{\Pi_P - \Pi_{SQ}}{\Pi_{SQ}} = \frac{(1-\pi_a\alpha_a)(c_a - c_p + \pi_a\alpha_a)^2}{\pi_a\alpha_a(1-c_a-\pi_a\alpha_a)^2} + \Bigg((\pi_a\alpha_a - c_a + c_p)^2 M -$$
$$4\pi_a\alpha_a(1-\pi_a\alpha_a)(1-\pi_a\alpha_a-c_a)(-\pi_a\alpha_a+c_a+c_p)\big(\pi_a\alpha_a((c_a+2)c_p+c_a)+$$
$$c_a(1-c_p)(c_p-c_a)-2c_p(\pi_a\alpha_a)^2\big)(c_a-c_p(1-\pi_a\alpha_a))+4\pi_i\alpha_i(\pi_a\alpha_a-1)(-\pi_a\alpha_a+c_a+1)$$
$$(-\pi_a\alpha_a+c_a-c_p)\big((\pi_a\alpha_a)^3(\pi_a\alpha_a)^2(c_a+c_p+1)-\pi_a\alpha_a(c_a-1)(c_a+c_p)+c_a(c_a-c_p)^2\big)$$
$$(c_a+c_p(\pi_a\alpha_a-1))\Bigg)\Bigg(\pi_a\alpha_a\pi_s\alpha_s(-\pi_a\alpha_a+c_a+1)^2(\pi_a\alpha_a+c_a-1)^3$$
$$(\pi_a\alpha_a-c_a+c_p)^2\Bigg)^{-1} + K_h\,, \quad (2.124)$$

where

$$M = 4c_a(\pi_a\alpha_a-1)(\pi_a\alpha_a+c_a-c_p)\Bigg((\pi_a\alpha_a+c_a)\big(\pi_a\alpha_a(2-\pi_a\alpha_a)+$$
$$c_a(\pi_a\alpha_a-2)+2c_p(\pi_a\alpha_a-1)^2\big)-\pi_a\alpha_a\Bigg). \quad (2.125)$$

Moreover, the reduction in the size of the unpatched population when patching rights are priced is given by

$$\hat{u}(\sigma^*|SQ) - \hat{u}(\sigma^*|PPR) =$$
$$\frac{\big(c_a^2(\pi_a\alpha_a-2)+c_a(\pi_a\alpha_a+c_p(2-3\pi_a\alpha_a))+\pi_a\alpha_a(\pi_a\alpha_a-1)(c_p-\pi_a\alpha_a)\big)}{\pi_s\alpha_s(-\pi_a\alpha_a+c_a+1)(\pi_a\alpha_a-c_a+c_p)} +$$
$$\sum_{k=2}^{\infty} d_k\left(\frac{1}{\pi_s\alpha_s}\right)^k, \quad (2.126)$$

which is strictly positive when $c_a < \min\left[\pi_a\alpha_a - c_p, c_p(1-\pi_a\alpha_a)\right]$.

Similarly, in the second case, the induced profit under optimal pricing in the status quo case when patching rights aren't priced is given by (2.54), and induced profit when patching rights are priced is given by (2.120). The percentage increase in profit is given by

$$\frac{\Pi_P - \Pi_{SQ}}{\Pi_{SQ}} = \frac{(1-\pi_a\alpha_a)(c_a - c_p + \pi_a\alpha_a)^2}{\pi_a\alpha_a(1-c_a-\pi_a\alpha_a)^2} + \Big(M + (-\pi_a\alpha_a + c_a + 1)\big(4\pi_i\alpha_i c_a(\pi_a\alpha_a - 1)(\pi_a\alpha_a + c_a - c_p)(c_a + c_p(\pi_a\alpha_a - 1))\big)\Big)\Big(\pi_a\alpha_a\pi_s\alpha_s(-\pi_a\alpha_a + c_a + 1)^2(\pi_a\alpha_a + c_a - 1)^3\Big)^{-1} K_k\,. \tag{2.127}$$

Moreover, the reduction in the size of the unpatched population when patching rights are priced is given by

$$\hat{u}(\sigma^*|SQ) - \hat{u}(\sigma^*|PPR) = \frac{(1-\pi_a\alpha_a)(\pi_a\alpha_a + c_a - c_p)}{\pi_s\alpha_s(-\pi_a\alpha_a + c_a + 1)} + \sum_{k=2}^{\infty} d_k \left(\frac{1}{\pi_s\alpha_s}\right)^k, \tag{2.128}$$

which is strictly positive when $|\pi_a\alpha_a - c_p| < c_a < c_p(1-\pi_a\alpha_a)$.

In either case, pricing patching rights increases profits and reduces the equilibrium size of the unpatched population as compared to the case when patching rights aren't priced. $\square$

Proof of Corollary 1: Differentiating (2.124) and (2.127) with respect to α_i, we have that in either case, the profit difference between pricing patching rights and the status quo case decreases in α_i.

In the first case when $0 < v_a < v_b < v_p < 1$ is induced under pricing patching rights, the derivative of the profit difference with respect to α_i is given as

$$\frac{d}{d\alpha_i}\left(\Pi_P^* - \Pi_{SQ}^*\right) = \frac{\pi_i\big(-c_a^3 + c_a^2(2c_p+1) + c_a\big(\pi_a\alpha_a(\pi_a\alpha_a - 1) - c_p^2 + c_p\pi_a\alpha_a\big) + c_p(\pi_a\alpha_a - 1)(c_p - \pi_a\alpha_a)\big)}{\pi_s\alpha_s(-\pi_a\alpha_a + c_a + 1)(-\pi_a\alpha_a + c_a - c_p)} + \sum_{k=2}^{\infty} c_k \left(\frac{1}{\pi_s\alpha_s}\right)^k, \tag{2.129}$$

which is negative when $c_a < \min\left[\pi_a\alpha_a - c_p, c_p(1-\pi_a\alpha_a)\right]$.

In the second case when $0 < v_b < v_a < v_p < 1$ is induced under pricing patching rights, the derivative of the profit difference with respect to α_i is given as

$$\frac{d}{d\alpha_i}\left(\Pi_P^* - \Pi_{SQ}^*\right) = -\frac{c_a\pi_i(\pi_a\alpha_a + c_a - c_p)}{\pi_s\alpha_s(-\pi_a\alpha_a + c_a + 1)} + \sum_{k=2}^{\infty} c_k \left(\frac{1}{\pi_s\alpha_s}\right)^k \tag{2.130}$$

which is negative when $|\pi_a\alpha_a - c_p| < c_a < c_p(1 - \pi_a\alpha_a)$. □

Proof of Corollary 2: Following the proof of Proposition 2, the reduction in the size of the unpatched population under the conditions of the corollary is given either by (2.126) or (2.128), depending on the parameters (with the conditions also given in the proof of Proposition 2).

In the first case, $\frac{d}{dc_p}\left[\hat{u}(\sigma^*|SQ) - \hat{u}(\sigma^*|PPR)\right] = \frac{2\pi_a\alpha_a(c_a - \pi_a\alpha_a)}{(c_a - c_p - \pi_a\alpha_a)^2\pi_s\alpha_s} + \sum_{k=2}^{\infty} d_k \left(\frac{1}{\pi_s\alpha_s}\right)^k$. Using $c_a < \pi_a\alpha_a - c_p$ (one of the conditions for this case, from Lemma 5), there exists $\hat{\alpha}_s$ such that $\alpha_s > \hat{\alpha}_s$ implies that $\frac{d}{dc_p}\left[\hat{u}(\sigma^*|SQ) - \hat{u}(\sigma^*|PPR)\right] < 0$.

Also,

$$\frac{d}{d(\pi_a\alpha_a)}\left[\hat{u}(\sigma^*|SQ) - \hat{u}(\sigma^*|PPR)\right] = \frac{1}{\pi_s\alpha_s}\Big(1+(2(c_p-c_a)(c_a^3-(1+2c_a(1+c_a))c_p)+ 4(c_a - c_p)(c_a^2 - (1 + c_a)c_p)\pi_a\alpha_a - 2(c_a^2 - c_ac_p + c_p^2)(\pi_a\alpha_a)^2)((1 + c_a - \pi_a\alpha_a)^2 (-c_a + c_p + \pi_a\alpha_a)^2)^{-1}\Big) + \sum_{k=2}^{\infty} d_k \left(\frac{1}{\pi_s\alpha_s}\right)^k. \tag{2.131}$$

There exists $\hat{\alpha}_s$ such that $\alpha_s > \hat{\alpha}_s$ implies that this is strictly positive, since $c_a < \pi_a\alpha_a - c_p$ in this parameter region and $\pi_a\alpha_a + c_a < 1$ from our initial model assumptions.

In the second case, $\frac{d}{dc_p}\left[\hat{u}(\sigma^*|SQ) - \hat{u}(\sigma^*|PPR)\right] = \frac{-(1-\pi_a\alpha_a)}{(1+c_a-\pi_a\alpha_a)\pi_s\alpha_s} + \sum_{k=2}^{\infty} d_k \left(\frac{1}{\pi_s\alpha_s}\right)^k$. There exists $\hat{\alpha}_s$ such that $\alpha_s > \hat{\alpha}_s$ implies that $\frac{d}{dc_p}\left[\hat{u}(\sigma^*|SQ) - \hat{u}(\sigma^*|PPR)\right]$ is strictly negative.

Also, $\frac{d}{d(\pi_a\alpha_a)}\left[\hat{u}(\sigma^*|SQ) - \hat{u}(\sigma^*|PPR)\right] = \frac{-c_a^2 + c_a(1+c_p-2\pi_a\alpha_a)+(1-\pi_a\alpha_a)^2}{(1+c_a-\pi_a\alpha_a)^2\pi_s\alpha_s} + \sum_{k=2}^{\infty} d_k \left(\frac{1}{\pi_s\alpha_s}\right)^k$.

There exists $\hat{\alpha}_s$ so that $\alpha_s > \hat{\alpha}_s$ implies $\frac{d}{d(\pi_a\alpha_a)}\left[\hat{u}(\sigma^*|SQ) - \hat{u}(\sigma^*|PPR)\right]$ is positive, using $c_a < c_p(1 - \pi_a\alpha_a)$ and $c_a > \pi_a\alpha_a - c_p$ (from Lemma 5). □

Proof of Proposition 3: Under the conditions of the Proposition 3, when patching rights are priced, there are two cases in equilibrium by Lemma 5: either $0 < v_a < v_b < v_p < 1$ is induced or $0 < v_b < v_a < v_p < 1$ is induced. In either case, we show that the optimal discount $\delta^* < 1$. It follows that it suffices to show that the price of the automated patching option can be higher when patching rights are priced than compared to the status quo case for low c_a.

In the first region when $c_a < \min\left[\pi_a\alpha_a - c_p, c_p(1 - \pi_a\alpha_a)\right]$, then $\sigma^*(v)$ is characterized by $0 < v_a < v_b < v_p < 1$ under optimal pricing. The discount when pricing patching rights is given in (2.106). Under the conditions $c_a < \min\left[\pi_a\alpha_a - c_p, c_p(1 - \pi_a\alpha_a)\right]$, there exists $\tilde{\alpha}_3 > 0$ such that $\alpha_s > \tilde{\alpha}_3$ implies the expression for δ^* given in (2.106) is bounded above by 1. To prove the lemma for this case, it suffices to show that the price of the automated patching option can be greater than the common status quo price across both options.

The price of the automated patching option when patching rights aren't priced is given in (2.53). Using (2.107) and (2.106), when patching rights are priced, the automated patching option has price

$$\delta^* p^* = \frac{1}{2}\left(1 - c_a - \pi_a\alpha_a\right) + \sum_{k=2}^{\infty} a_k \left(\frac{1}{\pi_s\alpha_s}\right)^k. \tag{2.132}$$

Comparing (2.132) with (2.53), the price of the automated patching option when patching rights are priced is greater than the common price in the base case when $c_a < \frac{\pi_i\alpha_i(2\pi_a\alpha_a - 1) - 1}{2\pi_a\alpha_a + \pi_i\alpha_i - 3} - \pi_a\alpha_a$. The intersection of this with the parameter region of this case is non-empty when $\pi_a\alpha_a < \frac{1+\pi_i\alpha_i}{2}$.

The argument for the second case (in which $0 < v_b < v_a < v_p < 1$ is induced in equilibrium) is similar and is omitted for brevity. We find that $\delta^* p^*$ is greater than the price of the status quo case if $c_a < \frac{(1-\pi_a\alpha_a)(\pi_i\alpha_i - 2c_p + 1)}{-4\pi_a\alpha_a - \pi_i\alpha_i + 5}$. The intersection of this with the parameter region of this case, namely with the condition $c_a > \pi_a\alpha_a - c_p$, is non-empty also when $\pi_a\alpha_a < \frac{1+\pi_i\alpha_i}{2}$. $\square$

Proof of Proposition 4: Using Lemma 4, the status quo pricing induces $0 < v_b < v_a < v_p < 1$ market structure. Using Lemma 5, the vendor induces either

$0 < v_a < v_b < v_p < 1$ or $0 < v_b < v_a < v_p < 1$ under priced patching rights. Using the definition of social welfare in (2.34), the social welfare under status quo pricing is given by

$$W_{SQ} = \int_{v_b(p^*)}^{1} v dv - \Bigg(\int_{v_a(p^*)}^{v_p(p^*)} c_a + \pi_a \alpha_a v dv + \int_{v_b(p^*)}^{v_a(p^*)} ((v_a(p^*) - v_b(p^*))\pi_s \alpha_s + \pi_i \alpha_i) v dv + c_p(1 - v_p(p^*)) \Bigg), \quad (2.133)$$

where p^* is the equilibrium price in the status quo case, given in (2.53). Using its asymptotic expansion, this can be written as

$$W_{SQ} = \frac{1}{8}\big(\pi_a \alpha_a - \frac{3c_a^2}{\pi_a \alpha_a - 1} + \frac{4(c_a - c_p)^2}{\pi_a \alpha_a} + 2c_a - 8c_p + 3\big) + \sum_{k=1}^{\infty} c_k \left(\frac{1}{\pi_s \alpha_s}\right)^k. \quad (2.134)$$

When pricing patching rights, one of the two cases will arise. In the first case, the social welfare is given as

$$W_P = \int_{v_a(p^*,\delta^*)}^{1} v dv - \Bigg(\int_{v_a(p^*,\delta^*)}^{v_b(p^*,\delta^*)} c_a + \pi_a \alpha_a v dv + \int_{v_b(p^*,\delta^*)}^{v_p(p^*,\delta^*)} ((v_p(p^*,\delta^*) - v_b(p^*,\delta^*))\pi_s \alpha_s + \pi_i \alpha_i) v dv + c_p(1 - v_p(p^*,\delta^*)) \Bigg), \quad (2.135)$$

where δ^* and p^* are given in (2.106) and (2.107) respectively. In the second case, the social welfare is given as

$$W_P = \int_{v_b(p^*,\delta^*)}^{1} v dv - \Bigg(\int_{v_a(p^*,\delta^*)}^{v_p(p^*,\delta^*)} c_a + \pi_a \alpha_a v dv + \int_{v_b(p^*,\delta^*)}^{v_a(p^*,\delta^*)} ((v_a(p^*,\delta^*) - v_b(p^*,\delta^*))\pi_s \alpha_s + \pi_i \alpha_i) v dv + c_p(1 - v_p(p^*,\delta^*)) \Bigg), \quad (2.136)$$

where δ^* and p^* are given in (2.119) and (2.118) respectively.

In both cases, the asymptotic expression for the equilibrium welfare is given as

$$W_P = \frac{3}{8}\left(\frac{c_a^2}{1-\pi_a\alpha_a} + \frac{(c_a-c_p)^2}{\pi_a\alpha_a} - 2c_p + 1\right) + \sum_{k=1}^{\infty} c_k \left(\frac{1}{\pi_s\alpha_s}\right)^k . \tag{2.137}$$

Comparing (2.134) and (2.137) reveals that pricing patching rights in this region of the parameter space hurts welfare.

We further characterize which losses drive this result. We define the total attack-related losses under status quo pricing with

$$SL_{SQ} \triangleq \int_{\mathcal{V}} \mathbb{1}_{\{\sigma^*(v)=(B,NP)|SQ\}} \left(\pi_s\alpha_s u(\sigma^*|SQ) + \pi_i\alpha_i\right) v dv \,, \tag{2.138}$$

the total costs associated with automated patching under status quo pricing with

$$AL_{SQ} \triangleq \int_{\mathcal{V}} \mathbb{1}_{\{\sigma^*(v)=(B,AP)|SQ\}} c_a + \pi_a\alpha_a v dv \,, \tag{2.139}$$

and the total costs associated with standard patching under status quo pricing with

$$PL_{SQ} \triangleq \int_{\mathcal{V}} \mathbb{1}_{\{\sigma^*(v)=(B,P)|SQ\}} c_p dv \,. \tag{2.140}$$

Specifically, the loss measures when $0 < v_a < v_b < v_p < 1$ is induced in equilibrium under status quo pricing are given as follows.

$$\begin{aligned} SL_{SQ} = -\Bigg(&((\pi_a\alpha_a(\pi_a\alpha_a - 1) + c_a(\pi_a\alpha_a - 2))((\pi_a\alpha_a - 1)(\pi_a\alpha_a - \pi_i\alpha_i) + c_a(\pi_a\alpha_a + \\ &\pi_i\alpha_i - 2)))\Bigg)\Bigg(2(\pi_s\alpha_s(\pi_a\alpha_a - 1)(-\pi_a\alpha_a + c_a + 1))\Bigg)^{-1} + \sum_{k=2}^{\infty} c_k \left(\frac{1}{\pi_s\alpha_s}\right)^k , \end{aligned} \tag{2.141}$$

$$AL_{SQ} = \Big((\pi_a\alpha_a(\pi_a\alpha_a - 1) + c_a(\pi_a\alpha_a - 2))\big((\pi_a\alpha_a - 1)^3(\pi_a\alpha_a - \pi_i\alpha_i)+$$

$$c_a^3(3\pi_a\alpha_a+\pi_i\alpha_i-4)+c_a^2(\pi_a\alpha_a-1)(3\pi_a\alpha_a-\pi_i\alpha_i-2)+c_a(\pi_a\alpha_a-1)^2(\pi_a\alpha_a+\pi_i\alpha_i-2)\big)\Big)$$

$$\Big(2\pi_s\alpha_s(\pi_a\alpha_a - 1)(-\pi_a\alpha_a + c_a + 1)^3\Big)^{-1} - \Big(\big((\pi_a\alpha_a)^2 + \pi_a\alpha_a(c_a - 2c_p - 1)-$$

$$2c_a + 2c_p\big)(c_a(\pi_a\alpha_a - 2) + (\pi_a\alpha_a - 1)(\pi_a\alpha_a + 2c_p))\Big)$$

$$\Big(8\left(\pi_a\alpha_a(\pi_a\alpha_a - 1)^2\right)\Big)^{-1} + \sum_{k=2}^{\infty} c_k \left(\frac{1}{\pi_s\alpha_s}\right)^k , \quad (2.142)$$

and

$$PL_{SQ} = \frac{c_p(\pi_a\alpha_a + c_a - c_p)}{\pi_a\alpha_a} dv + \sum_{k=2}^{\infty} c_k \left(\frac{1}{\pi_s\alpha_s}\right)^k . \quad (2.143)$$

Similarly, we define the total attack-related losses under priced patching rights with

$$SL_P \triangleq \int_{\mathcal{V}} \mathbb{1}_{\{\sigma^*(v)=(B,NP)|PPR\}} \left(\pi_s\alpha_s u(\sigma^*|PPR) + \pi_i\alpha_i\right) v dv\,, \quad (2.144)$$

the total costs associated with automated patching under priced patching rights with

$$AL_P \triangleq \int_{\mathcal{V}} \mathbb{1}_{\{\sigma^*(v)=(B,AP)|PPR\}} c_a + \pi_a\alpha_a v dv\,, \quad (2.145)$$

and the total costs associated with standard patching under priced patching rights with

$$PL_P \triangleq \int_{\mathcal{V}} \mathbb{1}_{\{\sigma^*(v)=(B,P)|PPR\}} c_p dv\,. \quad (2.146)$$

In the first case, in which $0 < v_a < v_b < v_p < 1$ is induced under PPR, the loss measures in equilibrium are given as follows.

$$SL_P = \frac{c_p(\pi_i\alpha_i(\pi_a\alpha_a - c_a + c_p) - 2\pi_a\alpha_a c_p)}{\pi_s\alpha_s(-\pi_a\alpha_a + c_a - c_p)} + \sum_{k=2}^{\infty} c_k \left(\frac{1}{\pi_s\alpha_s}\right)^k , \quad (2.147)$$

$$AL_P = \Bigg(c_p \big(2c_p \pi_a \alpha_a \left((\pi_a \alpha_a)^2 + c_a^2 - \pi_a \alpha_a (2c_a + c_p) - 3c_a c_p + 2c_p^2 \right) + \pi_i \alpha_i (-\pi_a \alpha_a + c_a - c_p) \left((\pi_a \alpha_a)^2 - 2c_a \pi_a \alpha_a + (c_a - c_p)^2 \right) \big) \Bigg) \Bigg(\pi_s \alpha_s (-\pi_a \alpha_a + c_a - c_p)^3 \Bigg) + \frac{c_p (\pi_a \alpha_a + c_a - c_p)}{2\pi_a \alpha_a} + \sum_{k=2}^{\infty} c_k \left(\frac{1}{\pi_s \alpha_s} \right)^k , \quad (2.148)$$

and

$$PL_P = \Bigg(c_p \pi_a \alpha_a (\pi_a \alpha_a + c_a + c_p)(\pi_i \alpha_i (\pi_a \alpha_a - c_a + c_p) + c_p(-3\pi_a \alpha_a - c_a + c_p)) \Bigg) \Bigg(\pi_s \alpha_s (\pi_a \alpha_a - c_a + c_p)^3 \Bigg)^{-1} + \frac{(c_a + c_p(\pi_a \alpha_a - 1))(c_a(2\pi_a \alpha_a - 3) + (\pi_a \alpha_a - 1)(2\pi_a \alpha_a + c_p))}{8\pi_a \alpha_a (\pi_a \alpha_a - 1)^2} + \sum_{k=2}^{\infty} c_k \left(\frac{1}{\pi_s \alpha_s} \right)^k . \quad (2.149)$$

Comparing these measures across status quo pricing and pricing patching rights, we have that $AL_P > AL_{SQ}$, $PL_{SQ} > PL_P$, and if $\frac{4c_p^2 \pi_a \alpha_a}{-c_a + c_p + \pi_a \alpha_a} - \frac{(c_a(2 - \pi_a \alpha_a) + \pi_a \alpha_a (1 - \pi_a \alpha_a))^2}{(1 + c_a - \pi_a \alpha_a)(1 - \pi_a \alpha_a)} > 0$, then $SL_P > SL_{SQ}$. Otherwise, $SL_P \leq SL_{SQ}$.

In the second case, in which $0 < v_b < v_a < v_p < 1$ is induced under PPR, the loss measures in equilibrium are given as follows.

$$SL_P = -\frac{((c_a - \pi_a \alpha_a c_p + c_p)(-\pi_i \alpha_i (-\pi_a \alpha_a + c_a + 1) + c_a - \pi_a \alpha_a c_p + c_p))}{2(\pi_s \alpha_s (\pi_a \alpha_a - 1)(-\pi_a \alpha_a + c_a + 1))} + \sum_{k=2}^{\infty} c_k \left(\frac{1}{\pi_s \alpha_s} \right)^k , \quad (2.150)$$

$$AL_P = \Bigg(\big(c_a^4(1-\pi_i\alpha_i) - c_a^3(\pi_a\alpha_a-1)(\pi_i\alpha_i(c_p-2)+c_p) - c_a^2(\pi_a\alpha_a-1)\big(-\pi_a\alpha_a + \pi_a\alpha_a c_p^2 + \pi_i\alpha_i(2\pi_a\alpha_a + c_p(4-3\pi_a\alpha_a)-3) + \pi_a\alpha_a c_p - 4c_p + 3\big) + c_a(\pi_a\alpha_a-1)^2(2\pi_a\alpha_a - 2\pi_i\alpha_i + c_p(-\pi_a\alpha_a(\pi_i\alpha_i+5) + 3\pi_i\alpha_i + 4c_p(\pi_a\alpha_a-1)+3)) + \pi_a\alpha_a(c_p-1)(\pi_a\alpha_a-1)^3 (c_p-\pi_i\alpha_i)\big)\Bigg)\Big(2\pi_s\alpha_s(\pi_a\alpha_a-1)(-\pi_a\alpha_a+c_a+1)^3\Big)^{-1} + \big((c_a+c_p(\pi_a\alpha_a-1)) (c_a(2\pi_a\alpha_a-3) + (\pi_a\alpha_a-1)(2\pi_a\alpha_a+c_p))\big)\big(8\pi_a\alpha_a(\pi_a\alpha_a-1)^2\big)^{-1} + \sum_{k=2}^{\infty} c_k \left(\frac{1}{\pi_s\alpha_s}\right)^k, \quad (2.151)$$

and

$$PL_P = \frac{c_a c_p(c_a(\pi_i\alpha_i-1) + (\pi_a\alpha_a-1)(-\pi_i\alpha_i+2c_p-1))}{\pi_s\alpha_s(-\pi_a\alpha_a+c_a+1)^2} + \frac{c_p(\pi_a\alpha_a+c_a-c_p)}{2\pi_a\alpha_a} + \sum_{k=2}^{\infty} c_k \left(\frac{1}{\pi_s\alpha_s}\right)^k. \quad (2.152)$$

Comparing these measures across status quo pricing and pricing patching rights, we have that $AL_P > AL_{SQ}$, $PL_{SQ} > PL_P$, and $SL_P < SL_{SQ}$ always holds under the conditions of this case.

□

Proof of Proposition 5: When $1-\pi_a\alpha_a-(1-c_p)\sqrt{1-\pi_a\alpha_a} \leq c_a < c_p(1-\pi_a\alpha_a)$, then by Lemmas 4 and 5, status quo pricing induces $0<v_b<v_a<v_p<1$ while pricing patching rights induces either $0<v_a<v_b<v_p<1$ or $0<v_b<v_a<v_p<1$.

So we have the same welfare expression under priced patching rights as in the proof of Proposition 4. Now however, the welfare under status quo pricing is given as

$$W_{SQ} = \int_{v_b(p^*)}^{1} v\,dv - \left(\int_{v_b(p^*)}^{v_p(p^*)} ((v_p(p^*)-v_b(p^*))\pi_s\alpha_s + \pi_i\alpha_i)v\,dv + c_p(1-v_p(p^*))\right), \quad (2.153)$$

where p^* is the equilibrium price in the status quo case, given in (2.55). Using its asymptotic expansion, this can be written as

$$W_{SQ} = \frac{3}{8}(1-c_p)^2 + \sum_{k=1}^{\infty} c_k \left(\frac{1}{\pi_s \alpha_s}\right)^k . \tag{2.154}$$

Comparing (2.154) to (2.137) reveals that pricing patching rights in this region of the parameter space improves welfare. Defining the specific losses for when $0 < v_b < v_p < 1$ arises under status quo pricing, we derive the following loss measures.

$$SL_{SQ} = -\frac{c_p(\pi_i \alpha_i (c_p+1) - 2c_p)}{(c_p+1)\pi_s \alpha_s} + \sum_{k=2}^{\infty} c_k \left(\frac{1}{\pi_s \alpha_s}\right)^k , \tag{2.155}$$

$$AL_{SQ} = 0 , \tag{2.156}$$

and

$$PL_{SQ} = \frac{c_p \left(\pi_i \alpha_i (c_p+1)\left(c_p^2+1\right) + 2c_p\left(-2c_p^2 + c_p - 1\right)\right)}{(c_p+1)^3 \pi_s \alpha_s} - \frac{1}{2}(c_p-1)c_p + \sum_{k=2}^{\infty} c_k \left(\frac{1}{\pi_s \alpha_s}\right)^k . \tag{2.157}$$

Comparing them with their respective measures given in Proof of Proposition 4 completes the proposition. Specifically, $SL_P < SL_{SQ}$, $PL_P < PL_{SQ}$, and $AL_P > AL_{SQ}$. $\square$

Chapter 2, in full, has been submitted for publication of the material as it may appear in “Market Segmentation and Software Security: Pricing Patching Rights,” in *Management Science.* Terrence August, Duy Dao, and Kihoon Kim. The dissertation author was the primary researcher and author of this material.

Chapter 3

Patch Management for Open-Source Security

3.1 Introduction

In the previous chapter, I studied the impact of optimally pricing patching rights for a commercial product on the security of the user population, profitability of the software, and the overall value of the product to the economy. For free open-source software, it is comparatively more difficult for companies to remain up-to-date with patches, especially since many open-source components are embedded within commercial applications that they use. For example, many companies didn't even know that they were vulnerable to the Heartbleed OpenSSL vulnerability. An organization that commits to patching on a regular basis would find it comparatively more costly to remain up-to-date on patches for open-source software than for commercial products. The presence of a large unpatched user population creates disincentives for software usage which, in turn, not only hurts profitability for a proprietary vendor but also reduces market coverage for free open-source software vendors. Moreover, without a price to deter low-valuation consumers who would otherwise be unpatched from entering the market, free open-source software can attract a larger population of risky users.

One would hope that just offering an automated patching option to consumers in this setting would alleviate the risks to users associated with using software utilized by a large unpatched user population. However, as we'll see, this may not necessarily be the case. As in the previous chapter, I construct a model of security where users can choose whether to purchase a software product and additionally whether to remain patched, unpatched, or have their systems automatically updated by a software vendor. In this chapter, I examine how our insights extend to the open-source software (OSS) domain. Many OSS products are made available for free, thus I propose how a priced patching rights policy can be implemented in this domain through taxes. I compare the relative value of priced patching rights in proprietary and open-source settings and characterize how the advancement of automated patching technologies affects the magnitude of taxing required.

3.2 Literature Review

Which is more secure: open source or proprietary software? This question has been the center of intense debate for years. OSS security vulnerabilities can be more easily spotted by developers, some of whom may offer fixes. However, these vulnerabilities can also be more easily spotted and attacked by malicious hackers as well. On the other hand, proprietary software might have secrecy working in its favor although security through obscurity is largely considered not a good strategy. The security community has argued both sides of the nuances of these observations. Schneier (1999) contends that in cryptography the algorithm is typically open to assure correctness. Thus, public algorithms which are designed to be secure even though they are open necessarily need to be more secure than proprietary ones. In his view, OSS security should be similar since it cannot simply rely on keeping the code secret. Other cryptography experts agree that opposing claims by proprietary vendors are refutable (Diffie 2003). Schneider (2000) believes that the incentives structure for developers and hackers has created a security landscape that is not governed by bugs discoverable from opening source code - the real security problems lie elsewhere. Many experts take the view that opening the source is necessary to build more secure systems, but certainly not sufficient (Hoepman and Jacobs 2007, Wheeler 2003). Others suggest that whether open source can improve security really comes down to the underlying economic incentives of the firms, users, and hackers (Witten et al. 2001, Anderson 2002).

We contribute to the discussion of this question by comparing security measures across OSS and proprietary contexts. Our focus however lies on how the difference in pricing (vendor-optimized proprietary price and free OSS) leads to starkly different equilibrium usage and unpatched behavior, with varying security implications in turn. Using our model, we are able to examine how the efficacy of priced patching rights policies compares across source code strategies. In the case of OSS, we study how a welfare-motivated project organizer (social planner) would price (tax) these rights. Thus, our work tackles this debate from a unique perspective that focuses on security as driven by user incentives.

Next, we discuss the literature that studies the management of security patches. Several papers examine the optimal *timing* of security patch release and application, along with its close connection to vulnerability disclosure. Beattie et al. (2002) characterize the optimal time to apply patches when trading off patch instability and security risk exposure. August and Tunca (2006) present a base model of software purchasing and patching in the presence of negative security externalities and patching costs, and then study the impact of patching mandates, rebates, and taxes. Choi et al. (2010) study the link between users' patching incentives and the issue of vulnerability disclosure. Lahiri (2012) and Kannan et al. (2013) study various aspects of the relationship between security patches and piracy. Our work is closer in spirit to the latter group of papers which employ models with a focus on users' patching incentives. We build on this body of work and focus on the inclusion of an automated patching option for users within a game theoretical context accenting negative externalities stemming from unpatched behavior. This inclusion serves two purposes. It permits a characterization of the natural consumer market segmentation that arises in equilibrium as users strategically respond to security risk and expanded patching options. In particular, the impact of automated patching on consumer behavior in equilibrium is different in a free open-source setting than in a proprietary setting, and formally examining those dynamics helps inform how security improvements should be made.

3.3 Model Description and Consumer Market Equilibrium

3.3.1 Model

There is a continuum of consumers whose valuations of a software product lie uniformly on $\mathcal{V} = [0, 1]$. Consumers are exposed to security risks associated with the software's use. In particular, a vulnerability can arise in the software, in which case the vendor makes a security patch available to all users of the software. Because the security vulnerability can be used by malicious hackers to exploit systems, users who do not apply the security patch are at risk.

The vendor offers two options for users to protect their respective systems: one with only automated patching as an option and one in which customers can choose when, if ever, to patch. In the current patching approach, both options are free. In this paper, we study the value of taxing patching rights by charging those who wish to retain patching rights a small fee. If a consumer elects to purchase the software and retain full patching rights, she pays the tax $\tau \geq 0$. Having this right means she can choose whether to patch the software or not patch the product and do so according to her own preferences. If she decides to patch the software, she will incur an expected cost of patching denoted $c_p > 0$. This *standard patching* cost accounts for the money and effort that a consumer must exert in order to verify, test, and roll-out patched versions of existing systems.

If she decides not to patch the software, then she faces the risk of an attack. The probability she is hit by a security attack is given by $\pi_s u$, where $\pi_s > 0$ is the probability an attack appears and u is the size of the unpatched population of users.[1] This reflects the negative security externality imposed by unpatched users of the software. If she is successfully attacked, she will incur expected security losses that are positively correlated with her valuation. That is, consumers with high valuations will suffer higher losses than consumers with lower valuations due to opportunity costs, higher criticality of data and loss of business. For simplicity, we assume that the correlation is of first order, i.e., the loss that a consumer with valuation v suffers if she is hit by an attack is $\alpha_s v$ where $\alpha_s > 0$ is a constant. We refer to $\pi_s \alpha_s$ as the *effective security risk* factor throughout the paper. This risk directly captures any attack that spreads through vulnerable populations and is agnostic to the specific attack vector or mechanism by which spreading occurs. This also indirectly captures any type of security attack where the incentives of the malicious individual for constructing the attack is positively related to the unpatched population size. For example, if large vulnerable populations are more attractive to hackers because it becomes easier to penetrate hosts or the return on their efforts becomes higher when infecting more hosts, then our model formulation of security risk will apply.

[1]The size of the unpatched population u is determined by the consumer strategies in equilibrium. Therefore, by the definition of $\mathcal{V}$, $u \in [0, 1]$.

If the consumer instead elects to purchase the software and relinquish patching rights, she pays no tax and obtains the software for free. In this case, the vendor retains full control over patching the software and will automatically and immediately do so to better protect the user population. From an implementation point of view, this software version would not give users much or any control over patch deployment (e.g., the typical options can be grayed out in this version). The user incurs a cost of *automated patching*, $c_a > 0$, which is associated with both inconvenience and configuration of the system to handle automatic deployment of security patches. Our model can examine any relationship between c_p and c_a. For example, it can capture the commonly observed situation in which users are choosing between: (i) completing all tasks associated with the rigorous, standard patching approach and incurring c_p, or (ii) doing the bare minimum tasks to deploy patches automatically without verification and incurring a lower cost, $c_a < c_p$, related to deployment. The model can also handle situations where $c_a \geq c_p$, to study scenarios in which users aim to achieve all tasks associated with standard patching but in an automated manner. In general, a software vendor who releases a security patch cannot test for compatibility of the patch with every possible user system configuration. Thus, there is always some risk associated with an automatically deployed patch causing a user's system to become unstable or even crash. We denote the probability that the automated patch is problematic with $\pi_a > 0$. We assume that the loss associated with an automated patch deployment failure is again positively correlated with her valuation, and that this correlation is of first order, denoted as $\alpha_a > 0$.

Each consumer makes a decision to buy, B, or not buy, NB. Similarly, the patching decision is denoted by one of patch, P, not patch, NP, and automatically patch, AP. In order to choose P or NP, the consumer must pay the tax τ to retain patching rights. By choosing AP, the consumer delegates patching rights to the vendor and obtains the software for free, and again we focus on vendors who deploy those patches on users' systems in an expeditious manner. The consumer action space is then given by $S = (\{B\} \times \{P, NP, AP\}) \cup (NB, NP)$. In a consumer market equilibrium, each consumer maximizes her expected utility given the equilibrium strategies for all consumers. For a given strategy profile $\sigma : \mathcal{V} \to S$, the expected utility for consumer v is given by:

$$U(v,\sigma) \triangleq \begin{cases} v-\tau-c_p & if \quad \sigma(v) = (B,P)\,; \\ v-\tau-\pi_s\alpha_s u(\sigma)v & if \quad \sigma(v) = (B,NP)\,; \\ v-c_a-\pi_a\alpha_a v & if \quad \sigma(v) = (B,AP)\,; \\ 0 & if \quad \sigma(v) = (NB,NP)\,, \end{cases} \tag{3.1}$$

where

$$u(\sigma) \triangleq \int_{\mathcal{V}} 1_{\{\sigma(v)\,=\,(B,NP)\}}\, dv\,. \tag{3.2}$$

To avoid trivialities and without loss of generality, we reduce the parameter space to $c_p, c_a \in (0,1)$, $\pi_s, \pi_a \in (0,1]$, $\alpha_s, \alpha_a \in (0,\infty)$, and $\pi_a\alpha_a \in (0, 1-c_a)$. The latter restriction, $\pi_a\alpha_a + c_a < 1$, ensures automated patching is economical.

3.3.2 Consumer Market Equilibrium

Before examining how patching rights should be taxed, we first must characterize how consumers segment across strategies for an arbitrary set of prices in equilibrium. The consumer with valuation v selects an action that solves the following maximization problem:

$$\max_{s\,\in\,S} \quad U(v,\sigma)\,, \tag{3.3}$$

where the strategy profile σ is composed of σ_{-v} (which is taken as fixed) and the choice being made, i.e., $\sigma(v) = s$. We denote her optimal action that solves (3.3) with $s^*(v)$. Further, we denote the equilibrium strategy profile with σ^*, and it satisfies the requirement that $\sigma^*(v) = s^*(v)$ for all $v \in \mathcal{V}$.

Lemma 1 *There exists a unique equilibrium consumer strategy profile σ^* that is* *characterized by thresholds* v_b, v_a, $v_p \in [0,1]$. *For each* $v \in \mathcal{V}$, *it satisfies either*

$$\sigma^*(v) = \begin{cases} (B,P) & if \quad v_p < v \le 1\,; \\ (B,NP) & if \quad v_b < v \le v_p\,; \\ (B,AP) & if \quad v_a < v \le v_b\,; \\ (NB,NP) & if \quad 0 \le v \le v_a\,, \end{cases} \tag{3.4}$$

or

$$\sigma^*(v) = \begin{cases} (B,P) & if \quad v_p < v \le 1\,; \\ (B,AP) & if \quad v_a < v \le v_p\,; \\ (B,NP) & if \quad v_b < v \le v_a\,; \\ (NB,NP) & if \quad 0 \le v \le v_b\,. \end{cases} \tag{3.5}$$

Lemma 1 establishes that if a population of patched consumers arises in equilibrium, it will consist of a segment of consumers with the highest valuations. Similar to the propertiary case, these consumers prefer to shield themselves from any valuation-dependent losses seen with either remaining unpatched and bearing security losses or selecting automated patching and bearing patch instability losses. Importantly, this segment need not arise, and $v_p = 1$ in cases where the valuation-dependent losses are smaller than the patching costs. For the middle segment on the other hand, the segment of consumers who elect for automated patching and the segment of consumers who elect to remain unpatched can be ordered either way. This ordering depends on the relative strength of the losses under each strategy. However as we will see, the optimal tax for patching rights will not be set to be so high that a restructuring of the market segments occurs when all segments are present in equilibrium under no tax. This is in contrast to how a vendor can set a significant premium on patching rights when c_a is low to induce a flip in the market structure. Before we examine the value of taxing patching rights, we first study how consumers respond in equilibrium to the current patching approach for open-source software in which there is no tax on patching rights.

3.4 Current OSS Patching Approach

With proprietary software, the combination of an automated patching option being available and the vendor's equilibrium pricing behavior tend to together help limit the size of the unpatched population that develops, even under status quo pricing. However, for open-source software (OSS) that is freely available, an automated patching option simply being available has a quite different impact on equilibrium outcomes. One might think that because a large unpatched population arises in the absence of a price, an automated patching option would help to reduce this unpatched population drastically. However, as we will see, the unpatched population that results even with the availability of automated patching can be substantial. Therefore, a policy that effectively prices patching rights in the open-source domain has significant potential.

In the following proposition, we examine how offering an automated patching option affects the size of the unpatched population in equilibrium. To do so, we compare it to an alternative scenario in which the consumer strategy set is restricted to $\tilde{S} = (\{B\} \times \{P, NP\}) \cup (NB, NP)$ thus excluding AP. We denote the equilibrium size of the unpatched population in this case with $\tilde{u}(\sigma^*)$. Our focus is on the most relevant parameter regime where the total security costs associated with automated patching ($\pi_a \alpha_a$, c_a) are reasonably close in magnitude to the standard patching cost (c_p).

Proposition 1 *Suppose $c_p - c_a < \pi_a \alpha_a < 1 - c_a/c_p$. Then, the inclusion of an automated patching option for consumers has the following impact on the size of the unpatched population in equilibrium, dependent upon the level of the effective security loss factor:*

(i) if $\pi_s \alpha_s \leq \frac{c_p(\pi_a\alpha_a)^2}{(c_p - c_a)^2}$, then $u(\sigma^*) = \tilde{u}(\sigma^*)$;

(ii) if $\frac{c_p(\pi_a\alpha_a)^2}{(c_p - c_a)^2} < \pi_s \alpha_s < \frac{1-\pi_a\alpha_a}{c_a}$, then $u(\sigma^*) < \tilde{u}(\sigma^*)$;

(iii) if $\pi_s \alpha_s \geq \frac{1-\pi_a\alpha_a}{c_a}$, then $u(\sigma^*) = \tilde{u}(\sigma^*)$.

Proposition 1 makes an important statement about the role of automated patching as it relates to the security of OSS: Even with some consumers choosing to

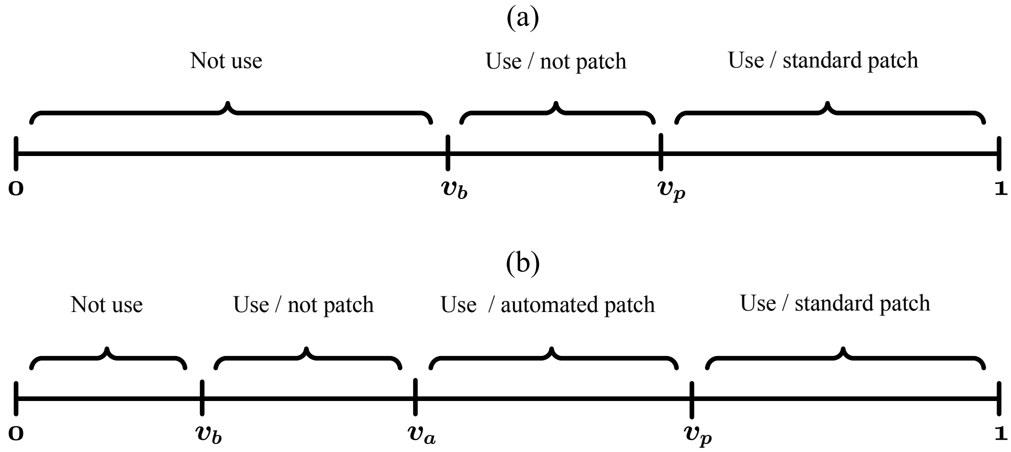


Figure 3.1: Equilibrium consumer market structure illustration for open-source software under a high effective security loss factor.
Panel (a) depicts the structure when automated patching is not available, and panel (b) depicts the structure when it is available.

use automated patching in equilibrium, the size of the unpatched population may simply remain unchanged. Proposition 1 establishes this behavior occurs under both a low and high effective security loss factor. Only within a medium range of security losses can an automated patching lead to a reduction in the size of the unpatched population. Perhaps most unsettling is part (iii) of Proposition 1. One would hope that the beneficial impact on security of an automated patching option would be highest when the effective security losses are also high. This is not the case, and, in fact, the unpatched population is exactly the same size with or without an automated patching option.

To understand why, it is useful to describe how users segment in equilibrium. When an automated patching option is not available, users with the highest valuations still prefer the standard patching option because they are unwilling to bear valuation-dependent losses, i.e., $\pi_s \alpha_s u(\sigma^*) v$ by remaining unpatched. When OSS has zero price, all consumers with positive valuations would prefer to use the software but cannot because of the security losses associated with a large unpatched population. Therefore, consumers with lower valuations begin to enter until the security externality becomes large enough that no other consumer elects to use. The equilibrium consumer market structure is illustrated in panel (a) of

Figure 3.1. Notably, under a high effective security loss factor, there can be a significant fraction of would-be users out of the market because of the unpatched users (those with valuations between v_b and v_p) causing risk.

When an automated patching option becomes available, users with the highest valuations still prefer the standard patching option because automated patching is also associated with valuation-dependent losses, i.e., $\pi_a \alpha_a v$. For users with moderate valuations, the trade-off shifts in favor of automated patching because they have lower value-at-risk which does not justify the higher costs associated with standard patching (recall $c_p > c_a$) to fully protect their valuations. For even lower valuation users, they will choose to remain unpatched, not even being willing to incur the cost of automated patching, c_a. This lowest segment of consumers faces the same trade-off as described in the case without automated patching - these users will continue to enter until there is a sufficiently large unpatched population that the next marginal user prefers not to use. Thus, the existence of an automated patching option only serves to shift the risky usage down the valuation space, which is depicted in panel (b) of Figure 3.1. What is important is that these consumers must be willing to use the software in the face of some risk, and it is precisely the lack of a price which creates a large potential user population. Therefore, the actual impact of an automated patching option for OSS is for market expansion. More consumers can become users and separate across protected forms, creating the opportunity for additional consumers (who would have been non-users without automated patching) to enter into the market. However, these additional users who enter and do not patch continue to cause an equivalent security externality on the network.

Part (ii) of Proposition 1 demonstrates that when the effective security loss factor is moderate, the size of the unpatched population can shrink when an automated patching option is available. Building on the preceding discussion, what changes in this case is that the consumer market gets covered and there is no more room to expand risky usage at the lower end of the valuation space. Thus, OSS products with moderate effective security losses that are necessarily in widespread use can benefit from an automated patching option by effecting an even further reduction in security risk. Finally, part (i) of Proposition 1 identifies conditions under which the security risk is sufficiently low such that automated patching is not a viable option, in which case there is no difference between the two scenarios.

From the preceding discussion, we see that OSS may have a sizable mass of unpatched users even when an automated patching option is offered. We next investigate the value of a policy analogous to priced patching rights in the open-source domain. In particular, we study a tax on patching rights set by a social planner to help mitigate the negative externality associated with unpatched usage. We begin by characterizing the equilibrium consumer market structures that emerge in both the open-source status quo and under the prospective tax policy, focusing on the case where security losses are appreciable and the use of automated patching solutions is incentive compatible.

Lemma 2 (OSS, Status Quo) *Suppose that $\pi_s \alpha_s > \omega$. If $c_p - \pi_a \alpha_a < c_a < c_p(1 - \pi_a \alpha_a)$, then σ^* is characterized by $0 < v_b < v_a < v_p < 1$ such that the lower tier of users remain unpatched and the middle tier prefers automated patching. If $c_a \leq c_p - \pi_a \alpha_a$, then σ^* is characterized by $0 < v_b < v_a < 1$ such that no consumer elects for standard patching in equilibrium.*

A comparison of equilibrium consumption under the status quo across proprietary and OSS cases is revealing. Examining Lemmas 4 and 2 where the market outcome is characterized by $0 < v_b < v_a < v_p < 1$, it becomes clear that the region of the parameter space in which we observe automated patching is larger in an OSS setting than in a proprietary one. In this sense, we are currently more likely to see consumers choosing automated patching options with OSS relative to proprietary software, across software classes; this behavior is good for security. On the flip side, there will also be a relatively larger mass of unpatched users in the OSS case, which is detrimental to security. Thus, it is useful to examine how the taxing of patching rights can mitigate this downside.

3.5 Taxing Patching Rights

Analogous to a vendor's pricing of patching rights for proprietary software, we study a government's taxing of patching rights for OSS. As a way to decrease the unpatched population size and as a result increase social welfare, the government may charge a tax ($\tau > 0$) on patching rights. For a given strategy profile $\sigma : \mathcal{V} \to S$, the expected utility for consumer v is then given by:

$$U(v,\sigma) = \begin{cases} v - \tau - c_p & if \quad \sigma(v) = (B,P)\,; \\ v - \tau - \pi_s \alpha_s u(\sigma) v & if \quad \sigma(v) = (B,NP)\,; \\ v - c_a - \pi_a \alpha_a v & if \quad \sigma(v) = (B,AP)\,; \\ 0 & if \quad \sigma(v) = (NB,NP)\,. \end{cases} \tag{3.6}$$

We denote the expected losses associated with security attacks stemming from the unpatched population $u(\sigma^*)$ with

$$SL \triangleq \int_{\mathcal{V}} \mathbb{1}_{\{\sigma^*(v)=(B,NP)\}} \pi_s \alpha_s u(\sigma^*) v dv\,. \tag{3.7}$$

In a similar fashion, we denote the expected losses associated with configuration and instability of automated patching with

$$AL \triangleq \int_{\mathcal{V}} \mathbb{1}_{\{\sigma^*(v)=(B,AP)\}} c_a + \pi_a \alpha_a v dv\,, \tag{3.8}$$

and the total costs associated with standard patching with

$$PL \triangleq \int_{\mathcal{V}} \mathbb{1}_{\{\sigma^*(v)=(B,P)\}} c_p dv\,. \tag{3.9}$$

The net impact of consumers changing their patching strategies (standard patching, remaining unpatched, electing for automated patching) on these security-related costs is unclear. In order to examine these concerns in aggregate, we also define total security-related costs as the sum of these three components:

$$L \triangleq SL + AL + PL\,, \tag{3.10}$$

in which case social welfare can be expressed as

$$W \triangleq \int_{\mathcal{V}} \mathbb{1}_{\{\sigma^*(v)\in\{(B,NP),(B,AP),(B,P)\}\}} v dv - L. \tag{3.11}$$

Next, we provide a characterization of both the equilibrium tax set by the social planner and the equilibrium consumer market structure that is induced by the welfare-maximizing tax.

Lemma 3 (OSS, Taxed Patching Rights) *Suppose that* $\pi_s\alpha_s > \omega$ *and patching rights are taxed. Then,*

(i) if $c_p - \pi_a\alpha_a < c_a < c_p(1-\pi_a\alpha_a)$*, then*

$$\tau^* = \frac{\pi_a\alpha_a}{\pi_s\alpha_s} + K_a\,, \tag{3.12}$$

and σ^* *is characterized by* $0 < v_b < v_a < v_p < 1$*.*

(ii) if $\frac{(\pi_a\alpha_a)^2}{1-\pi_a\alpha_a} < c_a \le c_p - \pi_a\alpha_a$*, then*

$$\tau^* = \frac{c_a}{2(1-\pi_a\alpha_a)} - \frac{1-3\pi_a\alpha_a}{16(1-\pi_a\alpha_a)\pi_s\alpha_s} + K_b\,, \tag{3.13}$$

and σ^* *is characterized by* $0 < v_b < v_a < 1$*.*

(iii) if $c_a \le c_p - \pi_a\alpha_a$ *and* $c_a \le \frac{(\pi_a\alpha_a)^2}{1-\pi_a\alpha_a}$*, then*

$$\tau^* = \frac{c_a + \pi_a\alpha_a}{2} + \frac{(c_a - 3\pi_a\alpha_a)(c_a+\pi_a\alpha_a)}{16\pi_s\alpha_s} + K_c\,, \tag{3.14}$$

and σ^* *is characterized by* $0 < v_a < v_b < 1$*.*[2]

First, we examine part (i) of Lemma 3 where c_a is within an intermediate range such that both standard patching and automated patching populations are present in equilibrium. In this region, an optimally configured tax reduces the size of the unpatched population in such a way that the consumer indifferent between

[2]We will represent constants that are of order $O\left(1/(\pi_s\alpha_s)^2\right)$ using the notation K and an enumerated subscript. Similarly, we will represent constants that are of order $O\left(1/(\pi_s\alpha_s)^3\right)$ using the notation J and an enumerated subscript.

automated patching and remaining unpatched under the status quo now switches to being unpatched under the tax. In other words, the threshold v_a under the optimal tax is higher than its counterpart in the status quo. Thus, the purchasing threshold v_b (i.e., the consumer indifferent between being an unpatched user and a non-user) also moves relatively even higher in comparison to the status quo due to the tax imposition. The net effect of both thresholds increasing in this manner yields a smaller unpatched population in equilibrium. Moreover, the tax induces some users who are engaged in standard patching practices under the status quo to forgo their patching rights. This characteristic can be seen in (3.12); the optimal tax τ^* increases as $\pi_a\alpha_a$ increases because a social planner needs to further incentivize users of standard patching to adopt automated patching solutions when these solutions are associated with increased instability. The planner essentially achieves this by increasing the cost of patching rights to these users. Although some consumers switch away from automated patching in the status quo to being unpatched as discussed above, the movement of consumers who elect for standard patching under the status quo toward automated patching yields a larger population of users of automated patching in aggregate under the optimal tax. In the following proposition, we formally state these findings.

Proposition 2 *For sufficiently high $\pi_s\alpha_s$ and $c_p - \pi_a\alpha_a < c_a < c_p(1-\pi_a\alpha_a)$, the optimal tax decreases the size of the unpatched population by $\frac{\pi_a\alpha_a(1-\pi_a\alpha_a)}{c_a(\pi_s\alpha_s)^2} + J_a$ such that $SL_P < SL_{SQ}$, increase the size of the automated patching population by $\frac{1}{\pi_s\alpha_s} + K_d$ such that $AL_P > AL_{SQ}$, and increase social welfare by $\frac{\pi_a\alpha_a}{2(\pi_s\alpha_s)^2} + J_b$.*

By Lemmas 2 and 3, the consumer market structure induced in equilibrium under the conditions of Proposition 2 is $0 < v_b < v_a < v_p < 1$ under both the status quo and optimally configured tax. Proposition 2 demonstrates that as the effective security loss factor increases, the tax has a relatively larger impact on increasing automated patching behavior in comparison to reducing unpatched behavior. Notably, as the effective security loss factor grows quite large, users have a significant incentive not to remain unpatched even under the status quo which limits the marginal benefit of a tax. However, when the effective security loss factor has a

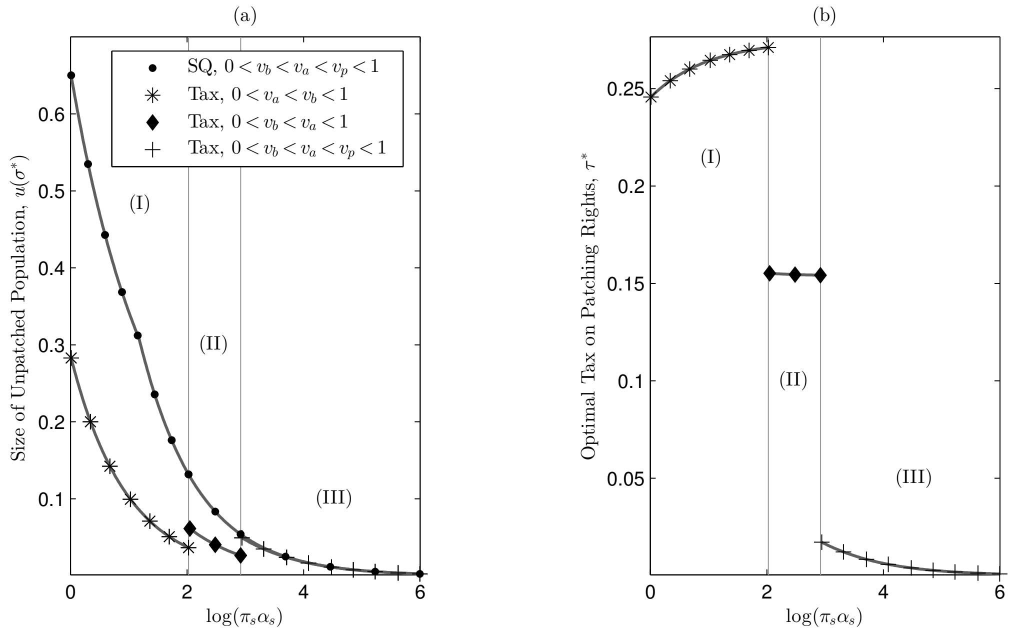


Figure 3.2: How the optimal tax on patching rights and equilibrium unpatched usage are influenced by the effective security loss factor when c_a is intermediate. Panel (a) illustrates the size of the unpatched population under both the status quo and in the presence of the tax. The optimal tax is depicted in panel (b). The parameter values are $\alpha_a = 3.5$, $\pi_a = 0.1$, $c_p = 0.5$, and $c_a = 0.2$.

moderate to moderately high magnitude, a tax can have an even greater impact on security risk.

Figure 3.2 illustrates both the equilibrium unpatched population size, $u(\sigma^*)$, and the optimal tax, τ^*, as a function of the effective security loss factor. First, we discuss the status quo. In panel (a), the upper curve represents the size of the unpatched population in equilibrium under the status quo. As $\pi_s \alpha_s$ initially increases, unpatched users with higher valuations switch to automated patching in order to bear relatively lower security risk. As $\pi_s \alpha_s$ increases further, unpatched users with lower valuations begin to drop out of the market, and the size of the unpatched population shrinks as is depicted. The impact of a tax on patching rights on the unpatched population is reflected by the lower curve in panel (a). Starting from the right-hand portion, Region (III) of panel (a) is consistent with Proposition 2 and illustrates how all consumer market segments are represented both under the status quo and under taxed patching rights. Panel (b) demonstrates how the optimal tax is a modest one and leads to a modest reduction in unpatched usage as seen in panel (a). As $\pi_s \alpha_s$ decreases into Region (II) and then (I), Figure 3.2 demonstrates how a much more significant tax is required to address mis-aligned incentives and induce a substantial reduction in unpatched usage.

Regions (I) and (II) illustrate what can happen when the effective security loss factor is not too high. Despite the consumer market structure being characterized by $0 < v_b < v_a < v_p < 1$ under the status quo, the optimal tax in these regions essentially precludes the existence of a segment of consumers who prefer standard patching in equilibrium. To see why, we begin by discussing Region (I) where the optimal tax induces the structure $0 < v_a < v_b < 1$. As can be seen in panel (b) of Figure 3.2, the optimal tax is set at a high level (in fact, higher than c_a) to provide incentives for consumers to forgo patching rights. In response, all consumers who were unpatched under the status quo either exit the market or choose the automated patching option. However, because of the large reduction in the unpatched population, the security risk is low and consumers with higher valuations who would be patching under the status quo now find it preferable to remain unpatched and bear the low, expected security losses in equilibrium. These consumers pay the tax to retain patching rights but need not exercise these rights. Instead, they are in spirit paying the tax to reduce security risk and hence the costly burden of standard patching processes. Notably, as panel (b) indicates, as the effective security loss factor increases through Region (I), a higher tax is needed to reduce unpatched usage by low valuation users and achieve these effects.

However, examining Region (II) in both panels of Figure 3.2, at some point the tax required is quite high and becomes too detrimental to total software usage in equilibrium; social welfare can be further improved by a different strategy here. In particular, in that with higher potential security risk, high valuation consumers prefer not to be exposed to higher valuation-dependent losses, more surplus would be created if a planner expands usage in the market to lower valuation consumers and provides incentives for high valuation users to switch to the automated patching option. In this case, a lower tax optimally expands usage and benefits welfare, while still limiting (albeit, to a lesser extent) the amount of unpatched behavior and associated expected security losses. In equilibrium, the consumer market structure is characterized by $0 < v_b < v_a < 1$; together, the security risk associated with expanded usage and the tax on patching rights are both sufficiently large that high valuation consumers prefer neither to pay the tax nor risk security losses. Instead, they elect for automated patching. In fact, the highest valuation consumers (who were incurring cost c_p under the status quo by standard patching) now incur cost $c_a + \pi_a \alpha_a v > c_p$ under automated patching but do not pay the tax τ.

As we discussed above, the value of taxing patching rights diminishes in security risk because it becomes more incentive compatible for users to choose patching and automated patching options. Thus, as the effective security loss factor increases even further, from a welfare perspective it is preferable for high valuation users to maintain patching rights and patch to prevent large security losses. In this case, a planner should set a small tax to encourage these users to retain rights and patch while also providing a modest disincentive for low valuation users to remain unpatched. The tax and its impact on an already smaller unpatched population is illustrated in Region (III) of Figure 3.2.

Having discussed parameter regions that correspond more closely to today's computing environment, we turn our attention to a region likely to unfold in the future when automated patching technology improves and becomes a less costly endeavor from the perspective of users. In Figure 3.3, we depict this case and illustrate how a larger tax on patching rights is utilized to significantly reduce unpatched usage in equilibrium. The following proposition formalizes our findings for a low cost of automated patching.

Proposition 3 *For sufficiently high $\pi_s \alpha_s$ and as automated patching solutions become more cost-effective to users, patching rights should be taxed relatively more heavily. Further,*

(i) if $\frac{(\pi_a \alpha_a)^2}{1-\pi_a \alpha_a} < c_a \leq c_p - \pi_a \alpha_a$, then optimally taxed patching rights decrease the size of the unpatched population by $\frac{1}{2\pi_s \alpha_s} + K_e$ such that $SL_P < SL_{SQ}$, decrease the size of the automated patching population by $\frac{1}{4(1-\pi_a \alpha_a)\pi_s \alpha_s} + K_f$ such that $AL_P < AL_{SQ}$, and increase social welfare by $\frac{c_a}{4(1-\pi_a \alpha_a)\pi_s \alpha_s} + K_g$;

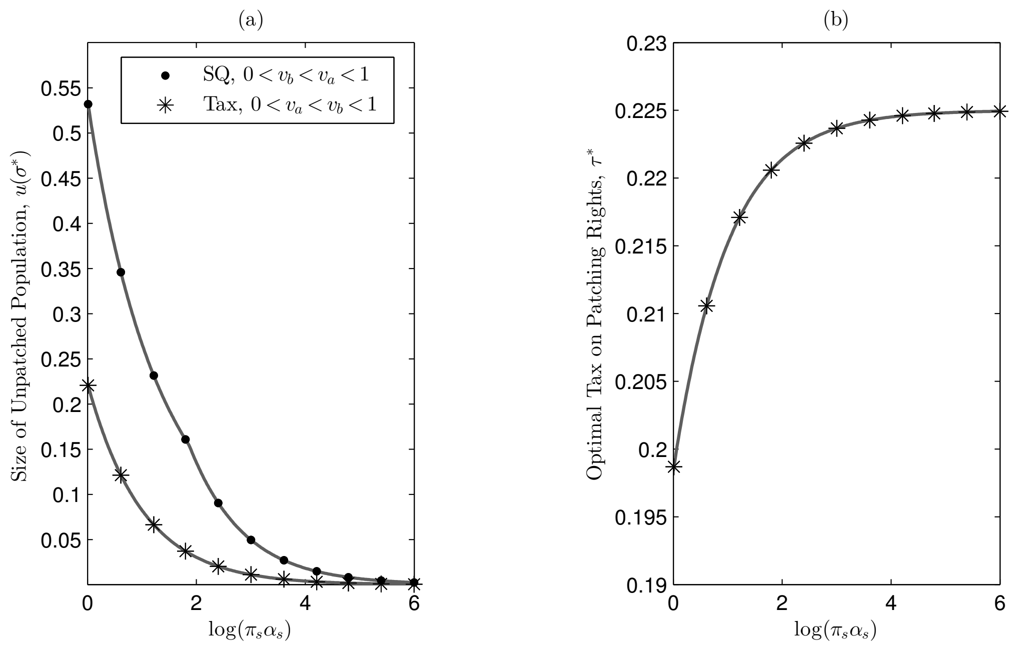


Figure 3.3: How the optimal tax on patching rights and equilibrium unpatched usage are influenced by the effective security loss factor when c_a is low. Panel (a) illustrates the size of the unpatched population under both the status quo and in the presence of the tax. The optimal tax is depicted in panel (b). The parameter values are $\alpha_a = 3.5, \pi_a = 0.1, c_p = 0.5,$ and $c_a = 0.1$.

(ii) if $c_a \leq \min\left(c_p - \pi_a\alpha_a, \frac{(\pi_a\alpha_a)^2}{1-\pi_a\alpha_a}\right)$*, then optimally taxed patching rights decrease the size of the unpatched population by* $\frac{2-c_a-\pi_a\alpha_a}{2\pi_s\alpha_s} + K_h$ *such that* $SL_P < SL_{SQ}$ *if and only if* $c_a > \frac{2\left(1-\sqrt{1-\pi_a\alpha_a(1-\pi_a\alpha_a)}\right)-\pi_a\alpha_a(1-\pi_a\alpha_a)}{1-\pi_a\alpha_a}$*, decrease the size of the automated patching population by* $\frac{c_a+\pi_a\alpha_a}{2\pi_s\alpha_s} + K_i$ *such that* $AL_P < AL_{SQ}$*, and increase social welfare by* $\frac{(c_a+\pi_a\alpha_a)^2}{4\pi_s\alpha_s} + K_j$ *.*

As Proposition 3 indicates, the socially optimal tax in this case is structured with a different concern in mind. In particular, Lemma 2 establishes that when c_a becomes lower, the equilibrium consumer market structure under the status quo is characterized by $0 < v_b < v_a < 1$ which is to say that consumers no longer find standard patching necessary due to the improvement in automated patching technology. In this case, parts (ii) and (iii) of Lemma 3 demonstrate that an optimally configured tax on patching rights will either induce the same consumer market structure or even $0 < v_a < v_b < 1$ if c_a is sufficiently reduced. In neither case can the planner induce standard patching behavior despite the effective security loss factor potentially being high. Therefore, in contrast to our findings in Proposition 2, the role of the optimal tax here is fundamentally different. Because high valuation users can have natural incentives to choose automated patching and thus need not be compelled to retain patching rights, the social planner has the ability to leverage high taxes on patching rights without affecting their behavior. For example, these consumers will choose not to retain patching rights under the conditions of part (i) of Proposition 3, in that $0 < v_b < v_a < 1$ is induced in equilibrium. As the proposition states, this large tax helps to significantly reduce unpatched usage by lower valuation consumers and increase social welfare. Because a large mass of lower valuation users are pushed out of the market by the tax, some users of the automated patching option under the status quo now switch to being unpatched, hence the size of the automated patching population also shrinks relative to the status quo.

More generally, we find that as automated patching solutions improve (across regions), a social planner should tend to utilize larger taxes to disincentivize the retainment of patching rights and significantly throttle unpatched usage. Interestingly, part (ii) of Proposition 3 establishes that if this technology improves sufficiently, large taxes will actually result in higher valuation consumers strategically retaining patching rights (despite the premium) and remaining unpatched. In this case, they benefit from the fact that only a small mass of consumers will find this behavior in their best interest which, in turn, limits the expected security losses they will incur. In this case, as can be seen in part (iii) of Lemma 3, the optimal tax is larger than c_a. Despite this outcome being an improvement to social welfare, because higher valuation users are the ones incentivized to remain unpatched, expected security losses can be increased relative to the status quo.

3.6 Conclusion

In the current state of affairs, both software end users and system administrators are faced with a barrage of security patches arriving weekly. However, because users are endowed with the right to choose whether or not to apply these security updates, a large portion of the user base ultimately chooses to remain unpatched, leaving their systems prone to security attacks. These users contribute to a security externality that affects all users of the software, which degrades its value and has a negative impact on its profitability. In this paper, we propose an adapted business model where a software vendor also differentiates its product based on patching rights. In this model, the right to choose whether or not to patch is no longer endowed. Instead, consumers who prefer to retain these rights and hence control of the patching status of their systems must pay a relative premium. Consumers who prefer to relinquish these rights have their software automatically updated by the vendor and, in exchange, end up paying a discounted price. The market segmentation induced carries a reduction in security risk and an increase in profitability to the vendor. In this way, a PPR policy can be a beneficial marketing strategy driving revenue growth; a vendor can market its product offerings as being more secure because its differentiated products incentivize better security behaviors by users.

Unlike with proprietary software, merely the inclusion of an automated patching option with open-source software often does not reduce the size of the unpatched population. In fact, under a high effective security loss factor, the size of the unpatched population is unchanged because consumers with lower valuations enter the market, re-establishing the externality. However, its inclusion does lead to a greater prevalence of automated patching as an equilibrium strategy across software classes (comparing OSS to proprietary software). Because of the large unpatched population that remains despite having an automated patching option, there is arguably an even greater need for patching rights to be managed in the OSS domain. We find that a tax on patching rights can be quite effective, and we characterize how it should be structured. In particular, when automated patching costs are comparable to standard patching costs and all segments are represented under the status quo, the optimal tax tends to be higher under a moderate effective security loss factor and more modest under a higher one. In this case, the tax is effective at reducing unpatched populations, expanding automated patching and achieving reduced security losses. When automated patching technology improves and costs become even smaller, a relatively large tax is warranted and leads to both a reduced unpatched population and a reduced automated patching population, in comparison to the status quo. Because of the significant drop in unpatched usage, in this case it is possible for expected security losses to increase when higher valuation consumers are the ones who become incentivized to remain unpatched in equilibrium. Nevertheless, the taxing of patching rights leads to a substantial improvement in social welfare.

3.7 Appendix

3.7.1 Consumer Market Equilibrium

Lemma A.1 *The consumer market equilibrium for an open-source software product made available for free is given by the following:*

(I) *If either (i) $\pi_a\alpha_a > c_p - c_a$ and $\pi_s\alpha_s \leq c_p$, or (ii) $\pi_a\alpha_a \leq c_p - c_a$ and $\pi_s\alpha_s \leq c_a + \pi_a\alpha_a$, then $0 = v_b < v_a = v_p = 1$;*

(II) *If either (i) $\pi_a\alpha_a \geq 1 - c_a/c_p$ and $c_p < \pi_s\alpha_s \leq 1/c_p$, or (ii) $c_p - c_a < \pi_a\alpha_a < 1 - c_a/c_p$ and $c_p < \pi_s\alpha_s \leq \frac{c_p(\pi_a\alpha_a)^2}{(c_p-c_a)^2}$, then $0 = v_b < v_a = v_p < 1$, where $v_p = \sqrt{\frac{c_p}{\pi_s\alpha_s}}$;*

(III) *If $\pi_a\alpha_a \geq 1 - c_a/c_p$ and $\pi_s\alpha_s > 1/c_p$, then $0 < v_b < v_a = v_p < 1$, where $v_b = c_p - 1/\pi_s\alpha_s$ and $v_p = c_p$;*

(IV) *If $c_p - c_a < \pi_a\alpha_a < 1 - c_a/c_p$ and $\pi_s\alpha_s > \frac{1-\pi_a\alpha_a}{c_a}$, then $0 < v_b < v_a < v_p < 1$, where $v_b = \frac{c_a}{1-\pi_a\alpha_a} - \frac{1}{\pi_s\alpha_s}$, $v_a = \frac{c_a}{1-\pi_a\alpha_a}$ and $v_p = \frac{c_p-c_a}{\pi_a\alpha_a}$;*

(V) *If $\pi_a\alpha_a \leq c_p - c_a$ and $\pi_s\alpha_s > \frac{1-\pi_a\alpha_a}{c_a}$, then $0 < v_b < v_a < v_p = 1$, where $v_b = \frac{c_a}{1-\pi_a\alpha_a} - \frac{1}{\pi_s\alpha_s}$ and $v_a = \frac{c_a}{1-\pi_a\alpha_a}$;*

(VI) *If $\pi_a\alpha_a \leq c_p - c_a$ and $c_a + \pi_a\alpha_a < \pi_s\alpha_s \leq \frac{1-\pi_a\alpha_a}{c_a}$, then $0 = v_b < v_a < v_p = 1$, where $v_a = \frac{\pi_a\alpha_a + \sqrt{(\pi_a\alpha_a)^2 + 4\pi_s\alpha_s c_a}}{2\pi_s\alpha_s}$;*

(VII) *If $c_p - c_a < \pi_a\alpha_a < 1 - c_a/c_p$ and $\frac{c_p(\pi_a\alpha_a)^2}{(c_p-c_a)^2} < \pi_s\alpha_s \leq \frac{1-\pi_a\alpha_a}{c_a}$, then $0 = v_b < v_a < v_p < 1$, where $v_a = \frac{\pi_a\alpha_a + \sqrt{(\pi_a\alpha_a)^2 + 4\pi_s\alpha_s c_a}}{2\pi_s\alpha_s}$ and $v_p = \frac{c_p-c_a}{\pi_a\alpha_a}$.*

Proof of Lemma A.1: We will prove parts (I) and (IV) which are representative of the arguments required. The remaining parts will be omitted due to similarity. For part (I), suppose $\pi_s\alpha_s \leq c_p$ is satisfied. This implies $(B, NP) \succ (B, P)$ for all $v \in \mathcal{V}$. Further, $\pi_a\alpha_a > c_p - c_a$ implies $\pi_a\alpha_a + c_a > \pi_s\alpha_s$ which ensures $(B, NP) \succ (B, AP)$ for all $v \in \mathcal{V}$. Because $\pi_s\alpha_s \leq c_p < 1$, by (2.58), $U(v, \sigma) \geq 0$ for all $v \in \mathcal{V}$ if $\sigma(v) = (B, NP)$. Second, $\pi_a\alpha_a \leq c_p - c_a$ and $\pi_s\alpha_s \leq c_a + \pi_a\alpha_a$ together imply the same preferences, and the characterization directly follows.

Next, we consider part (IV). By (2.58), $\sigma(v) = (B, P)$ if and only if $v \geq \max\left(\frac{c_p - c_a}{\pi_a\alpha_a}, \frac{c_p}{\pi_s\alpha_s u}, c_p\right)$, hence there exists a threshold $v_p \in (0,1]$ such that for all $v \in \mathcal{V}$, $\sigma^*(v) = (B, P)$ if and only if $v \geq v_p$. Moreover, $\sigma(v) \in \{\{B\} \times \{P, NP, AP\}\}$ if and only if $v \geq \max\left(c_p, \frac{c_a}{1-\pi_a\alpha_a}\right)$ and $v(1 - \pi_s\alpha_s u) \geq 0$. Thus, provided that $u \leq 1/\pi_s\alpha_s$, there exists a $\underline{v} \in (0,1]$ such that a consumer with valuation $v \in \mathcal{V}$ will purchase if and only if $v \geq \underline{v}$. Lastly, $(B, AP) \succ (B, NP)$ if and only if both $u \geq \pi_a\alpha_a/\pi_s\alpha_s$ and $v \geq \frac{c_a}{\pi_s\alpha_s u - \pi_a\alpha_a}$ are satisfied. Then if $\pi_a\alpha_a/\pi_s\alpha_s \leq u \leq 1/\pi_s\alpha_s$, then $\underline{v} = v_b \leq v_a \leq v_p$. Together $u \leq 1/\pi_s\alpha_s$ and $\pi_a\alpha_a < 1 - c_a/c_p$ imply that $v_p = \frac{c_p - c_a}{\pi_a\alpha_a}$. Suppose $u < 1/\pi_s\alpha_s$. Then, $v_b = 0$, hence $v_a = \frac{\pi_a\alpha_a + \sqrt{(\pi_a\alpha_a)^2 + 4\pi_s\alpha_s c_a}}{2\pi_s\alpha_s}$. But, $u = v_a \geq 1/\pi_s\alpha_s$ because $\pi_s\alpha_s > \frac{1-\pi_a\alpha_a}{c_a}$ which is a contradiction. Therefore, $u = 1/\pi_s\alpha_s$, from which it follows that $v_a = \frac{c_a}{1-\pi_a\alpha_a}$ and $v_b = \frac{c_a}{1-\pi_a\alpha_a} - \frac{1}{\pi_s\alpha_s}$. $\square$

Lemma A.2 *The consumer market equilibrium for an open-source software product under a tax τ is given by the following: If $\tau \leq c_a$, then*

(I) If $c_p + \tau - c_a < \pi_a\alpha_a < 1 - \frac{c_a}{c_p+\tau}$ and $(c_p + \tau - c_a)^2(c_p + \tau - c_a - c_p\pi_a\alpha_a - \tau\pi_a\alpha_a)\pi_s\alpha_s > c_p(\pi_a\alpha_a)^2(c_p + \tau - c_a - c_p\pi_a\alpha_a)$, then $0 < v_b < v_a < v_p < 1$, where $v_p = \frac{c_p+\tau-c_a}{\pi_a\alpha_a}$, $v_a = \frac{c_a-\tau}{\pi_s\alpha_s u - \pi_a\alpha_a}$, and $v_b = \frac{\tau}{1-\pi_s\alpha_s u}$. If $\tau < c_a$, $u \in (\frac{\pi_a\alpha_a}{\pi_s\alpha_s}, \frac{1}{\pi_s\alpha_s})$ satisfies $u(\pi_s\alpha_s u - \pi_a\alpha_a)(\pi_s\alpha_s u - 1) = c_a\pi_s\alpha_s u - (c_a - \tau + \tau\pi_a\alpha_a)$. If $\tau = c_a$, $u = \frac{\pi_a\alpha_a}{\pi_s\alpha_s}$;

(II) If $\pi_a\alpha_a \leq c_p + \tau - c_a$ and $(1 - c_a - \pi_a\alpha_a)\pi_s\alpha_s > (\pi_a\alpha_a + c_a - \tau)(1 - c_a - \pi_a\alpha_a + \tau)$, then $0 < v_b < v_a < 1$, where $v_a = \frac{c_a-\tau}{\pi_s\alpha_s u - \pi_a\alpha_a}$ and $v_b = \frac{\tau}{1-\pi_s\alpha_s u}$. If $\tau < c_a$, $u \in (\frac{\pi_a\alpha_a}{\pi_s\alpha_s}, \frac{1}{\pi_s\alpha_s})$ satisfies $u(\pi_s\alpha_s u - \pi_a\alpha_a)(\pi_s\alpha_s u - 1) = c_a\pi_s\alpha_s u - (c_a - \tau + \tau\pi_a\alpha_a)$. If $\tau = c_a$, $u = \frac{\pi_a\alpha_a}{\pi_s\alpha_s}$;

(III) If either (i) $\pi_a\alpha_a \geq 1 - \frac{c_a}{c_p+\tau}$ and $(1 - c_p - \tau)\pi_s\alpha_s > c_p(1 - c_p)$, or (ii) $c_p + \tau - c_a < \pi_a\alpha_a < 1 - \frac{c_a}{c_p+\tau}$, $c_p(1 - c_p) < (1 - c_p - \tau)\pi_s\alpha_s$, and $(c_p + \tau - c_a)^2(c_p + \tau - c_a - c_p\pi_a\alpha_a - \tau\pi_a\alpha_a)\pi_s\alpha_s \leq c_p(\pi_a\alpha_a)^2(c_p + \tau - c_a - c_p\pi_a\alpha_a)$, then $0 < v_b < v_p < 1$, where $v_p = \frac{c_p}{\pi_s\alpha_s u}$, $v_b = \frac{\tau}{1-\pi_s\alpha_s u}$, and $u \in (0, \frac{1}{\pi_s\alpha_s})$ satisfies $\pi_s\alpha_s u^2(\pi_s\alpha_s u - 1) = (c_p + \tau)\pi_s\alpha_s u - c_p$;

(IV) If either (i) $\pi_a\alpha_a \leq c_p + \tau - c_a$ and $(1 - \pi_a\alpha_a - c_a)\pi_s\alpha_s \leq (1 + \tau - \pi_a\alpha_a - c_a)(\pi_a\alpha_a + c_a - \tau)$, or (ii) $\pi_a\alpha_a > c_p + \tau - c_a$ and $(1 - c_p - \tau)\pi_s\alpha_s \leq c_p(1 - c_p)$,

then $0 < v_b < 1$*, where* $v_b = \frac{\tau}{1-\pi_s\alpha_s u}$ *and* $u = \frac{1+\pi_s\alpha_s-\sqrt{(1+\pi_s\alpha_s)^2-4\pi_s\alpha_s(1-\tau)}}{2\pi_s\alpha_s}$.

If $\tau > c_a$*, then*

(I) If either (i) $c_p + \tau - c_a < \pi_a\alpha_a \le (c_a c_p + \tau - c_a)/\tau$ *and* $(\pi_a\alpha_a - c_p - \tau + c_a)\pi_s\alpha_s > c_p(\pi_a\alpha_a - c_p)$*, or (ii)* $1 - c_a(1-c_p)/\tau < \pi_a\alpha_a < 1 - c_a/(c_p+\tau)$ *and* $c_a^2((c_p+\tau)(1-\pi_a\alpha_a) - c_a)\pi_s\alpha_s > (c_a - \tau(1-\pi_a\alpha_a))^2(1-\pi_a\alpha_a)$*, then* $0 < v_a < v_b < v_p < 1$*, where* $v_a = \frac{c_a}{1-\pi_a\alpha_a}$*,* $v_b = \frac{\tau - c_a}{\pi_a\alpha_a - \pi_s\alpha_s u}$*, and* $v_p = \frac{c_p}{\pi_s\alpha_s u}$*.* $u \in (0, \frac{\pi_a\alpha_a}{\pi_s\alpha_s})$ *satisfies* $\pi_s\alpha_s u^2(\pi_s\alpha_s u - \pi_a\alpha_a) = \pi_s\alpha_s(c_p + \tau - c_a)u - c_p\pi_a\alpha_a$*;*

(II) If either (i) $\tau - c_a < \pi_a\alpha_a \le c_p + \tau - c_a$ *and* $(1-\pi_a\alpha_a - c_a)c_a\pi_s\alpha_s > (c_a - \tau + \tau\pi_a\alpha_a)(1-\pi_a\alpha_a)$*, or (ii)* $c_p + \tau - c_a < \pi_a\alpha_a < 1 - \frac{c_a(1-c_p)}{\tau}$*,* $(c_a - \tau + \tau\pi_a\alpha_a)(1-\pi_a\alpha_a) < (1-\pi_a\alpha_a - c_a)c_a\pi_s\alpha_s$*, and* $(\pi_a\alpha_a - c_p - \tau + c_a)\pi_s\alpha_s \le c_p(\pi_a\alpha_a - c_p)$*, then* $0 < v_a < v_b < 1$*, where* $v_a = c_a/(1-\pi_a\alpha_a)$*,* $v_b = \frac{\tau - c_a}{\pi_a\alpha_a - \pi_s\alpha_s u}$*, and* $u = \frac{\pi_a\alpha_a + \pi_s\alpha_s - \sqrt{(\pi_a\alpha_a+\pi_s\alpha_s)^2 - 4\pi_s\alpha_s(\pi_a\alpha_a - \tau + c_a)}}{2\pi_s\alpha_s}$*;*

(III) If either (i) $1 - \frac{c_a(1-c_p)}{\tau} < \pi_a\alpha_a < 1 - \frac{c_a}{c_p+\tau}$*,* $c_p(1-c_p) < (1 - c_p - \tau)\pi_s\alpha_s$*, and* $c_a^2((c_p+\tau)(1-\pi_a\alpha_a) - c_a)\pi_s\alpha_s \le (\tau\pi_a\alpha_a - \tau + c_a)^2(1-\pi_a\alpha_a)$*, or (ii)* $\pi_a\alpha_a \ge 1 - \frac{c_a}{c_p+\tau}$ *and* $\pi_s\alpha_s(1 - c_p - \tau) > c_p(1-c_p)$*, then* $0 < v_b < v_p < 1$*, where* $v_b = \frac{\tau}{1-\pi_s\alpha_s u}$ *and* $v_p = \frac{c_p}{\pi_s\alpha_s u}$*.* $u \in (0, \frac{1}{\pi_s\alpha_s})$ *solves* $\pi_s\alpha_s u^2(\pi_s\alpha_s u - 1) = (c_p + \tau)\pi_s\alpha_s u - c_p$*;*

(IV) If either (i) $1 - c_a/\tau < \pi_a\alpha_a \le 1 - \frac{c_a(1-c_p)}{\tau}$ *and* $\pi_s\alpha_s c_a(1-\pi_a\alpha_a - c_a) \le (1-\pi_a\alpha_a)(c_a - \tau + \tau\pi_a\alpha_a)$*, or (ii)* $\pi_a\alpha_a > 1 - \frac{c_a(1-c_p)}{\tau}$ *and* $\pi_s\alpha_s(1-\tau-c_p) \le c_p(1-c_p)$*, then* $0 < v_b < 1$*, where* $v_b = 1 - u$ *and* $u = \frac{1+\pi_s\alpha_s-\sqrt{(1+\pi_s\alpha_s)^2-4\pi_s\alpha_s(1-\tau)}}{2\pi_s\alpha_s}$*;*

(V) If $\pi_a\alpha_a \le \tau - c_a$*, then* $0 < v_a < 1$*, where* $v_a = \frac{c_a}{1-\pi_a\alpha_a}$ *and* $u = 0$*.*

Note that the consumer market equilibrium structures of (III) and (IV) when $\tau \le c_a$ *are the same as those when* $\tau > c_a$*.*

Proof of Lemma A.2: The results can be derived as a special case of Lemma A.4 in Chapter 2 by setting $p = \tau$ and $\delta = 0$. □

Proof of Lemma 2: This follows from taking $\pi_s\alpha_s \to \infty$ in Lemma A.1. □

Proof of Lemma 3: Technically, we prove that there exists an α such that for $\alpha_s > \alpha$, τ^* is set so that

(i) if $c_p - \pi_a\alpha_a < c_a < c_p(1-\pi_a\alpha_a)$, then σ^* is characterized by $0 < v_b < v_a < v_p < 1$ under the optimal tax.

(ii) if $\frac{(\pi_a\alpha_a)^2}{1-\pi_a\alpha_a} < c_a \leq c_p - \pi_a\alpha_a$, then σ^* is characterized by $0 < v_b < v_a < 1$ under the optimal tax.

(iii) if $c_a \leq c_p - \pi_a\alpha_a$ and $c_a \leq \frac{(\pi_a\alpha_a)^2}{1-\pi_a\alpha_a}$, then σ^* is characterized by $0 < v_a < v_b < 1$ under the optimal tax.

Suppose that $0 < v_b < 1$. By Lemma A.2, $v_b = \frac{-1+\pi_s\alpha_s+\sqrt{(1-\pi_s\alpha_s)^2+4\tau\pi_s\alpha_s}}{2\pi_s\alpha_s}$. Plugging this in the welfare function $W_N(\tau) = \int_{v_b(\tau)}^{1} v - (1 - v_b(\tau))\pi_s\alpha_s v dv$ for this case, we find that the social welfare function is then given by

$$W_N(\tau) = \frac{\tau(1-\tau)}{\pi_s\alpha_s} - \frac{\tau(1-\tau)(1-3\tau)}{2(\pi_s\alpha_s)^2} + O\left(\frac{1}{(\pi_s\alpha_s)^3}\right). \tag{3.15}$$

The interior maximizing price is given by

$$\tau_N = \frac{1}{2} - \frac{3}{16\pi_s\alpha_s} - \frac{3}{64(\pi_s\alpha_s)^2} + O\left(\frac{1}{(\pi_s\alpha_s)^3}\right). \tag{3.16}$$

By substituting (3.16) into (3.15), we have that the social welfare in this case, when the solution is interior, is given by

$$W_N(\tau_N) = \frac{1}{4\pi_s\alpha_s} - \frac{1}{16(\pi_s\alpha_s)^2} + O\left(\frac{1}{(\pi_s\alpha_s)^3}\right). \tag{3.17}$$

On the other hand, suppose that $0 < v_a < v_b < v_p < 1$ is induced. By Lemma A.2, we obtain $v_a = \frac{c_a}{1-\pi_a\alpha_a}$, $v_b = \frac{c_p-c_a+\tau}{\pi_a\alpha_a} - \frac{c_p^2\pi_a\alpha_a}{(c_p-c_a+\tau)^2\pi_s\alpha_s} + \frac{2c_p^3(c_a-\tau)(\pi_a\alpha_a)^3}{(c_a-c_p-\tau)^5(\pi_s\alpha_s)^2} + O\left(\frac{1}{(\pi_s\alpha_s)^3}\right)$, and $v_p = \frac{c_p-c_a+\tau}{\pi_a\alpha_a} + \frac{c_p(\tau-c_a)\pi_a\alpha_a}{(c_p-c_a+\tau)^2\pi_s\alpha_s} - \frac{c_p^2(\tau-c_a)(c_p+c_a-\tau)(\pi_a\alpha_a)^3}{(c_p-c_a+\tau)^5(\pi_s\alpha_s)^2} + O\left(\frac{1}{(\pi_s\alpha_s)^3}\right)$.

Plugging this in the social welfare function

$$W_A(\tau) = \int_{v_a(\tau)}^{1} v dv - \Bigg(\int_{v_a(\tau)}^{v_b(\tau)} c_a + \pi_a \alpha_a v dv + \int_{v_b(\tau)}^{v_p(\tau)} (v_p(\tau) - v_b(\tau)) \pi_s \alpha_s v dv + c_p(1 - v_p(\tau)) \Bigg), \quad (3.18)$$

we find that the social welfare function is then given by

$$W_A(\tau) = \frac{1}{2}\left(1 - 2c_p + \frac{(c_p - c_a)^2 - \tau^2}{\pi_a \alpha_a} + \frac{c_a^2}{1 - \pi_a \alpha_a}\right) + \frac{c_p^2 \tau \pi_a \alpha_a}{(c_p - c_a + \tau)^2 \pi_s \alpha_s} - \frac{c_p^3 (c_p - c_a - 3\tau)(c_a - \tau)(\pi_a \alpha_a)^3}{2(c_p - c_a + \tau)^5 (\pi_s \alpha_s)^2} + O\left(\frac{1}{(\pi_s \alpha_s)^3}\right). \quad (3.19)$$

The interior maximizing price is given by

$$\tau_A = \frac{(c_p \pi_a \alpha_a)^2}{(c_p - c_a)^2 \pi_s \alpha_s} - \frac{7 c_p^3 (\pi_a \alpha_a)^4}{2(c_p - c_a)^4 (\pi_s \alpha_s)^2} + O\left(\frac{1}{(\pi_s \alpha_s)^3}\right). \quad (3.20)$$

By substituting (3.20) into (3.19), we have that the social welfare in this case, when the solution is interior, is given by

$$W_A(\tau_A) = \frac{1}{2}\left(1 - 2c_p + \frac{(c_p - c_a)^2}{\pi_a \alpha_a} + \frac{c_a^2}{1 - \pi_a \alpha_a}\right) + \frac{(c_p \pi_a \alpha_a)^3}{2(c_p - c_a)^3 (\pi_s \alpha_s)^2} + O\left(\frac{1}{(\pi_s \alpha_s)^3}\right). \quad (3.21)$$

Next, suppose that $0 < v_b < v_a < v_p < 1$ is induced. By Lemma A.2, we obtain $v_a = \frac{c_a}{1-\pi_a\alpha_a} + \frac{\tau(c_a - \tau(1-\pi_a\alpha_a))}{c_a^2 \pi_s \alpha_s} + \frac{\tau(c_a-\tau)(1-\pi_a\alpha_a)(c_a - \tau(1-\pi_a\alpha_a))(c_a - 2\tau(1-\pi_a\alpha_a))}{c_a^5 (\pi_s\alpha_s)^2} + O\left(\frac{1}{(\pi_s\alpha_s)^3}\right)$, $v_b = -\frac{(c_a-\tau)(c_a-(1-\pi_a\alpha_a)\tau}{c_a^2\pi_s\alpha_s} - \frac{\tau(c_a-\tau)(1-\pi_a\alpha_a)(-2c_a+2\tau+(c_a-2\tau)\pi_a\alpha_a)(c_a-\tau(1-\pi_a\alpha_a))}{c_a^5(\pi_s\alpha_s)^2} + \frac{c_a}{1-\pi_a\alpha_a} + O\left(\frac{1}{(\pi_s\alpha_s)^3}\right)$, and $v_p = \frac{c_p - (c_a - \tau)}{\pi_a \alpha_a}$. Again, substituting this into the social welfare function, which for this case is given by

$$W(\tau) = \int_{v_b(\tau)}^{1} v dv - \Bigg(\int_{v_b(\tau)}^{v_a(\tau)} (v_a(\tau) - v_b(\tau)) \pi_s \alpha_s v dv + \int_{v_a(\tau)}^{v_p(\tau)} c_a + \pi_a \alpha_a v dv + c_p(1 - v_p(\tau)) \Bigg), \quad (3.22)$$

we find that the social welfare function is then given by

$$W_B(\tau) = \frac{1}{2}\left(1 - 2c_p + \frac{(c_p - c_a)^2 - \tau^2}{\pi_a\alpha_a} + \frac{c_a^2}{1 - \pi_a\alpha_a}\right) +$$
$$\frac{\tau(c_a - \tau(1 - \pi_a\alpha_a))}{c_a\pi_s\alpha_s} + \frac{(c_a - \tau)\tau(1 - \pi_a\alpha_a)(c_a - \tau(1 - \pi_a\alpha_a))(c_a - 3\tau(1 - \pi_a\alpha_a))}{2c_a^4(\pi_s\alpha_s)^2} +$$
$$O\left(\frac{1}{(\pi_s\alpha_s)^2}\right). \quad (3.23)$$

The interior maximizing price is given by

$$\tau_B = \frac{\pi_a\alpha_a}{\pi_s\alpha_s} + \frac{\pi_a\alpha_a(1 - \pi_a\alpha_a)(1 - 4\pi_a\alpha_a)}{2c_a(\pi_s\alpha_s)^2} + O\left(\frac{1}{(\pi_s\alpha_s)^3}\right). \quad (3.24)$$

By substituting (3.24) into (3.23), we have that the social welfare in this case, when the solution is interior, is given by

$$W_B(\tau_B) = \frac{1}{2}\left(1 - 2c_p + \frac{(c_p - c_a)^2}{\pi_a\alpha_a} + \frac{c_a^2}{1 - \pi_a\alpha_a}\right) + \frac{\pi_a\alpha_a}{2(\pi_s\alpha_s)^2} + O\left(\frac{1}{(\pi_s\alpha_s)^3}\right). \quad (3.25)$$

Next, suppose that $0 < v_b < v_p < 1$ is induced. By Lemma A.2, we obtain $v_b = c_p + \tau - \frac{c_p^2}{(c_p+\tau)^2\pi_s\alpha_s} + O\left(\frac{1}{(\pi_s\alpha_s)^2}\right)$ and $v_p = c_p + \tau + \frac{\tau c_p}{(c_p+\tau)^2\pi_s\alpha_s} + O\left(\frac{1}{(\pi_s\alpha_s)^2}\right)$. Again, substituting this into the social welfare function, which for this case is given by

$$W_P(\tau) = \int_{v_b(\tau)}^{1} v dv - \left(\int_{v_b(\tau)}^{v_p(\tau)} (v_p(\tau) - v_b(\tau))\pi_s\alpha_s v dv + c_p(1 - v_p(\tau))\right)$$

, we find that the social welfare function is then given by

$$W_P(\tau) = \frac{1}{2}\left((1 - c_p)^2 - \tau^2\right) + \frac{c_p^2\tau}{(c_p + \tau)^2\pi_s\alpha_s} + O\left(\frac{1}{(\pi_s\alpha_s)^2}\right). \quad (3.26)$$

The interior maximizing price is given by

$$\tau_P = \frac{1}{\pi_s\alpha_s} + O\left(\frac{1}{(\pi_s\alpha_s)^2}\right). \quad (3.27)$$

By substituting (3.27) into (3.26), we have that the social welfare in this case, when the solution is interior, is given by

$$W_P(\tau_P) = \frac{1}{2}(1-c_p)^2 + O\left(\frac{1}{(\pi_s\alpha_s)^2}\right). \tag{3.28}$$

Next, suppose that $0 < v_a < 1$ is induced in equilibrium. By Lemma A.2, we obtain $v_a = \frac{c_a}{1-\pi_a\alpha_a}$. Again, substituting this into the social welfare function, which for this case is given by $W(\tau) = \int_{v_a(\tau)}^{1} vdv - \int_{v_a(\tau)}^{1} c_a + \pi_a\alpha_a vdv$, we find that the optimal social welfare function is then

$$W_a(\tau) = \frac{(1-c_a-\pi_a\alpha_a)^2}{2(1-\pi_a\alpha_a)}, \tag{3.29}$$

which doesn't depend on τ.

Next, suppose that $0 < v_a < v_b < 1$ is induced in equilibrium. By Lemma A.2, we obtain $v_a = \frac{c_a}{1-\pi_a\alpha_a}$ and $v_b = \frac{\pi_s\alpha_s - \pi_a\alpha_a + \sqrt{(\pi_s\alpha_s-\pi_a\alpha_a)^2 + 4\pi_s\alpha_s(\tau-c_a)}}{2\pi_s\alpha_s}$. Again, substituting this into the social welfare function, which for this case is given by

$$W_{ab}(\tau) = \int_{v_a(\tau)}^{1} vdv - \left(\int_{v_a(\tau)}^{v_b(\tau)} c_a + \pi_a\alpha_a vdv + \int_{v_b(\tau)}^{1} (1-v_b(\tau))\pi_s\alpha_s vdv\right), \tag{3.30}$$

we find that the optimal social welfare function is then

$$W_{ab}(\tau) = \frac{(1-c_a-\pi_a\alpha_a)^2}{2(1-\pi_a\alpha_a)} + \frac{\tau(c_a-\tau+\pi_a\alpha_a)}{\pi_s\alpha_s} - \frac{(c_a-\tau)(c_a-3\tau+\pi_a\alpha_a)(c_a-\tau+\pi_a\alpha_a)}{2(\pi_s\alpha_s)^2} + O\left(\frac{1}{(\pi_s\alpha_s)^3}\right). \tag{3.31}$$

The interior maximizing price is given by

$$\tau_{ab} = \frac{1}{2}(c_a+\pi_a\alpha_a) + \frac{(c_a-3\pi_a\alpha_a)(c_a+\pi_a\alpha_a)}{16\pi_s\alpha_s} + \frac{(c_a-\pi_a\alpha_a)(3c_a-\pi_a\alpha_a)(c_a+\pi_a\alpha_a)}{64(\pi_s\alpha_s)^2} + O\left(\frac{1}{(\pi_s\alpha_s)^3}\right). \tag{3.32}$$

By substituting (3.32) into (3.31), we have that the social welfare in this case, when the solution is interior, is given by

$$W_{ab}(\tau_{ab}) = \frac{(1-c_a-\pi_a\alpha_a)^2}{2(1-\pi_a\alpha_a)} + \frac{(c_a+\pi_a\alpha_a)^2}{4\pi_s\alpha_s} + \frac{(c_a-\pi_a\alpha_a)(c_a+\pi_a\alpha_a)^2}{16(\pi_s\alpha_s)^2} + O\left(\frac{1}{(\pi_s\alpha_s)^3}\right). \quad (3.33)$$

Finally, suppose that $0 < v_b < v_a < 1$ is induced in equilibrium. By Lemma A.2, we have $v_a = \frac{c_a}{1-\pi_a\alpha_a} + \frac{\tau(c_a-\tau(1-\pi_a\alpha_a))}{c_a^2\pi_s\alpha_s} + \frac{\tau(c_a-\tau)(1-\pi_a\alpha_a)(c_a-\tau(1-\pi_a\alpha_a))(c_a-2\tau(1-\pi_a\alpha_a))}{c_a^5(\pi_s\alpha_s)^2} + O\left(\frac{1}{(\pi_s\alpha_s)^3}\right)$ and $v_b = \frac{c_a}{1-\pi_a\alpha_a} - \frac{(c_a-\tau)(c_a-\tau(1-\pi_a\alpha_a))}{c_a^2\pi_s\alpha_s} - \frac{\tau(c_a-\tau)(1-\pi_a\alpha_a)(-2c_a+2\tau+(c_a-2\tau)\pi_a\alpha_a)(c_a-\tau(1-\pi_a\alpha_a))}{c_a^5(\pi_s\alpha_s)^2} + O\left(\frac{1}{(\pi_s\alpha_s)^3}\right)$. Again, substituting this into the social welfare function, which for this case is given by

$$W_{ba}(\tau) = \int_{v_b(\tau)}^{1} v dv - \left(\int_{v_b(\tau)}^{v_a(\tau)} (v_a(\tau)-v_b(\tau))\pi_s\alpha_s v dv + \int_{v_a(\tau)}^{1} c_a + \pi_a\alpha_a v dv\right),$$

we find that the optimal social welfare function is then

$$W_{ba}(\tau) = \frac{(1-c_a-\pi_a\alpha_a)^2}{2(1-\pi_a\alpha_a)} + \frac{\tau(c_a-\tau(1-\pi_a\alpha_a))}{c_a\pi_s\alpha_s} + \frac{\tau(c_a-\tau)(1-\pi_a\alpha_a)(c_a-\tau(1-\pi_a\alpha_a))(c_a-3\tau(1-\pi_a\alpha_a))}{2c_a^4(\pi_s\alpha_s)^2} + O\left(\frac{1}{(\pi_s\alpha_s)^2}\right). \quad (3.34)$$

The interior maximizing price is given by

$$\tau_{ba} = \frac{1}{2}\left(\frac{c_a}{1-\pi_a\alpha_a}\right) - \frac{1-3\pi_a\alpha_a}{16(1-\pi_a\alpha_a)\pi_s\alpha_s} - \frac{(2-\pi_a\alpha_a)(1-3\pi_a\alpha_a)}{(64c_a(1-\pi_a\alpha_a)(\pi_s\alpha_s)^2} + O\left(\frac{1}{(\pi_s\alpha_s)^3}\right). \quad (3.35)$$

By substituting (3.35) into (3.34), we have that the social welfare in this case, when the solution is interior, is given by

$$W_{ba}(\tau_{ba}) = \frac{(1-c_a-\pi_a\alpha_a)^2}{2(1-\pi_a\alpha_a)} + \frac{c_a}{4(1-\pi_a\alpha_a)\pi_s\alpha_s} - \frac{1-2\pi_a\alpha_a}{32(1-\pi_a\alpha_a)(\pi_s\alpha_s)^2} + O\left(\frac{1}{(\pi_s\alpha_s)^3}\right). \quad (3.36)$$

To prove Lemma 3, we proceed as follows. We first focus on the case when $0 < v_b < v_a < v_p < 1$ is induced in equilibrium. We find the conditions under which the interior optimal price for this case indeed induces the conjectured market structure. Further, we show that the induced welfare is greater than the maximal welfare of the other cases (under their respective optimal taxes) when the conditions specified in the lemma are met. We then proceed to the remaining cases of $0 < v_b < v_a < 1$ and $0 < v_a < v_b < 1$ using the same steps.

Suppose that $0 < v_b < v_a < v_p < 1$ is induced in equilibrium. For the optimal tax of this case, in (3.24), to induce this case, we need to have that (3.24) satisfies (I) of Lemma A.2 (when $\tau \leq c_a$) for sufficient high $\pi_s\alpha_s$. Omitting the algebra, the condition under which the optimal tax of $0 < v_b < v_a < v_p < 1$, given in (3.24), indeed induces the market structure $0 < v_b < v_a < v_p < 1$ is given by

$$c_p - \pi_a\alpha_a < c_a < c_p(1 - \pi_a\alpha_a).$$

Next, we show that under the above conditions, the interior optimal welfare of $0 < v_b < v_a < v_p < 1$ dominates the interior optimal welfares (and possible boundary extrema) in all the other cases. Just by comparing the expressions, we see that the interior optimal welfare of $0 < v_b < v_a < v_p < 1$ dominates the interior optimal welfares of all the cases except for possibly $0 < v_a < v_b < v_p < 1$. Specifically, (3.25) is greater than (3.36), (3.33), (3.29), (3.28), and (3.17). To complete the proof for this case, we next show that the interior solution for $0 < v_a < v_b < v_p < 1$ can't induce that market structure. To show that the interior solution for $0 < v_a < v_b < v_p < 1$ doesn't induce $0 < v_a < v_b < v_p < 1$, note that from (I) of Lemma A.2 (when $\tau > c_a$), one of the conditions is $\tau > c_a$. However, looking at (3.20), we see that for sufficiently high $\pi_s\alpha_s$, we'll have that $\tau_A < c_a$. Therefore, the interior solution for $0 < v_a < v_b < v_p < 1$ can't induce that market structure. In other words, if, given the parameters, $0 < v_a < v_b < v_p < 1$ can be induced by some τ, then the maximal welfare of $0 < v_a < v_b < v_p < 1$ occurs at a boundary point for that market structure. It follows that (3.25) dominates the boundary between $0 < v_a < v_b < v_p < 1$ and any other market structures, since it dominates the interior optimal welfares in the other structures. This proves part

(i) of Lemma 3.

Next, suppose that $0 < v_b < v_a < 1$ is induced in equilibrium. For the optimal tax of this case, in (3.35), to induce this case, we need to have that (3.35) satisfies (II) of Lemma A.2 (when $\tau \le c_a$) for sufficient high $\pi_s\alpha_s$. Omitting the algebra, the condition under which the optimal tax of $0 < v_b < v_a < 1$, given in (3.35), indeed induces the market structure $0 < v_b < v_a < 1$ is given by $\pi_a\alpha_a < \frac{1}{2}$ and $c_a \le \frac{2(c_p - \pi_a\alpha_a)(1-\pi_a\alpha_a)}{1-2\pi_a\alpha_a}$.

We show next that under the conditions specified in the lemma for this case, the interior optimal welfare of $0 < v_b < v_a < 1$ dominates the interior optimal welfares (and possible boundary extrema) in all the other cases. Just by comparing the expressions, we see that the interior optimal welfare of $0 < v_b < v_a < 1$, given in (3.36), always dominates the interior optimal welfares of cases $0 < v_a < 1$ and $0 - v_b - 1$, given in (3.29) and (3.17) respectively. Moreover, when $c_a \le c_p - \pi_a\alpha_a$, we'll have that the interior optimal welfare of $0 < v_b < v_a < 1$ is greater than the interior optimal welfare of $0 - v_b - v_p - 1$, given in (3.28). We'll now show that under the conditions specified for this case, namely $\frac{(\pi_a\alpha_a)^2}{1-\pi_a\alpha_a} < c_a \le c_p - \pi_a\alpha_a$, it will be the case that there's no τ which can induce either $0 < v_b < v_a < v_p < 1$ or $0 < v_a < v_b < v_p < 1$. Specifically, using part (I) of Lemma A.2 for when $\tau \le c_a$, for the case $0 < v_b < v_a < v_p < 1$ to be induced for some τ when $\pi_s\alpha_s$ gets sufficiently big, we need $(1+c_p)(1-\pi_a\alpha_a) > c_a > c_p - \pi_a\alpha_a$ to hold. Similarly, using part (I) of Lemma A.2 for when $\tau > c_a$, for the case $0 < v_b < v_a < v_p < 1$ to be induced for some τ when $\pi_s\alpha_s$ gets sufficiently big, we also need $c_a > c_p - \pi_a\alpha_a$. Therefore, when $c_a \le c_p - \pi_a\alpha_a$ (as specified in the lemma for this case), it'll be the case that no τ can induce either $0 < v_b < v_a < v_p < 1$ or $0 < v_a < v_b < v_p < 1$ in equilibrium. Lastly, we need to compare the interior optimal welfare of $0 < v_b < v_a < 1$ against the interior optimal welfare of $0 < v_a < v_b < 1$. Note that the conditions for the interior optimal tax of $0 < v_a < v_b < 1$ to indeed induce $0 < v_a < v_b < 1$ are given by $2c_p \ge c_a + \pi_a\alpha_a$ and $c_a < \pi_a\alpha_a$. By comparing (3.36) and (3.33), we see that the interior optimal welfare of $0 < v_b < v_a < 1$ dominates the interior optimal welfare of $0 < v_a < v_b < 1$ when $c_a \ge \frac{(\pi_a\alpha_a)^2}{1-\pi_a\alpha_a}$. Since $\frac{(\pi_a\alpha_a)^2}{1-\pi_a\alpha_a} < \pi_a\alpha_a$ and $c_a + \pi_a\alpha_a > \frac{1}{2}(c_a + \pi_a\alpha_a)$ for $\pi_a\alpha_a < \frac{1}{2}$, it follows that we need to have $c_a \ge \frac{(\pi_a\alpha_a)^2}{1-\pi_a\alpha_a}$ as a condition for $0 < v_b < v_a < 1$

to be induced in equilibrium. We note that $\pi_a\alpha_a < \frac{1}{2}$ must hold in order for $\frac{(\pi_a\alpha_a)^2}{1-\pi_a\alpha_a} < c_a \leq c_p - \pi_a\alpha_a$ to hold for some c_a and c_p, and $\frac{(\pi_a\alpha_a)^2}{1-\pi_a\alpha_a} < c_a \leq c_p - \pi_a\alpha_a$ is a subset of $c_a \leq \frac{2(c_p-\pi_a\alpha_a)(1-\pi_a\alpha_a)}{1-2\pi_a\alpha_a}$, so that when the conditions of the lemma hold, $0 < v_b < v_a < 1$ will be induced in equilibrium. This completes the proof for this case.

Lastly, suppose that $0 < v_a < v_b < 1$ is induced in equilibrium. For the optimal tax of this case, in (3.32), to induce this case, we need to have that (3.32) satisfies (II) of Lemma A.2 (when $\tau > c_a$) for sufficient high $\pi_s\alpha_s$. Omitting the algebra, the conditions under which the optimal tax of $0 < v_a < v_b < 1$, given in (3.32), indeed induces the market structure $0 < v_a < v_b < 1$ are given by $c_a < \pi_a\alpha_a$ and $c_a \leq 2c_p - \pi_a\alpha_a$.

We show next that under the conditions specified in the lemma for this case, the interior optimal welfare of $0 < v_a < v_b < 1$ dominates the interior optimal welfares (and possible boundary extrema) in all the other cases. In the same way as the previous case, we have that the interior optimal welfare of $0 < v_a < v_b < 1$, given in (3.33), always dominates the interior optimal welfares of cases $0 < v_a < 1$ and $0 < v_b < 1$, given in (3.29) and (3.17) respectively. Moreover, when $c_a \leq c_p - \pi_a\alpha_a$, we'll have that the interior optimal welfare of $0 < v_a < v_b < 1$ is greater than the interior optimal welfare of $0 - v_b - v_p - 1$, given in (3.28). For the same reason as in the previous case of $0 < v_b < v_a < 1$, we have that there's no τ which can induce either $0 < v_b < v_a < v_p < 1$ or $0 < v_a < v_b < v_p < 1$. Therefore, when $c_a \leq c_p - \pi_a\alpha_a$ (as specified in the lemma for this case), it'll be the case that no τ can induce either $0 < v_b < v_a < v_p < 1$ or $0 < v_a < v_b < v_p < 1$ in equilibrium. Lastly, we need to compare the interior optimal welfare of $0 < v_a < v_b < 1$ against the interior optimal welfare of $0 < v_b < v_a < 1$. Again, from the previous case, we have that the condition is $c_a < \frac{(\pi_a\alpha_a)^2}{1-\pi_a\alpha_a}$. Note that when $\pi_a\alpha_a \leq \frac{1}{2}$, we have $\frac{(\pi_a\alpha_a)^2}{1-\pi_a\alpha_a} < \pi_a\alpha_a$ and when $\pi_a\alpha_a > \frac{1}{2}$, we have $c_p - \pi_a\alpha_a < \pi_a\alpha_a$. Therefore, $c_a < \frac{(\pi_a\alpha_a)^2}{1-\pi_a\alpha_a}$ and $c_a \leq c_p - \pi_a\alpha_a$ imply $c_a < \pi_a\alpha_a$. Also, $c_a \leq c_p - \pi_a\alpha_a$ implies $c_a \leq 2c_p - \pi_a\alpha_a$. From the above, we have that the conditions $c_a < \frac{(\pi_a\alpha_a)^2}{1-\pi_a\alpha_a}$ and $c_a \leq c_p - \pi_a\alpha_a$ imply that the optimal tax of $0 < v_a < v_b < 1$ indeed induces that market structure in equilibrium. This completes the proof for this case and concludes the lemma. $\square$

3.7.2 Proofs of Propositions

Proof of Proposition 1: From Lemma A.1, it becomes clear that the absence of an automated patching option is a special case characterized by:

(A) If $\pi_s\alpha_s \leq c_p$, then $0 = v_b < v_p = 1$;

(B) If $c_p < \pi_s\alpha_s \leq 1/c_p$, then $0 = v_b < v_p < 1$, where $v_p = \sqrt{\frac{c_p}{\pi_s\alpha_s}}$;

(C) If $\pi_s\alpha_s > 1/c_p$, then $0 < v_b < v_p < 1$, where $v_b = c_p - 1/\pi_s\alpha_s$ and $v_p = c_p$.

For part (i) of the proposition, suppose that $c_p < \pi_s\alpha_s \leq \frac{c_p(\pi_a\alpha_a)^2}{(c_p-c_a)^2}$. By part (II) of Lemma A.1, $u = \sqrt{\frac{c_p}{\pi_s\alpha_s}}$. Part (B) above is also satisfied since $\frac{c_p(\pi_a\alpha_a)^2}{(c_p-c_a)^2} < 1/c_p$, and hence $\tilde{u} = \sqrt{\frac{c_p}{\pi_s\alpha_s}}$. Suppose $\pi_s\alpha_s \leq c_p$. By part (I) of Lemma A.1 and part (A) above, $u = \tilde{u} = 1$. For part (ii) of the proposition, by part (VII) of Lemma A.1, $u = \frac{\pi_a\alpha_a + \sqrt{(\pi_a\alpha_a)^2 + 4\pi_s\alpha_s c_a}}{2\pi_s\alpha_s}$. Because $c_p < \frac{c_p(\pi_a\alpha_a)^2}{(c_p-c_a)^2} < \frac{1}{c_p} < \frac{1-\pi_a\alpha_a}{c_a}$, $\tilde{u}$ satisfies either $\tilde{u} = \sqrt{\frac{c_p}{\pi_s\alpha_s}}$ or $\tilde{u} = 1/\pi_s\alpha_s$. In either case, $u < \tilde{u}$. For part (iii) of the proposition, by part (IV) of Lemma A.1 and part (C) above, $u = \tilde{u} = 1/\pi_s\alpha_s$, which completes the proof. ■

Proof of Proposition 2: By Lemma 2 and Lemma 3, for sufficiently high $\pi_s\alpha_s$, we have the market structure $0 < v_b < v_a < v_p < 1$ being induced both under the status quo setting without the tax as well as under the optimal tax when $c_p - \pi_a\alpha_a < c_a < c_p(1 - \pi_a\alpha_a)$. The social welfare function of this case is given by (3.23). The welfare under the status quo case of no tax is given as

$$W_{B,\text{ Status Quo}} = \frac{1}{2}\left(1 - 2c_p + \frac{(c_a - c_p)^2}{\pi_a\alpha_a} + \frac{c_a^2}{1 - \pi_a\alpha_a}\right) + O\left(\frac{1}{(\pi_s\alpha_s)}\right),$$

which is the limit of (3.23) as $\tau \to 0$. Comparing this expression with (3.25), the increase in social welfare upon imposing the optimal tax is given by $\frac{\pi_a\alpha_a}{2(\pi_s\alpha_s)^2} + O\left(\frac{1}{(\pi_s\alpha_s)^3}\right)$.

Next, the size of the unpatched population for this market structure is given by $u(\sigma^*|\tau) = v_a(\tau) - v_b(\tau)$, and the size of the automated patching population is given by $a(\sigma^*|\tau) = v_p(\tau) - v_a(\tau)$. Taking the limit of these as $\tau \to 0$, the sizes of the unpatched and automated populations under the status quo setting are given as $u_B = \frac{1}{\pi_s\alpha_s} + O\left(\frac{1}{(\pi_s\alpha_s)^3}\right)$ and $a_B = \frac{c_p - c_a}{\pi_a\alpha_a} - \frac{c_a}{1-\pi_a\alpha_a} + \frac{1}{\pi_s\alpha_s} + \frac{1-\pi_a\alpha_a(7-4\pi_a\alpha_a)}{2c_a(\pi_s\alpha_s)^2} + O\left(\frac{1}{(\pi_s\alpha_s)^3}\right)$. Comparing these to their respective population sizes under the optimal tax, $\tau = \tau_B$ (given in (3.24)), we establish that the size of the unpatched population decreases by $\frac{\pi_a\alpha_a(1-\pi_a\alpha_a)}{c_a(\pi_s\alpha_s)^2} + O\left(\frac{1}{(\pi_s\alpha_s)^3}\right)$ and the size of the automated patching population increases by $\frac{1}{\pi_s\alpha_s} + \frac{1-\pi_a\alpha_a(7-4\pi_a\alpha_a)}{2c_a(\pi_s\alpha_s)^2} + O\left(\frac{1}{(\pi_s\alpha_s)^3}\right)$.

Under both the status quo setting and under taxed patching rights, the total security loss from unpatched usage as a function of τ is given by $SL_B(\tau) = \int_{v_b(\tau)}^{v_a(\tau)} (v_a(\tau) - v_b(\tau))\pi_s\alpha_s v dv$ and total patching cost from automated patching is given by $AL_B(\tau) = \int_{v_a(\tau)}^{v_p(\tau)} c_a + \pi_a\alpha_a v dv$. Then the decrease in unpatched losses is given by $SL_B(0) - SL_B(\tau_B) = \frac{2\pi_a\alpha_a}{(\pi_s\alpha_s)^2} + O\left(\frac{1}{(\pi_s\alpha_s)^3}\right)$, and the increase in automated patching losses is given by $AL_B(\tau_B) - AL_B(0) = \frac{c_p}{\pi_s\alpha_s} + O\left(\frac{1}{(\pi_s\alpha_s)^2}\right)$. ■

Proof of Proposition 3: By Lemma 3, it can be seen that the magnitude of the optimal taxes imposed on patching rights significantly increases as c_a becomes less than or equal to $c_p - \pi_a\alpha_a$. The structure of the remaining proof is similar to the proof of Proposition 2.

(i) By Lemma 2 and Lemma 3, for sufficiently high $\pi_s\alpha_s$, we have the market structure $0 < v_b < v_a < 1$ being induced both under status quo setting without tax as well as under the optimal tax when $\frac{(\pi_a\alpha_a)^2}{1-\pi_a\alpha_a} < c_a \le c_p - \pi_a\alpha_a$. The social welfare function of this case is given by (3.34). The welfare under the status quo case of no tax is given as

$$W_{ba,\text{ Status Quo}} = \frac{(1 - c_a - \pi_a\alpha_a)^2}{2(1 - \pi_a\alpha_a)} + O\left(\frac{1}{(\pi_s\alpha_s)^3}\right), \tag{3.37}$$

which is the limit of (3.34) as $\tau \to 0$. Comparing this expression with (3.36), the increase in social welfare upon imposing the optimal tax is given by $\frac{c_a}{4(1-\pi_a\alpha_a)\pi_s\alpha_s} + O\left(\frac{1}{(\pi_s\alpha_s)^2}\right)$.

Next, the size of the unpatched population for this market structure is given by $u(\sigma^*|\tau) = v_a(\tau) - v_b(\tau)$, and the size of the automated patching population is given by $a(\sigma^*|\tau) = 1 - v_a(\tau)$. Taking the limit of these as $\tau \to 0$, the sizes of the unpatched and automated populations under the status quo setting are given as $u_{ba} = \frac{1}{\pi_s\alpha_s} + O\left(\frac{1}{(\pi_s\alpha_s)^3}\right)$ and $a_{ba} = 1 - \frac{c_a}{1-\pi_a\alpha_a} + O\left(\frac{1}{(\pi_s\alpha_s)^3}\right)$. Comparing these to their respective population sizes under the optimal tax, $\tau = \tau_{ba}$ (given in (3.35)), we establish that the size of the unpatched population decreases by $\frac{1}{2\pi_s\alpha_s} + O\left(\frac{1}{(\pi_s\alpha_s)^2}\right)$ and the size of the automated patching population decreases by $\frac{1}{4(1-\pi_a\alpha_a)\pi_s\alpha_s} + O\left(\frac{1}{(\pi_s\alpha_s)^3}\right)$.

Under both the status quo setting and under taxed patching rights, the total security loss from unpatched usage as a function of τ is given by $SL_{ba}(\tau) = \int_{v_b(\tau)}^{v_a(\tau)} (v_a(\tau) - v_b(\tau))\pi_s\alpha_s v dv$ and total patching cost from automated patching is given by $AL_{ba}(\tau) = \int_{v_a(\tau)}^{1} c_a + \pi_a\alpha_a v dv$. Then the decrease in unpatched losses is given by $SL_{ba}(0) - SL_{ba}(\tau_{ba}) = \frac{3c_a}{4(1-\pi_a\alpha_a)\pi_s\alpha_s} + O\left(\frac{1}{(\pi_s\alpha_s)^2}\right)$, and the increase in automated patching losses is given by $AL_{ba}(0) - AL_{ba}(\tau_{ba}) = \frac{c_a}{4(1-\pi_a\alpha_a)^2\pi_s\alpha_s} + O\left(\frac{1}{(\pi_s\alpha_s)^2}\right)$.

(ii) Similarly, by Lemma 2 and Lemma 3, for sufficiently high $\pi_s\alpha_s$ and $c_a \leq \min\left(c_p - \pi_a\alpha_a, \frac{(\pi_a\alpha_a)^2}{1-\pi_a\alpha_a}\right)$, we have the market structure $0 < v_b < v_a < 1$ being induced under status quo setting while $0 < v_a < v_b < 1$ is induced under the optimal tax. The welfare under the status quo case of no tax was found in the previous case. Comparing (3.37) with the welfare under the optimal tax (3.33), we find that the increase in social welfare upon imposing the optimal tax is given by $\frac{(c_a+\pi_a\alpha_a)^2}{4\pi_s\alpha_s} + O\left(\frac{1}{(\pi_s\alpha_s)^2}\right)$.

Next, the size of the unpatched population for this market structure under the optimal tax is given by $u(\sigma^*|\tau) = 1 - v_b(\tau)$, and the size of the automated patching population is given by $a(\sigma^*|\tau) = v_b(\tau) - v_a(\tau)$. Evaluating these under the optimal tax, $\tau = \tau_{ab}$ (given in (3.32)), the equilibrium sizes of the unpatched and automated patching populations under the optimal tax are $u_{ab} = \frac{c_a+\pi_a\alpha_a}{2\pi_s\alpha_s} + O\left(\frac{1}{(\pi_s\alpha_s)^2}\right)$ and $a_{ab} = \left(1 - \frac{c_a}{1-\pi_a\alpha_a}\right) - \frac{c_a+\pi_a\alpha_a}{2\pi_s\alpha_s} + O\left(\frac{1}{(\pi_s\alpha_s)^2}\right)$. From the previous case, the sizes of the unpatched and automated populations under the status quo setting are given as $u_{ba} = \frac{1}{\pi_s\alpha_s} + O\left(\frac{1}{(\pi_s\alpha_s)^3}\right)$ and $a_{ba} = 1 - \frac{c_a}{1-\pi_a\alpha_a} + O\left(\frac{1}{(\pi_s\alpha_s)^3}\right)$. Comparing these, we establish that the size of the unpatched population decreases by $\frac{2-c_a-\pi_a\alpha_a}{2\pi_s\alpha_s} + O\left(\frac{1}{(\pi_s\alpha_s)^2}\right)$ and the size of the automated patching population decreases by $\frac{c_a+\pi_a\alpha_a}{2\pi_s\alpha_s} + O\left(\frac{1}{(\pi_s\alpha_s)^2}\right)$ under the optimal tax.

Under taxed patching rights, the total security loss from unpatched usage as a function of τ is given by $SL_{ab}(\tau) = \int_{v_b(\tau)}^{1} (1 - v_b(\tau))\pi_s\alpha_s v dv$ and total patching cost from automated patching is given by $AL_{ab}(\tau) = \int_{v_a(\tau)}^{v_b(\tau)} c_a + \pi_a\alpha_a v dv$. Then the change in unpatched losses is given by $SL_{ba}(0) - SL_{ab}(\tau_{ab}) = \frac{1}{\pi_s\alpha_s}\left(\frac{c_a}{1-\pi_a\alpha_a} - \frac{1}{4}(c_a + \pi_a\alpha_a)^2\right) + O\left(\frac{1}{(\pi_s\alpha_s)^2}\right)$, and the increase in automated patching losses is given by $AL_{ba}(0) - AL_{ab}(\tau_{ab}) = \frac{(c_a+\pi_a\alpha_a)^2}{2\pi_s\alpha_s} + O\left(\frac{1}{(\pi_s\alpha_s)^2}\right)$. Note that $SL_{ba}(0) - SL_{ab}(\tau_{ab}) > 0$ if $c_a > \frac{2(1-\sqrt{1-\pi_a\alpha_a(1-\pi_a\alpha_a)})-\pi_a\alpha_a(1-\pi_a\alpha_a)}{1-\pi_a\alpha_a}$ and $SL_{ba}(0) - SL_{ab}(\tau_{ab}) \leq 0$ otherwise. $\blacksquare$

Chapter 3, in part, is currently being prepared for submission for publication. Terrence August, Duy Dao, and Kihoon Kim. The dissertation author was the primary researcher and author of this material.

Bibliography

Adams, W. J. and J. L. Yellen (1976). Commodity bundling and the burden of monopoly. *The Quarterly Journal of Economics 90*(3), 475–498.

Anderson, R. (2002). Security in open versus closed systems – The dance of Boltzmann, Coase and Moore. In *Proceedings of the Conference on Open Source Software Economics*, Toulouse, France.

Arora, A., R. Telang, and H. Xu (2008). Optimal policy for software vulnerability disclosure. *Management Science 54*(4), 642–656.

August, T., R. August, and H. Shin (2014). Designing user incentives for cybersecurity. *Communications of the ACM 57*(11), 43–46.

August, T., M. F. Niculescu, and H. Shin (2014). Cloud implications on software network structure and security risks. *Information Systems Research 25*(3), 489–510.

August, T. and T. I. Tunca (2006). Network software security and user incentives. *Management Science 52*(11), 1703–1720.

August, T. and T. I. Tunca (2011). Who should be responsible for software security? A comparative analysis of liability policies in network environments. *Management Science 57*(5), 934–959.

Bakos, Y. and E. Brynjolfsson (1999). Bundling information goods: Pricing, profits, and efficiency. *Management Science 45*(12), 1613–1630.

Beattie, S., S. Arnold, C. Cowan, P. Wagle, and C. Wright (2002, November). Timing the application of security patches for optimal uptime. *Proceedings of LISA '02: Sixteenth Systems Administration Conference*, 233–242.

Beattie, S., S. Arnold, C. Cowan, P. Wagle, C. Wright, and A. Shostack (2002). Timing the application of security patches for optimal uptime. *Proceedings of the 16th USENIX conference on system administration*, 233–242.

Beres, Y. and J. Griffin (2009). Optimizing network patching policy decisions. *HP Laboratories*.

Bhargava, H. K. (2013). Mixed bundling of two independently valued goods. *Management Science 59*(9), 2170–2185.

Bhargava, H. K. and V. Choudhary (2001). Information goods and vertical differentiation. *Journal of Management Information Systems 18*(2), 89–106.

Bhargava, H. K. and V. Choudhary (2008). Research note: When is versioning optimal for information goods? *Management Science 54*(5), 1029–1035.

Bloor, B. (2003). The patch problem: It's costing your business real dollars.

Brillinger, D. (1966). The analyticity of the roots of a polynomial as functions of the coefficients. *Mathematics Magazine 39*(3), 145–147.

Brito, D. L., E. Sheshinski, and M. D. Intriligator (1991, June). Externalities and compulsory vaccinations. *J. Public Econ. 45*(1), 69–90.

Cavusoglu, H., H. Cavusoglu, and S. Raghunathan (2007). Efficiency of vulnerability disclosure mechanisms to disseminate vulnerability knowledge. *IEEE Transactions on Software Engineering 33*(3), 171–185.

Cavusoglu, H., H. Cavusoglu, and J. Zhang (2008). Security patch management: Share the burden or share the damage? *Management Science 54*(4), 657–670.

CCIRC (2014, October). Top 4 strategies to mitigate targeted cyber intrusions. Available at http://www.publicsafety.gc.ca/cnt/ntnl-scrt/cbr-scrt/tp-strtgs-eng.aspx.

Chen, F. (2001). Market segmentation, advanced demand information, and supply chain performance. *Manufacturing & Service Operations Management 3*(1), 53–67.

Choi, J. P., C. Fershtman, and N. Gandal (2010). Network security: Vulnerabilities and disclosure policy. *Journal of Industrial Economics 58*(4), 868–894.

Cormen, T. H., C. E. Leiserson, R. L. Rivest, and C. Stein (2009, July). *Introduction to Algorithms* (3 ed.). U.S.A.: The MIT Press.

Debo, L. G., L. B. Toktay, and L. N. V. Wassenhove (2005). Market segmentation and product technology selection for remanufacturable products. *Management Science 51*(8), 1193–1205.

Desai, P. S. (2001). Quality segmentation in spatial markets: When does cannibalization affect product line design? *Marketing Science 20*(3), 265–283.

Dey, D., A. Lahiri, and G. Zhang (2015). Optimal policies for security patch management. *Journal on Computing 27*(3), 462–477.

Diffie, W. (2003, Jan). Risky business: Keeping security a secret. *ZDNet News*.

Eckalbar, J. C. (2010). Closed-form solutions to bundling problems. *Journal of Economics & Management Strategy 19*(2), 513–544.

Fitzgerald, D. (2015, May). Level 3 tries to waylay hackers. The Wall Street Journal.

Forbath, T., P. Kalaher, and T. O'Grady (2005). The total cost of security patch management. *Wipro Technologies, Ltd.*.

Gabszewicz, J. J., A. Shaked, J. Sutton, and J.-F. Thisse (1986). Segmenting the market: The monopolist's optimal product mix. *Journal of Economic Theory 39*(2), 273–289.

Greenberg, A. (2015, July). Hackers remotely kill a jeep on the highway - with me in it. WIRED.

Greenberg, A. (2017, May). The wannacry ransomware hackers made some real amateur mistakes. Fortune.

Gupta, A. and D. Zhdanov (2012). Growth and sustainability of managed security services networks: An economic perspective. *MIS Quarterly 36*(4), 1109–1130.

Hick, M. (2012, June). Apple acknowledges Flashback malware with update. Huffington Post UK.

Hoepman, J.-H. and B. Jacobs (2007). Increased security through open source. *Communications of the ACM 50*(1), 79–83.

HP Security Research (2015, February). Hp cyber risk report 2015. Available at http://www8.hp.com/us/en/software-solutions/cyber-risk-report-security-vulnerability/.

Johnson, J. P. and D. P. Myatt (2003, Jun). Multiproduct quality competition: Fighting brands and product line pruning. *The American Economic Review 93*(3), 748–774.

Kannan, K., M. S. Rahman, and M. Tawarmalani (2013, Jan). Economic and policy implications of restricted patch distribution. Working paper, Purdue University and University of Calgary.

Krill, P. (2015, April). Oracle to end publicly available security fixes for java 7 this month. InfoWorld.

Laffont, J.-J. and J. Tirole (1988). The dynamics of incentive contracts. *Econometrica 56*(5), 1153–1175.

Lahiri, A. (2012). Revisiting the incentive to tolerate illegal distribution of software products. *Decision Support Systems 53*(2), 357–367.

Levchenko, K., A. Pitsillidis, N. Chachra, B. Enright, M. Félegyházi, C. Grier, T. Halvorson, C. Kanich, C. Kreibich, H. Liu, D. McCoy, N. Weaver, V. Paxson, G. M. Voelker, and S. Savage (2011). Click trajectories: End-to-end analysis of the spam value chain. In *Proceedings of the 2011 IEEE Symposium on Security and Privacy*, SP '11, Washington, DC, USA, pp. 431–446. IEEE Computer Society.

Li, L., R. D. McKelvey, and T. Page (1987). Optimal research for Cournot oligopolists. *Journal of Economic Theory 42*(1), 140–166.

Lohr, S. and L. Alderman (2017, May). The fallout from a global cyberattack: 'a battle we're fighting every day'. The New York Times.

MacLeod, W. B. and J. M. Malcomson (1993). Investments, holdup, and the form of market contracts. *American Economic Review 83*(4), 811–837.

Maskin, E. and J. Riley (1984). Monopoly with incomplete information. *RAND Journal of Economics 15*(2), 171–196.

McAfee, R. P., J. McMillan, and M. D. Whinston (1989, May). Multiproduct monopoly, commodity bundling, and correlation of values. *The Quarterly Journal of Economics 104*(2), 371–383.

Microsoft (2017a). New ransomware, old techniques: Petya adds worm capabilities. Microsoft Malware Protection Blog Post. Available at https://blogs.technet.microsoft.com/mmpc/2017/06/27/new-ransomware-old-techniques-petya-adds-worm-capabilities/.

Microsoft (2017b). Wannacrypt ransomware worm targets out-of-date systems. Microsoft Malware Protection Blog Post. Available at https://blogs.technet.microsoft.com/mmpc/2017/05/12/wannacrypt-ransomware-worm-targets-out-of-date-systems/.

Miller, P. D. (2006). *Applied Asymptotic Analysis*, Volume 75 of *Graduate Studies in Mathematics*. Providence, Rhode Island: American Mathematical Society.

Moorthy, K. S. (1984). Market segmentation, self-selection, and product line design. *Marketing Science 3*(4), 288–307.

Moorthy, K. S. and I. Png (1992). Market segmentation, cannibalization, and the timing of product introductions. *Management Science 38*(3), 345–359.

Muller, H. M. (2000). Asymptotic efficiency in dynamic principal-agent problems. *Journal of Economic Theory 91*, 292–301.

Mussa, M. and S. Rosen (1978). Monopoly and product quality. *Journal of Economic Theory 18*(2), 301–317.

Netessine, S. and T. A. Taylor (2007). Product line design and production technology. *Marketing Science 26*(1), 101–117.

Newman, J. (2015, July). Windows 10 will force automatic updates for all Home users. PCWorld.

OPSWAT (2014, October). Antivirus and operating system report. Available at https://www.opswat.com/resources/reports/antivirus-october-2014.

Pei, P. P.-E., D. Simchi-Levi, and T. I. Tunca (2011). Sourcing flexibility, spot trading, and procurement contract structure. *Operations Research 59*(3), 578–601.

Pesendorfer, W. and J. M. Swinkels (2000). Efficiency and information aggregation in auctions. *American Economic Review 90*(3), 499–525.

Png, I. P. and Q. Wang (2009). Information security: Facilitating user precautions vis-à-vis enforcement against attackers. *Journal of Management Information Systems 26*(2), 97–121.

Prasad, A., R. Venkatesh, and V. Mahajan (2010). Optimal bundling of technological products with network externality. *Management Science 56*(12), 2224–2236.

Ransbotham, S., S. Mitra, and J. Ramsey (2012). Are markets for vulnerabilities effective? *MIS Quarterly 36*(1), 43–64.

Schmalensee, R. (1984, January). Gaussian demand and commodity bundling. *The Journal of Business 57*(1), S211–S230.

Schneider, F. B. (2000). Open source in security: Visiting the bizarre. In *Proceedings 2000 IEEE Symposium on Security and Privacy*, pp. 126–127. IEEE Computer Society.

Schneier, B. (1999, Sep). Open source and security. *Crypto-Gram Newsletter*.

Stremersch, S. and G. J. Tellis (2002, January). Strategic bundling of products and prices: A new synthesis for marketing. *Journal of Marketing 66*, 55–72.

Tunca, T. I. and S. A. Zenios (2006). Supply auctions and relational contracts for procurement. *M&SOM 8*(1), 43–67.

US-CERT (2015). Top 30 targeted high risk vulnerabilities. US-CERT Alert (TA15-119A). Available at https://www.us-cert.gov/ncas/alerts/TA15-119A.

Venkatesh, R. and W. A. Kamakura (2003). Optimal bundling and pricing under a monopoly: Contrasting complements and substitutes from independently valued products. *The Journal of Business 76*(2), 211–231.

Villas-Boas, J. M. (2004). Communication strategies and product line design. *Marketing Science 23*(3), 304–316.

Villas-Boas, J. M. (2009). Product variety and endogenous pricing with evaluation costs. *Management Science 55*(8), 1338–1346.

Wheeler, D. A. (2003, Mar). *Secure Programming for Linux and Unix HOWTO*.

Witten, B., C. Landwehr, and M. Caloyannides (2001). Does open source improve system security? *IEEE Software 18*(1), 57–61.