University of Bern

New Insights for Regulation in Competition Policy

Philipp Brunner

Inauguraldissertation zur Erlangung der Würde eines

DOCTOR RERUM OECONOMICARUM

der Wirtschafts- und Sozialwissenschaftlichen Fakultät

der Universität Bern

22. Dezember 2023

Die Fakultät hat diese Arbeit am 22.02.2024 auf Antrag der beiden Gutachter Prof. Dr. Igor Letina und Prof. Dr. Alexander Rasch als Dissertation angenommen, ohne damit zu den darin ausgesprochenen Auffassungen Stellung nehmen zu wollen.

Acknowledgements

First, I want to thank my supervisor, Igor Letina. His support and guidance were immeasurable for my dissertation and my personal academic development. Working together on the first chapter was a great experience, where I learned many valuable things about research and the academic process during that time. Further, whenever I struggled with my other chapters, he always had the right suggestions to advance them. Lastly, because of Igor, I applied for a grant to finance my research stay at DICE in November 2022, which greatly impacted my academic career. I would also like to thank Alexander Rasch for enabling this research stay and generously accepting to evaluate my thesis as an external reviewer.

A big thanks to Christian Zihlmann for first battling Gerzensee together with me and then introducing me to the world of behavioral economics, which eventually shaped my main research interests significantly.

I am grateful to Jean-Michel Benkert, who was always open to discussions and provided plenty of comments, constructive criticism, and advice. I also want to thank Marc Möller and Ralph Winkler for helping me to achieve my PhD at the University of Bern. I am thankful to my co-author of the first chapter Armin Schmutzler, who also contributed to my learning during the early stage of my doctoral studies.

Further thanks to Sarah Spycher and Severin Lenhard for guiding me through the life of a PhD student and sharing many great moments at conferences.

Of course, a great thanks to my office mates Moritz, Anna, Leyla, Severin, and Marina for making my time in Bern so enjoyable.

Finally, I am grateful for the support from my family and Michelle. I never had to worry about anything and was able to focus fully on my research and completing my doctoral studies.

Contents

List of Figures

Introduction

This thesis comprises three chapters providing new theoretical frameworks and insights for regulators, policy-makers, and competition authorities. Each chapter studies a particular problem in the field of industrial organization, aiming to inform decisions for competition policy. The first analyzes the effects of horizontal cooperation on innovation when financial resources are scarce. The second and third chapters study markets in which firms sell add-ons and complementary products, respectively, when consumers are subject to context effects.

In the first chapter, together with Igor Letina and Armin Schmutzler, we investigate collaboration on research and development (R&D) between competitors that face financial constraints. Innovation often involves immense R&D investments such that even large companies struggle to finance it internally. While external financing is possible, it is more costly. Cooperation in the form of a research joint venture (RJV) can be cheaper because the members combine their (financial) resources. Particularly in the automobile industry, we observe an increasing number of RJVs for the development of electric vehicles. When forming an RJV, firms share the costs and benefits of innovation and coordinate their research effort, but stay competitors in the product market. This cooperation alleviates the financing problem, especially because coordination eliminates wasteful duplication of effort.

Typically, collusion among competitors is prohibited as it harms consumers. With innovation, the matter is different. Resources can be used more efficiently, promoting innovation and eventually benefiting consumers.

However, research incentives could also be lower with an RJV. If a firm innovates alone, it gets a technological, and thus competitive, advantage. In this case, it would escape competition and increase profits. In an RJV, this is not possible anymore.

We provide a novel reason when RJVs are beneficial and increase innovation probability, and thus, competition authorities should allow it. This depends on how intense the competition in the product market is and how costly external financing is. When competition is soft, the benefit of escaping competition it is relatively small. In that case, an RJV always increases innovation probability. If competition is intense, the benefit of inventing alone is large. Hence, the cost-saving aspects of an RJV must be sufficiently strong that it increases innovation probability, which is the case when external financing is expensive. Further, RJVs can be a better alternative to mergers. While efficiency gains, and thus the effect on innovation, are similar, the negative impact on competition is smaller because an RJV does not reduce the number of firms in a market.

In the second chapter, together with Christian Zihlmann, we study markets with add-on selling and consumers with context-dependent preferences, coming from relative thinking or salience effects. Firms initially sell a base good and then offer an additional product. Behavioral consumers relate the extra offer to the price of the base good, which affects their price sensitivity and temporarily increases demand for the add-on. For instance, a seat upgrade for $50 feels less expensive when the flight costs $1000 than if the same flight costs only $300.

We observe the decoupling of products into a base good and an add-on in many markets. This practice leads to a sequential presentation of prices called drip pricing. Experimental and empirical evidence document that consumers purchase more add-ons when shown sequentially than when base goods and add-ons are bundled products. Firms may exploit behavioral consumers by increasing the price of the add-on. Importantly, the behavioral effect is stronger the larger the reference point is, which, in our case, is the base good price. Thus, firms also have an incentive to increase the price of the main goods. However, this lowers the demand for this primary good, and, in turn, the demand for the add-on. It is well known in the add-on literature that firms attract consumers with a low price for the first good and then generate profits with expensive add-ons. In our model, firms face a trade-off between exploiting more behavioral consumers or selling the add-on at a higher price. We investigate how this affects classical consumers in the market, who are not subject to behavioral effects.

We find that the presence of behavioral consumers has mixed, non-monotonic effects on these classical buyers. Depending on how large the behavioral population is, classical consumers can be better or worse off. When firms exploit the relative thinking of customers, products can become cheaper or more expensive. This novel result raises new questions for regulators about whether we need to protect classical consumers from the behavior of others. We show that an effective policy that prevents firms from exploiting can help all consumers and increase consumer surplus. However, an inefficient regulation can also increase prices and worsen the situation for consumers. Thus, policy-makers must be careful when implementing new policies and regulations, as it could cause more harm than good.

In the third chapter, I consider consumers with context-dependent preferences who purchase complementary products. In contrast to the second chapter, the two products are offered by different firms. For instance, consumers must buy a flight and book a hotel room. A surprisingly good deal on one product may catch the attention of consumers, who then are less price-sensitive for the other complement. Or, when complements are purchased sequentially, the price of the first may impact how the price of the second complement is perceived.

In the model, two firms compete and offer one complement, while a monopolist produces the other complement. Because each firm maximizes its profits, we have the inefficiency problem of double marginalization. This can be solved by vertical integration, which usually lowers prices for the total product. In our case, one of the competitive firms merges with the monopolist. In the benchmark case without behavioral effects on consumers, vertical integration increases profits and, thus, occurs when formation costs are not too high.

I analyze how behavioral consumers in the market affect the incentives for vertical integration. More specifically, consumers underweight the price of the competitive good. This makes competition less intense, which decreases the efficiency gains of vertical integration for the merged firms. This results in lower merger incentives, and vertical integration is less likely to happen in markets with this type of consumer.

Chapter 1

Research Joint Ventures: The Role of Financial Constraints

joint with Igor Letina and Armin Schmutzler

Published as:
P. Brunner, I. Letina and A. Schmutzler, Research joint ventures: The role of financial constraints. *European Economic Review* (2024).
https://doi.org/10.1016/j.euroecorev.2024.104742

1.1 Introduction

Innovation often involves large R&D investments. A well-known example is the pharmaceutical industry where blockbuster drugs can require high upfront R&D expenses.[1] Similarly, automobile producers have recently spent £341 billion within five years to become successful players in the electric vehicle industry.[2] The necessary investments and the required technological skills are so large that even industry giants rarely attempt to take on the task on their own. In the last few years, major players have agreed on research joint ventures (RJVs). For instance, *Daimler* and *Geely* jointly develop battery-driven *Smart* cars. *PSA* and *Opel* hooked up with *Saft*, a subsidiary of *Total*, to develop batteries. Together with *BP*, *Daimler* and *BMW* develop charging stations. *Renault*, *Nissan* and *Mitsubishi Motors* agreed on investing $26 billion to develop common platforms for electric vehicles. Further up in the value chain, suppliers of essential inputs have also joined forces.[3] Not only are the required R&D investments large, there is also significant uncertainty about which technology the vehicles of the future will rely on. Today, most electric vehicles are powered by lithium-ion batteries, but this technology has significant drawbacks and automotive companies are additionally investing in alternative approaches. For example, *Volvo* and *Daimler* are collaborating on fuel-cell driven cars, while *Ford* and *BMW* have jointly invested in a startup developing solid-state batteries. In all these partnerships, at least some of the firms are competing or planning to compete in the product market.[4]

Though competition policy investigates RJVs in various ways, it typically treats them more leniently than other forms of horizontal cooperation. For instance, the European Union addresses RJVs either under its merger

[1] For example, see “Largest video game industry acquisitions worldwide as of November 2022”, available at https://www.cbo.gov/publication/57126.

[2] Sources for all RJVs mentioned in this section are listed in Appendix 1.A.10.

[3] For instance, the German chemical firm *BASF* and the Chinese firm *Shanshan* jointly search for better materials to produce cathodes for batteries.

[4] While we will focus on such horizontal research joint ventures, purely vertical collaborations are common as well. For instance, *Panasonic* engages in a joint venture with *Toyota* to develop batteries. Moreover, there are joint ventures between *Volkswagen* and *Stellantis* with *Enel* and *ENGIE*, respectively, to develop networks of charging stations.

regulation or under Article 101 of the EU treaty, depending on whether it is a full-function joint venture or not. In the latter case, even if an RJV has been found to have anti-competitive object or actual or potential competition-restricting effects (Article 101(1)), it may still be justified on the basis of efficiency gains under certain conditions (Article 101(3)).[5] The legal situation in the United States is similar, with the *1993 National Cooperative Research and Production Act* specifying that horizontal cooperation in RJVs is not per se illegal, but is to be evaluated under the "rule of reason".

An important prerequisite to justify a friendly approach of competition policy towards RJVs is that they have beneficial effects on R&D activities. Existing literature focuses on knowledge spillovers as the main justification for RJVs.[6] Our paper analyzes a different channel through which RJVs can lead to more innovation: When R&D costs are high (so that firms are financially constrained) and there is significant uncertainty about the right way to generate the desired innovative outcome, an RJV can help reduce investments in duplicate R&D projects, thereby freeing up funds that can be invested in previously unexplored approaches. To clarify conditions under which this is indeed the case, we introduce a model that combines financial constraints and uncertainty about the right way to generate the desired innovation. Contrary to previous theoretical literature on RJVs, the firms not only choose how much to invest in, but also how to spread investments over different R&D projects. This feature of our model allows us to investigate how the members of an RJV can benefit from reallocating scarce resources across projects. Thereby, we can separate the decisions on how much to invest from the decision in which projects to invest in. To the best of our knowledge, we provide the first analysis of research joint ventures that explicitly considers project choice.

More precisely, in our benchmark model, we analyze a duopoly with two symmetric firms. These firms choose in which set of R&D projects from a continuum of alternatives to invest. Only one of all possible projects will

[5]See, in particular, Commission Regulation No. 1217/2010 of 14. December 2010.

[6]Early examples include Katz (1986), d'Aspremont and Jacquemin (1988), Kamien, Muller, and Zang (1992). See Section 1.5 for a detailed literature discussion.

lead to an innovation, resulting in a positive effect on the firm's product market profits. Therefore, when firms invest in a wider range of projects, they are more likely to find the right approach. We assume that projects are identical except that some are more costly than others. Each firm has a fixed budget, which can be used for R&D investments.[7] In addition, firms can borrow externally. In line with the empirical literature (see Section 1.5), we assume that such external financing is costly and that firms who borrow externally have to pay a positive interest rate on the external loans. The firm chooses its investment strategy so as to maximize expected profits. We assume that the budget is sufficiently small that, in equilibrium, both firms borrow positive amounts from the financial markets. Our analysis compares the outcome of this R&D competition game with the alternative that the firms form an RJV in an otherwise identical setting. In the latter case, the two firms combine their budgets, and the RJV chooses R&D investments to maximize joint payoffs. Firms share the research costs equally and, if successful, both receive the innovation. After the R&D outcomes materialize, the firms compete in the product market.

Our central results give conditions under which an RJV increases the probability of innovation. The intensity of product market competition is a first important determinant. To see this, note that, in the absence of an RJV, an innovating firm may benefit from *escaping competition*, moving ahead by being the only one who has access to a superior technology. Under an RJV, it is obviously impossible to escape competition by innovation, because firms have agreed to share the fruits of their research efforts. Instead, a successful RJV symmetrically increases the profits of both firms. When competition is soft, so that the increase in industry profits from successful joint innovation is large relative to the benefit from escaping competition, the innovation probability is higher under an RJV than under R&D competition. Interestingly, this result does *not* rely on the existence of financial constraints. Moreover, like all our main results, it does not require spillovers, which are

[7]We can also interpret the limited budget as the firm's (internally) available time of researchers or the laboratory's infrastructural capacity, which can be expanded through (more expensive) external researchers or laboratories.

the driving force behind innovation-enhancing research cooperation in the literature. As an example, we show that the soft competition case applies to a model of price competition with sufficiently differentiated goods.

Next, we suppose that competition is not soft, including for instance homogeneous quantity competition as well as price competition with weakly differentiated products. In this case, the value of escaping competition would always be higher than the joint profit increase from innovating together, so that, in the absence of budget constraints, the probability of innovation under an RJV would be lower than under R&D competition. This is precisely where our modeling choices play a critical role, because they allow us to identify features of RJVs that are absent in standard models. In an RJV, the participating parties can not only coordinate the decisions on how much to invest, but also in which projects to invest. This allows them to reduce duplication and free up resources, which they can spend on further projects without having to access the capital market. When the amount of internal funding that an RJV frees up is large enough, then that RJV can *potentially* invest in a wider range of projects, compared to independent firms, using just internal funding. Whether an RJV *actually* makes use of this opportunity or whether it just enjoys the cost savings from avoided duplication depends not only on the nature of competition, but also on financial constraints: When external financing conditions are sufficiently bad, then the RJV increases the innovation probability even when competition is not soft. To repeat, this result relies on the existence of financial constraints: Without them, the RJV would invest in less projects than the two independent firms together.

In the situation with relatively intense competition just described, the RJV not only increases the probability of a successful innovation, but at the same time it also reduces overall R&D spending. This result means that total industry R&D costs and the probability of a successful innovation do not necessarily move in the same direction. This is in stark contrast with the existing literature, which typically views an RJV as innovation-enhancing if and only if it increases total investment cost. This feature of our model underlines the importance of allowing for different R&D projects.

While understanding how RJVs impact innovation outcomes is of independent interest, maximizing consumer surplus is often emphasized as a policy goal. Under very mild assumptions, we show that any RJV that increases the probability of innovation also increases expected consumer surplus. This occurs because consumers are better off if innovation is more likely, and conditional on being discovered, if it is used by as many firms as possible. Since all RJVs increase the diffusion of innovation among firms (because all firms in an RJV get access to the innovation), then RJVs which increase the innovation probability unambiguously benefit consumers. Thus, the conditions that we identify for which RJVs increase the innovation probability are also sufficient to guarantee that RJVs increase consumer welfare.

Overall, the results just discussed show that RJVs are helpful for inducing innovation and improving consumer welfare under a wider range of circumstances than identified by previous literature. However, in line with existing worries in EU circles, we also found circumstances under which RJVs are harmful to innovation.[8] Thus, to evaluate the innovation effects of RJVs, it is decisive to understand the incentives of firms to form an RJV. If firms only had an incentive to form RJVs that reduce innovation, then lenient policy towards them would be misguided. We thus ask: Will firms have incentives to engage in RJVs for which our analysis has shown that they enhance innovation? Or will they rather engage in RJVs that reduce innovation? We find general and widely applicable conditions under which firms benefit from forming RJVs that increase the innovation probability. In particular, this will always be true unless competition is very intense. However, we also find circumstances under which firms engage in RJVs even though they reduce overall innovation – the cost savings in these cases suffice to make the RJVs profitable.

Next, we compare RJVs and mergers. Which of the two forms of cooperation is more conducive to innovation depends on the nature of product market cooperation and the stringency of financial constraints. This result relates to a recent discussion in merger control that has emphasized

[8]See "Guidelines on the applicability of Article 101 of the Treaty on the Functioning of the European Union to horizontal co-operation agreements." Official Journal of the European Union (2011/C 11/01).

R&D effects, asking whether (potentially) beneficial effects of mergers on innovation provide a justification for waving them through in spite of their well-known mark-up increasing effects. We identify a wide range of parameters, for which even mergers that lead to a higher innovation probability than R&D competition should be prohibited, as an RJV would have the same social benefits without the social costs of eliminating competition.[9]

Moreover, we explore the link between our analysis and the more familiar rationale for RJVs that relies on knowledge spillovers. In an extension of our model, we find that knowledge spillovers and financial constraints are complements in the sense that RJVs with financial constraints are more likely to increase the probability of innovation the stronger spillovers are, and vice versa. Finally, we analyze the relation between licensing and RJVs. In line with previous literature, the chance to earn licensing fees increases innovation incentives under R&D competition. As a result, the conditions under which an RJV yields a higher probability of innovation than R&D competition become more restrictive. Moreover, with licensing, if an RJV increases the probability of innovation, it *always* results in lower R&D spending.

All told, our paper attempts to shed light on how the consideration of project choice and financial constraints affects the analysis of research joint ventures. While we ignore important aspects such as firm asymmetries, costs of RJV formation and governance issues, and we work under the debatable assumption that the RJV does not induce collusive behaviour in the product market, we are confident that our approach can be a useful input for a more comprehensive welfare analysis.[10]

In Section 1.2, we provide the benchmark duopoly model. Section 1.3 analyzes the innovation effects of RJVs and identifies conditions under which they are profitable. In Section 1.4, we compare RJVs and mergers. Further, we extend the analysis to the case of spillovers and to multiple firms, and we

[9] More broadly, authors such as Farrell and Shapiro (2000) have emphasized that, even if efficiency gains outweigh the competition-softening effects of a merger, competition authorities still have to ask whether the merger is actually necessary to achieve these gains.

[10] See Duso, Röller, and Seldeslachts (2014) and Sovinsky (2022) for evidence suggesting that RJVs may foster collusion. However, note that our analysis of mergers for the duopoly case can alternatively be interpreted as an RJV with full collusion in the product market.

discuss licensing. In Section 1.5, we discuss the model in the light of existing literature. Section 1.6 concludes.

1.2 Model

Our model of R&D with project choice builds on previous work of Letina (2016) and Letina, Schmutzler, and Seibel (2021).[11] However, neither of these papers deals with research joint ventures or budget constraints. We assume that two ex-ante symmetric firms ($i \in \{1, 2\}$) can invest in R&D before they compete in the product market. There are two possible levels of technology – current technology, which is available to both firms, and new technology, which is only available to the firms that innovate. To improve their technology level, firms can invest in multiple projects θ from the set of available projects $\Theta = [0, 1)$. Only one these projects is correct, that is, leads to an innovation. Let $\hat{\theta} \in \Theta$ be the correct project. Nature chooses which of the available projects is correct, but firms are not informed about it, hence firms see the location of the correct project as a random variable. We assume that the location of the correct project is uniformly distributed on $\Theta = [0, 1)$. For each $\theta \in [0, 1)$, each firm chooses whether to invest in that research project ($r_i(\theta) = 1$) or not ($r_i(\theta) = 0$). If $r_i(\hat{\theta}) = 1$, then firm i will innovate for sure and if $r_i(\hat{\theta}) = 0$, then firm i will not innovate.[12] We restrict the firm's choices to the set of measurable functions $r : \Theta \to \{0, 1\}$, which we denote with $\mathcal{R}$. The cost of developing a project θ is given by $C(\theta)$, where we assume that the function $C : [0, 1) \to \mathbb{R}^+$ is differentiable and strictly increasing and that $C(0) = 0$ and $\lim_{\theta \to 1} C(\theta) = \infty$. Therefore, the total research costs of firm i are $\int_0^1 r_i(\theta)C(\theta)d\theta$.[13]

If a firm has chosen $r_i(\hat{\theta}) = 1$ in the investment stage, it has access to an innovation and enters the product market competition with technology state

[11] Accordingly, the model description follows those papers closely.

[12] It is possible to formulate a version of the model where firms can partially invest in research projects, that is $r_i(\theta) \in [0, 1]$. One benefit of such richer model is that it admits a symmetric equilibrium. However, all economic insights remain the same as in the current version. For this reason, we decided to present the simpler model. The interested reader can find the model with intermediate investment levels in the previous version of this paper, Brunner, Letina, and Schmutzler (2022).

[13] If this integral does not converge, we assign the value ∞ to it.

$t_i = I$. If it has not invested in $\hat{\theta}$, it does not have access to an innovation, and its technology state is $t_i = 0$.

For now, we do not explicitly model product market competition. Instead, we formulate weak general assumptions that we show to hold in familiar models of product market competition in Section 1.3.6. We assume that the product market profits of firm i are given in reduced form by the expression $\pi_{t_i t_j}$ for $j \neq i$. If both firms innovate, then they will compete with the new technology, and their market profits are given by π_{II}. Similarly, if both firms compete with the current technology, then each of them obtains profits π_{00}. If a single firm innovates, it obtains profits π_{I0}, while the other firm obtains π_{0I}. We will impose the following regularity assumptions on the profit functions.

Assumption 1 (Regularity of profit functions)**.**

(i) Profits are non-negative: $\pi_{t_i t_j} \geq 0$ *for all* t_i *and* t_j*.*

(ii) Innovation increases profits: $\pi_{II} \geq \pi_{00}$*.*

(iii) Competitor innovation reduces profits: $\pi_{t_i 0} \geq \pi_{t_i I}$ *for* $t_i \in \{0, I\}$*.*

(iv) Escaping competition is more valuable than catching up: $\pi_{I0} - \pi_{00} \geq \pi_{II} - \pi_{0I}$*.*

Obviously, Assumptions 1(i)-(iii) are compatible with most standard oligopoly models. Furthermore, authors such as Bagwell and Staiger (1994), Leahy and Neary (1997), Farrell and Shapiro (2000) and Schmutzler (2013) have argued that submodularity conditions like (iv) hold for many innovation games with standard models of price and quantity competition unless knowledge spillovers are strong. Intuitively, a successful innovation of the competitor reduces own equilibrium outputs and margins, which reduces the benefits from increasing margins and outputs through own innovation.

While we will always maintain that competition is sufficiently intense that Assumption 1(iv) holds, we will distinguish between three different regimes according to the intensity of competition.[14]

[14] Boone (2008a,b) similarly uses the relation between efficiency differences and profit differences in his definition of intensity of competition.

Definition 1 (Intensity of competition).

(i) Competition is ***intense*** *if avoiding the competitor catching up is more valuable than catching up:* $\pi_{I0} - \pi_{II} > \pi_{II} - \pi_{0I}$.

(ii) Competition is ***soft*** *if avoiding the competitor catching up is less valuable than improving together:* $\pi_{I0} - \pi_{II} < \pi_{II} - \pi_{00}$.

(iii) Competition is ***moderate*** *if neither of the above cases holds, so that:* $\pi_{II} - \pi_{0I} \geq \pi_{I0} - \pi_{II} \geq \pi_{II} - \pi_{00}$.

For cost-reducing investments, competition is typically intense in a homogeneous

Bertrand market, but also for a homogeneous Cournot market with linear demand (see Section 1.3.6). In Section 1.3.6, we will see that all three regimes arise with differentiated price competition, depending on the degree of substitution.

Each firm has a research budget B. If a firm spends more than B, it has to borrow from the capital market at the interest rate $\rho > 0$, reflecting the well-known difficulties of external financing of R&D investments (see Section 1.5).[15] We will assume (in a way which will be made precise in Assumption 2) that without a research joint venture the budget is binding and both firms find it optimal to borrow positive amounts from the capital market.

The expected total payoff of firm i, given the strategy of competitor j is then

$$\begin{aligned}\mathbb{E}\Pi_i(r_i, r_j) = &\int_0^1 (1 - r_j(\theta)) \left[r_i(\theta)\pi_{I0} + (1 - r_i(\theta))\pi_{00}\right] d\theta \\ &+ \int_0^1 r_j(\theta) \left[r_i(\theta)\pi_{II} + (1 - r_i(\theta))\pi_{0I}\right] d\theta \\ &- \int_0^1 r_i(\theta)C(\theta)d\theta - \rho \max\left\{0, \int_0^1 r_i(\theta)C(\theta)d\theta - B\right\}.\end{aligned}$$

The first integral captures the expected payoffs when firm j does not innovate. Similarly, the second integral represents the payoffs when firm j innovates.

[15] Although our financing assumption is simple, it captures the essence of the idea that the marginal costs of own funds, as long as they are available, are lower than the marginal costs of borrowed funds.

The third line represents research costs, depending on whether the firm borrows from the capital market or not. Firms choose $r_i(\theta)$ and $r_j(\theta)$ simultaneously with the goal of maximizing $\mathbb{E}\Pi_i$ and $\mathbb{E}\Pi_j$, respectively. We will focus on pure strategy equilibria throughout.

1.3 Effects of Research Joint Ventures

1.3.1 Equilibrium under R&D Competition

We now characterize the equilibrium strategies under R&D competition. Given our assumptions on research costs, it is intuitive that both firms will invest in projects near $\theta = 0$, whereas neither firm will invest in projects near $\theta = 1$. One would thus expect equilibrium strategies to be of the following type.

Definition 2. *A* ***double cut-off strategy profile*** *is a profile (r_i, r_j) of research strategies for which $\theta_L \in [0, 1)$ and $\theta_H \in [\theta_L, 1)$ exist such that*

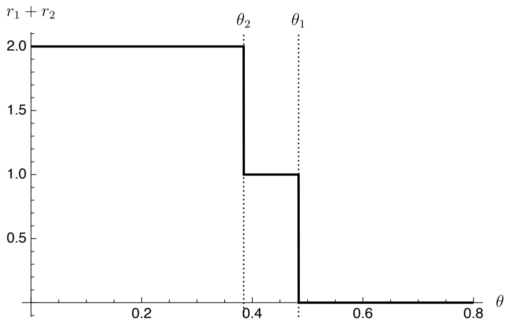
$$r_i(\theta) = r_j(\theta) = 1 \quad \text{if } \theta < \theta_L$$
$$r_i(\theta) = r_j(\theta) = 0 \quad \text{if } \theta > \theta_H.$$

Note that the definition does not specify which firm invests for $\theta \in (\theta_L, \theta_H)$. To find the equilibrium cut-off values, consider the equations

$$(1 + \rho)C(\theta_1) = \pi_{I0} - \pi_{00}$$
$$(1 + \rho)C(\theta_2) = \pi_{II} - \pi_{0I}.$$

θ_1 is the most expensive project in which a firm can profitably invest using external finance, assuming that the competitor does not invest in this project. Similarly, θ_2 is the most expensive project in which a firm can profitably invest using external finance, assuming that the competitor invests in this project. An immediate consequence of Assumption 1(iv) is that $\theta_2 \leq \theta_1$. The following assumption guarantees that both firms will borrow positive amounts in any equilibrium.

Figure 1.1: Equilibrium Portfolio



Industry portfolio of research projects in any equilibrium.

Assumption 2. $B < \int_0^{\theta_2} C(\theta)d\theta$.

Next, we characterize all equilibria of this game.[16]

Lemma 1 (Characterization of investment strategies under competition).
(i) The research competition game has multiple equilibria. A profile of double-cut off strategies (r_i^*, r_j^*) *is an equilibrium if it satisfies (a)* $\theta_L = \theta_2$ *and* $\theta_H = \theta_1$ *and (b) for each* $\theta \in (\theta_2, \theta_1)$ *either:*

$$r_i^*(\theta) = 1 \text{ and } r_j^*(\theta) = 0 \text{ or}$$
$$r_i^*(\theta) = 0 \text{ and } r_j^*(\theta) = 1.$$

(ii) No other pure-strategy equilibria of the research-competition game exist.

Thus, all equilibria share the double cut-off structure, which is determined by the marginal cost of research projects and the benefit of being successful. Both firms invest in the cheap projects $\theta \in [0, \theta_2)$, while neither firm invests in the expensive projects $\theta \in (\theta_1, 1]$. For intermediate projects, the marginal benefits of an innovation are higher than the marginal costs when only one firm finds the innovation, but not when both firms are successful. Hence, for each θ in the interval (θ_2, θ_1), one firm invests while the other does not invest,

[16]Of course, for any equilibrium strategies r_i^* and r_j^* there exist infinitely many equilibria which only differ on sets of measure zero. We ignore those differences and only regard strategies as distinct if they differ on sets of positive measure.

but the identity of the investing firm is not determined, which leads to the multiplicity of equilibria. However, all equilibria are equivalent – in the sense that they generate the innovation with the same probability and lead to the same market structure in each state of the world. In any equilibrium, the overall innovation probability is θ_1. Furthermore, the probability of a duopoly with an innovation is θ_2, the probability of a single firm with an innovation is $\theta_1 - \theta_2$, while the probability of a duopoly without an innovation is $1 - \theta_1$. Note that there is duplication of research efforts in equilibrium, as all projects in the interval $[0, \theta_2)$ are duplicated. Figure 1.1 depicts the industry portfolio of research projects in every equilibrium.

The symmetric setting of our model brings out clearly that the asymmetric outcome depicted in the figure exclusively reflects equilibrium considerations rather than exogeneous differences between firms. The value of investing in a particular project depends on the behavior of the competitor. Investing tends to be more worthwhile if the competitor does not invest than if he invests. In the former case, the resulting profit increase is given by the value of escaping competition $(\pi_{I0} - \pi_{00})$, which by Assumption 1 is larger than the value of catching up $(\pi_{II} - \pi_{0I})$, which determines the incentives for investing in projects that the competitor also invests in. The asymmetric investment behavior of firms for intermediate projects directly reflects these differences in incentives.

As the difference between the value of escaping competition and the value of catching up increases (reflecting greater intensity of competition), the area with asymmetric investment becomes larger. An increase in the value of escaping competition increases θ_1 and thus project variety and the probability of innovation. This tends to lead to more demand for external funding to finance more expensive projects. Conversely, in most standard oligopoly models, the value of catching up decreases with more intense competition, which lowers the amount of duplication and, thus, the need for external funds. Therefore, the overall effect of increased competition on external funding is ambiguous. Further, a higher borrowing cost ρ implies a lower probability of finding the innovation and less duplication of effort since θ_1 and θ_2 both decrease in ρ. Lastly, by Assumption 2, a marginal change in the budget size

B does not affect the equilibrium portfolio.

1.3.2 Optimal Project Choice of an RJV

In our model of RJVs, the firms combine their individual budgets and invest in research together. However, the two firms still compete in the product market after the successful project has been realized.[17] Moreover, the research costs are equally shared and both firms obtain the innovation if developed. This eliminates the possibility of an asymmetric product market structure. The firms will compete either with or without innovation. Like an individual firm, the RJV can borrow at the interest rate ρ on the external market if the total budget $2B$ is insufficient. The RJV chooses the research strategy r_v to maximize the expected total payoff

$$\begin{aligned}\mathbb{E}\Pi_v(r_v) = 2\int_0^1 &[r_v(\theta)\pi_{II} + (1-r_v(\theta))\pi_{OO}]\,d\theta \\ &- \int_0^1 r_v(\theta)C(\theta)d\theta - \rho\max\left\{\int_0^1 r_v(\theta)C(\theta)d\theta - 2B, 0\right\}. \quad (1.1)\end{aligned}$$

The optimal strategy will be of the following type.

Definition 3. *A* ***single cut-off strategy*** *is a research strategy r_v for which a $\theta^* \in [0,1)$ exists such that $r_v(\theta) = 1$ if $\theta < \theta^*$ and $r_v(\theta) = 0$ if $\theta > \theta^*$.*

Let θ^B be defined as the solution to $\int_0^{\theta^B} C(\theta)d\theta = 2B$ if $\int_0^1 C(\theta)d\theta > 2B$ and $\theta^B = 1$ otherwise. That is, a joint venture which invests in all projects in the set $(0, \theta^B)$ either has innovation costs equal to $2B$ or invests in all projects. Next, let θ^u and θ^ρ be the solutions to the following equations

$$\begin{aligned}(1+\rho)C(\theta^\rho) &= 2[\pi_{II} - \pi_{OO}] \\ C(\theta^u) &= 2[\pi_{II} - \pi_{OO}].\end{aligned}$$

Thus, θ^u is the most expensive research project in which an RJV that does not borrow from the capital market wants to invest in. Similarly, $\theta^\rho < \theta^u$

[17] This is the main difference to a merger, which will result in a monopolistic market in any case.

is the most expensive research project in which an RJV that has to borrow would choose to invest in. How θ^B relates to these two values will determine the optimal portfolio of the RJV.

Lemma 2 (Investment strategies of an RJV)**.**
The RJV chooses a single cut-off strategy with

$$\theta^* = \begin{cases} \theta^\rho & \textit{if} \quad \theta^B < \theta^\rho \\ \theta^B & \textit{if} \quad \theta^B \in [\theta^\rho, \theta^u] \\ \theta^u & \textit{if} \quad \theta^B > \theta^u. \end{cases}$$

Thus, the cut-off project always lies in the interval $[\theta^\rho, \theta^u]$. Which of the three cases in the lemma arises depends on the budget B, the interest rate ρ, on product market profits and on the cost function. If $\theta^B < \theta^\rho$, then the joint venture invests its entire budget $2B$ into research and, in addition, it borrows from the capital market in order to finance its research activities. In contrast with a marginal change in the cost of borrowing ρ, a marginal increase in the budget would not affect the investment strategy. When $\theta^B \in (\theta^\rho, \theta^u)$, the RJV invests the entire budget, but it does not borrow. Thus, a marginal increase in the budget would lead to an increase in investment, whereas a marginal change in ρ would have no effect. Finally, when $\theta^B > \theta^u$, the RJV does not borrow and furthermore only invests a portion of its budget into research. Hence, neither marginal changes in B nor in ρ would change investment behavior, which is fully determined by product market conditions.

Note that in standard oligopoly models, the expression $\pi_{II} - \pi_{00}$ is decreasing in standard parameterizations of the intensity of competition.[18] This implies that the critical cut-off projects θ^ρ and θ^u are larger when product market competition is softer. Thus, unless it is optimal to just invest the entire budget ($\theta^* = \theta^B$), the RJV uses more funding when competition becomes softer. This is in line with the findings of Kamien et al. (1992) and Goyal and Moraga-Gonzalez (2001) that softer product market competition increases incentives to cooperate and leads to higher research efforts.

[18]For instance, this is the case in our two examples with linear Cournot competition and differentiated price competition in Section 1.3.6.

1.3.3 R&D Competition vs. R&D Cooperation

Next, we present our central result that deals with the effect of the RJV on the probability that an innovation will be discovered. Define the interest threshold $\bar{\rho}$ and the budget threshold $\bar{B}(\rho)$ as

$$\bar{\rho} = \begin{cases} \dfrac{\pi_{I0} - \pi_{II} - (\pi_{II} - \pi_{00})}{2(\pi_{II} - \pi_{00})}, & \text{for } \pi_{II} > \pi_{00} \\ \infty, & \text{for } \pi_{II} = \pi_{00}. \end{cases}$$

$$\bar{B}(\rho) = \frac{\int_0^{\theta_1} C(\theta)d\theta}{2}$$

The budget threshold depends negatively on ρ because θ_1 does. The thresholds play a critical role for the effects of an RJV on innovation.

Proposition 1 (Comparison of R&D competition and RJV)**.**

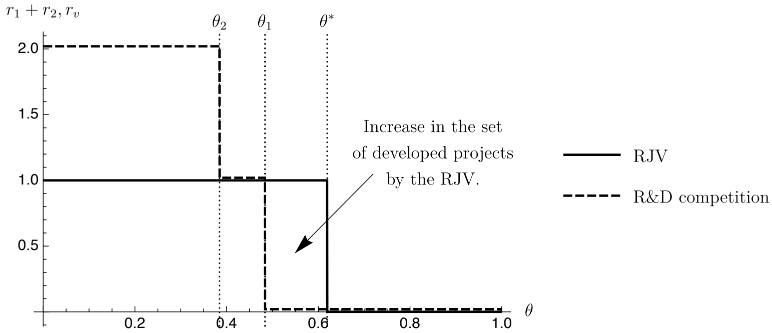
(i) Suppose competition is soft. Then the innovation probability is strictly larger under the RJV than under R&D competition.

(ii) Suppose competition is moderate or intense. Then:
(a) The innovation probability is strictly larger under the RJV than in any equilibrium under competition if and only if $B > \bar{B}(\rho)$ and $\rho > \bar{\rho}$.
(b) If the formation of the RJV strictly increases the innovation probability, then it weakly decreases total R&D spending.

The result reflects the subtle interplay between product market competition and financing conditions. In a model with R&D project choice, an RJV results in efficiency gains at the investment stage – it reduces the amount of duplication of research projects. This allows the RJV to "cast a wider net," as the funds that were previously used to finance duplicate research projects can now be redirected to other projects. This duplication reduction effect of the RJV makes it less costly to sustain high innovation probabilities. However, a potential countervailing effect needs to be taken into account: Escaping competition can be very valuable for each individual firm. Thus, compared with an RJV, incentives for innovation may be higher for a firm that can fully appropriate the benefits from innovation as the single successful innovator under R&D competition. If competition is soft, i.e., (i) holds, then this countervailing effect has no bite, as joint profits in an RJV are high enough that the innovation probability will be higher than under R&D competition. As we will see in Proposition 6, this result does not even require the existence of financial constraints.

By contrast, Proposition 1(ii) deals with the case that product market competition is moderate or intense. Then additional requirements are necessary for an RJV to increase innovation. Together, the condition that $\rho > \bar{\rho}$ and $B > \bar{B}(\rho)$ guarantee that the RJV will invest in more projects than both firms would in *any* equilibrium without the RJV, even though product market competition is not soft.[19] The advantages of the RJV in this setting come from the ability to avoid duplication and thereby finance a wider range of projects internally, thus avoiding the necessity to borrow from the capital markets. This is illustrated in Figure 1.2. When either $B \leq \bar{B}(\rho)$ or $\rho \leq \bar{\rho}$, so that the conditions in (iia) are not satisfied, then RJVs (weakly) decrease the innovation probability.

Figure 1.2: Innovation Effect of RJV



An example of an RJV increasing the set of developed projects.

Result (iib) deserves particular emphasis. It is common in the innovation literature to use the overall amount of R&D spending as a measure of

[19] Note that there is a tension between Assumption 2 which demands that the budget is not too high and the condition in Proposition 1(ii) that $B > \bar{B}(\rho)$. The Cournot example in Section 1.3.6 shows that the conditions can nevertheless be satisfied together for non-degenerate parameter regions.

the probability that an innovation will be discovered. Usually, a policy is said to promote innovation if it leads to more R&D spending. The result demonstrates that this approach can be misleading: When competition is not soft and financial constraints are severe, R&D competition leads to both higher R&D costs and a lower innovation probability than an RJV. Intuitively, in any equilibrium under R&D competition, both firms invest more than their available budget in R&D. Therefore, the marginal R&D project that they are willing to invest in has to be sufficiently profitable, so that incurring the higher marginal cost of borrowed funds is justified. However, whenever the conditions of Proposition 1(iia) are satisfied, an RJV optimally invests weakly less than its total budget. In spite of this reduction in R&D costs, the probability of innovation increases as the reduction in investment corresponds to avoided duplication rather than reductions in project variety. By contrast, in case (i), we cannot rule out that the RJV spends more on R&D than the firms under R&D competition: While the RJV can achieve the same innovation probability as the competitive firms with lower costs, it also faces stronger investment incentives.

1.3.4 Consumer Welfare

While supporting inventiveness of an industry can be a worthy goal in itself, competition policy often emphasizes consumer surplus. As our next result shows, the two are aligned under very mild conditions.

The model we have introduced so far does not specify the impact of innovation on consumers at all. However, a natural intuition is that consumers benefit from innovations. The following assumption formalizes this intuition. Denote consumer surplus resulting from market competition when both firms have innovated with CS_{II}, when neither firm has innovated with CS_{00} and with CS_{I0} when only one firm has innovated.

Assumption 3. *Consumers benefit from innovation:* $CS_{II} > CS_{00}$ *and* $CS_{II} > CS_{I0}$.

When innovations are aimed at developing better products or lowering production costs, we can expect that some of the benefits will be passed on to consumers, so that Assumption 3 will hold. With the addition of this assumption, we can show the following result.

Proposition 2 (Effect of RJVs on consumer surplus).
If an RJV strictly increases the innovation probability, then it also strictly increases expected consumer surplus.

Formation of an RJV affects consumer surplus in two ways: (i) it changes the probability that the innovation is discovered and (ii) it changes the diffusion of innovation among the competing firms. By Assumption 3, consumers benefit from both a higher innovation probability and more diffusion of innovation. Since RJVs always facilitate the diffusion of innovation (because whenever the RJV innovates, both firms can use the resulting innovation), then an RJV that increases the innovation probability clearly leads to higher consumer surplus.

Together with Proposition 1, this result gives simple conditions for an RJV to increase expected consumer surplus. It should be noted that these conditions are sufficient, but not necessary – because RJVs always increase the diffusion of innovation, it is possible for an RJV that slightly decreases the innovation probability to lead to a higher expected consumer surplus.

1.3.5 Profitability of RJVs

So far, we have analyzed how the formation of an RJV affects innovation probability and consumer welfare. We have not yet asked whether it is in the firms' interest to agree on an RJV. In the following, we will deal with this issue. We ask under which conditions joint profits are higher under an RJV than under R&D competition. If this requirement is not fulfilled, then at least one of the firms would not consent to an RJV. By contrast, if the RJV does increase joint profits and profits are symmetric under R&D competition, then the RJV will result in a Pareto improvement from the perspective of the two firms.[20] Even when profits are not symmetric ex ante, an RJV that

[20] A sufficient condition for equal profits under R&D competition to emerge is that the firms coordinate on an equilibrium where they innovate with an equal probability and where their research costs are equal.

increases joint profits could always be turned into a Pareto improvement using suitable transfers.

Then using Lemmas 1 and 2, we find that net profits with an RJV are at least as high as under competition if and only if

$$\begin{gathered}2\theta^{*}\pi_{II}+2\left(1-\theta^{*}\right)\pi_{00}-\gamma^{rjv}\geq \\ 2\theta_2\pi_{II}+\left(\theta_1-\theta_2\right)\left(\pi_{I0}+\pi_{0I}\right)+2\left(1-\theta_1\right)\pi_{00}-2\gamma^{com},\end{gathered} \tag{1.2}$$

where γ^{rjv} and γ^{com} capture total research cost (including the costs of external financing) incurred by the RJV and a single firm under competition, respectively.[21] In the following, we will shed more light on this condition by identifying transparent (sufficient) conditions on primitives under which it holds. Define

$$\Psi=\begin{cases}\dfrac{\pi_{I0}+\pi_{0I}-2\pi_{II}}{2(\pi_{II}-\pi_{00})} & \text{for } \pi_{II}>\pi_{00}\\ \infty, & \text{for } \pi_{II}=\pi_{00},\end{cases}$$

and note that whenever competition is intense, $\Psi>0$.

Proposition 3 (Profitable innovation-enhancing RJV)**.**
An RJV strictly increases net profits in each of the following constellations.

(i) Competition is soft.

(ii) Competition is moderate, $B>\bar{B}(\rho)$ and $\rho>\bar{\rho}$.

(iii) Competition is intense and $\frac{\min\{\theta^B,\theta^u\}-\theta_1}{\theta_1-\theta_2}>\Psi$.

In all three cases, an RJV strictly increases the innovation probability.

The distinction between the three cases reiterates the importance of the intensity of competition. In case (i), competition is soft, so that part (i) of Proposition 1 applies – the RJV increases innovation. Proposition 3 shows that, in this case, the firms' incentives for RJV formation are fully aligned

[21] Using Lemma 2, θ^* can be expressed in terms of (θ^B, θ^u and θ^ρ), which, in turn, can be expressed in terms of fundamentals.

with the goal of increasing the innovation probability. Like Proposition 1(ii), Proposition 3(ii) imposes that competition is moderate. In this region, an RJV increases profits as well as the innovation probability, provided the additional conditions on the budget and the interest rate hold. Finally, Part (iii) applies when competition is intense. Contrary to soft and moderate competition, the conditions guaranteeing that an RJV increases innovation by Proposition 1 and the condition under which it increases profits no longer coincide: The additional condition in Proposition 3(iii) limits the intensity of competition as captured by Ψ.[22] For instance, it does *not* hold with homogeneous Bertrand competition. It also requires that the budget of the RJV is sufficiently large.[23]

In most cases, the conditions in Proposition 3 also guarantee that the RJV does not spend more than its total budget, so that, by Assumption 2, it does not increase total expenditures. An exception arises in the subcase of (i) where the budget is sufficiently low that $\theta^B < \theta^\rho$: In this case, the RJV may spend more ($\theta^* = \theta^\rho$) than the two firms would have spent under R&D competition. Spending the same amount as before would have reduced costs without affecting innovation and thus would have already been profitable. The fact that the RJV chooses to spend more thus means that this is profitable, despite the increase in R&D costs.

An immediate corollary of Propositions 2 and 3 is that an RJV increases total welfare whenever conditions (i)–(iii) of Proposition 3 are satisfied. The reason for this is that such an RJV increases the innovation probability, so that, by Proposition 2, it increases consumer welfare and by Proposition 3, it increases net profits, therefore increasing total welfare.

As mentioned in the introduction, one concern about research joint ventures in the European Union is their potential adverse effect on the probability of innovation. While we have just seen that firms typically want to engage in RJVs if they increase innovation, we cannot rule out the case that firms engage in research joint ventures even when they reduce the probability

[22] Note that Ψ is high if the value of avoiding competition is high relative to the value of catching up.

[23] At the boundary between the intense and moderate competition regime, $\Psi = 0$. Thus, the second condition in (iii) reduces to $\theta^B > \theta_1$ and $\theta^u > \theta_1$, which is equivalent to the conditions $B > \bar{B}(\rho)$ and $\rho > \overline{\rho}$ in (ii).

of innovation:

Proposition 4 (Profitable innovation-reducing RJV).
Suppose that the following conditions hold:

(i) $2\pi_{II} - (\pi_{I0} + \pi_{00}) = 0$.

(ii) $B \leq \bar{B}(\rho)$ *or* $\rho \leq \bar{\rho}$.

(iii) $\pi_{II} > \pi_{0I}$.

Then there exists some $\hat{\pi}_{I0} > \pi_{I0}$ *such that for all* $\pi'_{I0} \in (\pi_{I0}, \hat{\pi}_{I0})$ *and keeping other parameters fixed, the RJV is profitable, but reduces the innovation.*

The result applies when competition is moderate, but close to the area where it would be soft. Then, under Condition (ii) in Proposition 4, an RJV would reduce innovation slightly, but without major adverse effects on gross profits. The cost-reducing effect of an RJV will then suffice to make it profitable.

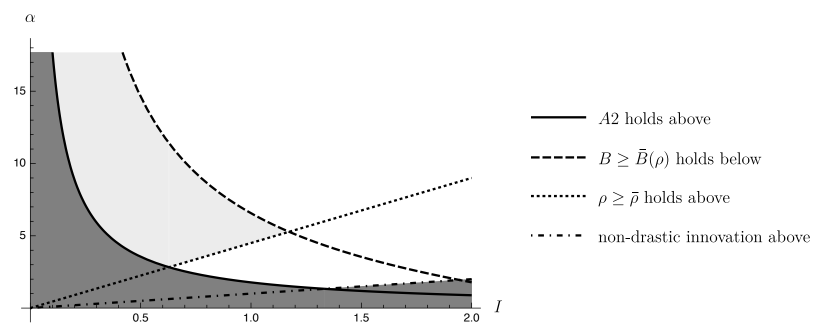
1.3.6 Examples

In this subsection, we illustrate the general analysis with two standard oligopoly models. For the first one, homogeneous linear Cournot competition, competition is moderate or intense, so that Proposition 1(ii) always applies and financial constraints are necessary for the innovation probability to be higher with an RJV than without. In the second example, differentiated price competition, competition can be soft. When this is the case, Proposition 1(i) applies and the RJV always increases the innovation probability. In each case, we only sketch the analysis; more details are in Appendix 1.A.6.

Cournot Competition

Suppose that two firms are choosing quantities q_1 and q_2, with $Q = q_1 + q_2$. Assuming an interior solution, the market price is given by $P(Q) = a - bQ$. Each firm can produce the good with some constant marginal cost c. The firms can invest in a potential process innovation that reduces the marginal cost of production to $c - I$ for some $I > 0$. Denoting $\alpha = a - c$, we assume for simplicity that $\alpha > I$, which guarantees that innovations are non-drastic. Calculating standard Cournot profits when firms have marginal costs c or $c-I$ yields the reduced form profits $\pi_{t_i t_j}$ and it is straightforward to verify that they satisfy Assumption 1. In fact, the stricter condition that competition is not soft, as required by Proposition 1(ii), holds for all parameter values.

Figure 1.3: Cournot Example



Comparison of R&D competition and RJV in a Cournot example with inverse demand $P(Q) = a - bQ$, constant marginal costs c, $B = 0.01$, $\rho = 0.1$ and $C(\theta) = \frac{\theta}{1-\theta^2}$. Axes depict cost reduction I and $\alpha = a - c$.

Thus, after calculating $\overline{\rho} = \frac{I}{2\alpha + I}$, we directly obtain:

Corollary 1. *In the linear Cournot model, the innovation probability is strictly higher with an RJV than with R&D competition if and only if* $\rho > \bar{\rho} = \frac{I}{2\alpha+I}$ *and* $B > \bar{B}(\rho)$. *If these conditions both hold, then the total R&D expenditures of the RJV are lower than those under R&D competition.*

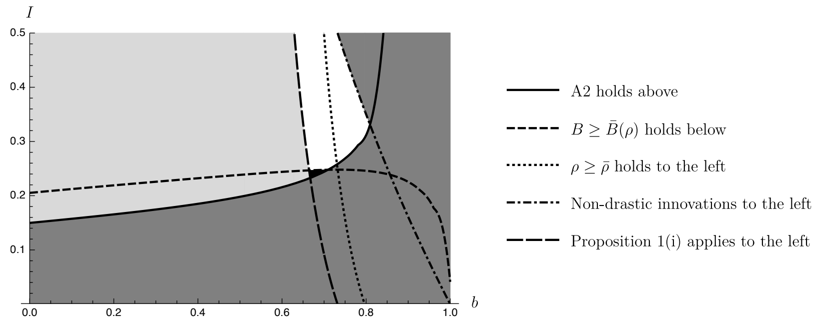
When competition is moderate or intense, the RJV only improves the innovation probability if the impact of pooling of resources is significant enough (as captured by the conditions on ρ and B). Importantly, Corollary 1 also identifies the role of the product market. A larger product market (captured by higher α) and a smaller innovation size I both increase the range of interest rates for which the RJV increases profits.

Figure 1.3 illustrates the result for specific parameter values. Assumption 2 and the focus on non-drastic innovations imply that we do not consider the darkly shaded region. The lightly shaded area depicts the parameter region for which the innovation probability is higher with an RJV than with R&D competition. The existence of this region means that the requirement of Assumption 2 that the budget is sufficiently small and the requirement from Proposition 1(ii) that it is sufficiently large are consistent. Note that all RJVs that increase the innovation probability compared to any equilibrium under R&D competition are profitable in this case.[24] In the parameter region colored in white, an RJV lowers the innovation probability compared to any equilibrium under competition.

Differentiated Price Competition

The linear homogeneous Cournot model is simple to analyze, but it restricts the possible outcomes, because competition is moderate or intense, so that Propositions 1(i) and 3(i) never apply. With differentiated goods, competition can be soft (as well as moderate or intense), so that these results become applicable. To see this, consider a standard model of differentiated price competition with inverse demand $p_i = 1 - q_i - bq_j$ for $b \in [0, 1)$ and constant marginal cost $c > 0$ where firms can engage in cost reductions $I \leq c$.

Figure 1.4: Bertrand Example



Comparison of R&D competition and RJV in a differentiated Bertrand example with inverse demands $p_i = 1 - q_i - bq_j$ and constant marginal costs $c = 0.5$. Axes depict substitution parameter b and cost reduction I.

[24]This is true because, for the given parameterization, competition is moderate. Hence, according to Proposition 1(ii), such innovation-enhancing RJVs must satisfy the conditions that are sufficient for a profitable RJV according to Proposition 3(ii).

In the appendix, we derive the equilibrium profits.

Figure 1.4 illustrates the stark contrast to the homogeneous Cournot example. We exclude parameter areas where the innovation is drastic (in which case π_{0I} would be negative) and/or Assumption 2 is violated (darkly shaded region). The central observation is that the RJV increases innovation and is profitable with sufficiently weak competition (in the large shaded grey area). This follows by applying Propositions 1(i) and 3(i). By contrast, the parameter region where Proposition 1(ii) and 3(ii) apply is very small (only the very small black area in the middle of the figure). Finally, note that, by Proposition 4, it will be profitable to engage in RJVs that reduce innovation (and costs) for parameter constellations near the left boundary of the white region.[25]

1.4 Further Results

In this section, we provide further results. We first compare the effects of RJVs with those of mergers. Then we allow for spillovers and licensing, respectively. Finally, we consider markets with more than two firms.

1.4.1 Mergers vs. RJVs

Competition policy usually views RJVs more favorably than full mergers as they allow the participants to reap some of the efficiency benefits that might arise in R&D, without necessarily eliminating product market competition between the firms involved.[26] However, a precise comparison needs to take differences in the effects of RJVs and mergers on innovation into account. In the following, we therefore analyze the innovation effects of a merger between the two firms, following the above analysis of the RJV closely. Contrary to the RJV, the merged entity not only combines the research budget, but its constituent parts give up competition entirely. We denote the (monopoly)

[25] Close to the left boundary of the white region, condition (i) of Proposition 4 holds. Moreover, $B < \bar{B}(\rho)$, so that (ii) holds.

[26] Note, however the empirical work suggesting that RJVs may foster collusion (Duso et al. (2014) and Sovinsky (2022)).

profit of the merged firm as π_{t_m}, where $t_m \in \{0, I\}$ indicates whether the firm has successfully innovated or not. In line with Assumption 1(ii), we assume that innovation increases profits.

Assumption 4. $\pi_I > \pi_0$.

The analysis for the merged firm is entirely analogous to the RJV case, except that we have to replace π_{II} with π_I and π_{00} with π_0 in the expected payoff formula (1.1).

Accordingly, we define critical values θ_m^u and $\theta_m^\rho < \theta_m^u$ which are analogous to θ^u and θ^ρ, except that we replace $2(\pi_{II} - \pi_{00})$ with $\pi_I - \pi_0$. It is straightforward that the merged firm optimally uses a single cut-off strategy like the RJV, with θ_m^u and θ_m^ρ instead of θ^u and θ^ρ (see Lemma 1.A.6 in Appendix 1.A.7). As a result, the comparison between investments with a merger and with R&D competition (see Proposition 1.A.1 in Appendix 1.A.7) is analogous to the comparison between the RJV and R&D competition (Proposition 1), except that we again need to replace $2(\pi_{II} - \pi_{00})$ with $\pi_I - \pi_0$, and the interest rate threshold thus becomes

$$\bar{\rho}_m = \frac{\pi_{I0} - \pi_{00} - (\pi_I - \pi_0)}{\pi_I - \pi_0}.$$

The following result compares the innovation probability under a merger and under an RJV.

Proposition 5 (Comparison of an RJV and a merger)**.**

(i) *If* $2(\pi_{II} - \pi_{00}) \geq \pi_I - \pi_0$, *the innovation probability under an RJV is weakly higher than under a merger. The difference is strict, except when* $\theta^B \in [\theta^\rho, \theta_m^u]$ *or* $2(\pi_{II} - \pi_{00}) = \pi_I - \pi_0$.

(ii) *If* $2(\pi_{II} - \pi_{00}) < \pi_I - \pi_0$, *the innovation probability under an RJV is weakly lower than under a merger. The difference is strict, except when* $\theta^B \in [\theta_m^\rho, \theta^u]$.

A merger leads to similar efficiency gains as an RJV – in both cases duplicate projects are eliminated, and those resources can be invested into new projects. However, the total profit increase for the members of the RJV will generally differ from those for the merged firm: Whereas innovation increases the joint profit of the RJV by $2[\pi_{II} - \pi_{00}]$, the corresponding value for the merged firm is $\pi_I - \pi_0$. The above result confirms the intuition that the relative size of these two profit differentials determines whether an RJV or a merged firm will be more likely to generate innovation.

However, there is a subtle effect of financial constraints: Even when the total profit effects of innovation differ for RJVs and mergers, the investments and thus the innovation probability are the same for non-degenerate parameter ranges. This happens when the budgets are intermediate, that is either $\theta^B \in [\theta^\rho, \theta_m^u]$ in case (i) or $\theta^B \in [\theta_m^\rho, \theta^u]$ in case (ii). In those cases, both the RJV and the merged firm invest their entire budgets, but the marginal return of additional research projects is not sufficient to justify the cost of borrowing from the capital market. Hence, the RJV and the merged entity each invest exactly the total research budget $2B$ into R&D.

Proposition 5 enables us to analyze whether a merger or an RJV would be better from the consumer surplus perspective, assuming that firms would want to engage in it.[27] Analogously to Proposition 3, we maintain the following weak assumptions: (a) When technology is the same, consumer surplus is higher with two active firms than with one; (b) for the same number of active firms, consumer surplus is higher if the firms have innovated than if they have not. Thus, in case (i) of Proposition 5, the RJV unambiguously increases consumer surplus. The reason is that the RJV both weakly increases the probability that the innovation will be discovered and increases competition for any level of technology. Even in case (ii), where the profit increase from innovation is larger for the merger than for the RJV, the RJV unambiguously leads to higher consumer surplus than the merger if $\theta^B \in [\theta_m^\rho, \theta^u]$, as the innovation probability is the same for both forms of cooperation. In case (ii), if $\theta^B \notin [\theta_m^\rho, \theta^u]$, the comparison is ambiguous: The merged firm would be more likely to discover the innovation, while the RJV would maintain the more competitive market structure. One consequence

[27] This will, for instance, be the case if Propositions 3 or 4 apply.

of this analysis is that from a consumer perspective, an RJV is preferable to a merger, except possibly when the innovation probability would be significantly higher under a merger. This suggests that firms should not only be required to show that a merger would have positive innovation effects, but also that these effects would not occur with an RJV.

1.4.2 Spillovers

Our model differs from the previous literature on RJVs not only by its focus on financial constraints as opposed to spillovers, but also by the feature that firms can choose between different R&D projects. To simplify the comparison with the existing literature, we first consider a variant of our project choice model without financial constraints, but with spillovers. Thereafter, we analyze the interaction between financial constraints and spillovers.

Spillovers without financial constraints

We modify the setting of Section 1.2 by assuming that the firms with cost functions $C(\theta)$ choose their investment portfolio without any budget constraint. Moreover, with R&D competition, if a firm has invested successfully in a project and the rival has not, then with probability $\sigma \in [0, 1]$ the rival will obtain access to the innovation. Thus, it is now possible that a firm obtains the innovation without investing itself.

We provide the equilibrium characterization for R&D competition in Appendix 1.A.8. As in the benchmark model, we obtain an equilibrium in double cut-off strategies. A full description of the equilibrium is given in Lemma 1.A.8. The analysis with RJVs is simpler than in the case with financial constraints. The increase in joint profit from a successful innovation is $2\pi_{II} - 2\pi_{00}$. Hence, the RJV invests in all projects up to a cut-off value, which is given by θ^u, and it does not invest in the remaining ones. The following result compares investments in the RJV with those under R&D competition.

Proposition 6. *Consider the model with spillovers, but without financial constraints. Assume that* $\pi_{I0} > \pi_{II}$. *Then the innovation probability is strictly larger under the RJV than under R&D competition if and only if*

$$\sigma > 1 - \frac{\pi_{II} - \pi_{00}}{\pi_{I0} - \pi_{II}}.$$

This condition is always satisfied if competition is soft.

The proof is in Appendix 1.A.8. As in the case with financial constraints, for RJVs to generate a higher innovation probability than R&D competition, it is crucial that the value of escaping competition is sufficiently small relative to the value of joint innovation. A simple, but important implication of Proposition 6 needs to be emphasized: When competition is soft, then the RHS of the inequality in Proposition 6 is negative and an RJV increases innovation for any level of spillovers (including $\sigma = 0$). When competition is moderate or intense, the RHS is positive, but an RJV can still increase the innovation probability if the spillovers are strong enough relative to the strength of the competition. The exception is (homogeneous) Bertrand competition, which is so intense that $\pi_{II} = \pi_{00} = 0$, so that the inequality cannot be satisfied for any $\sigma \in [0, 1]$.

Spillovers with financial constraints

In Appendix 1.A.8, we integrate the model with spillovers just discussed into the model with financial constraints. Large parts of the analysis follow directly from our results in Section 1.3. To apply those results, one needs to define the *expected* payoffs $\tilde{\pi}_{t_i t_j}$ of discovering the innovation (i.e., before any spillovers happen and taking into account the possibility of a spillover) and then observe that Assumption 1 holds with π replaced by $\tilde{\pi}$. Then, the results of Section 1.3 apply after replacing realized product market profits with expected payoffs. Adapting Assumption 2, we assume that the research budgets of the individual firms are sufficiently small that they will borrow positive amounts in any equilibrium. It is straightforward to show that there is an equilibrium in double cut-off strategies under R&D competition (see Lemma 1.A.9 in Appendix 1.A.8 for details). The comparison between R&D competition and RJV is also very similar to the case without spillovers (see Proposition 1.A.2 in Appendix 1.A.8): When the total profit increase $2\pi_{II} - 2\pi_{00}$ from innovation is high enough, then the RJV will lead to a greater innovation probability than R&D competition independent of financial constraints. If the total profit increase from innovation is lower, the RJV only leads to a greater innovation probability if both the interest rate ρ and the RJV budget $2B$ are above a threshold; in this case, the RJV saves investment costs by avoiding duplication.

The following differences to the benchmark model are relevant for the comparison between investments under R&D competition and under the RJV. First, RJVs unconditionally increase innovation whenever $2\pi_{II} - 2\pi_{00} > \pi_{I0} - \pi_{00} - \sigma(\pi_{I0} - \pi_{II})$, which is more likely to be satisfied when spillovers are strong (i.e., when σ is high). Second, when that condition is not satisfied, an increase in σ lowers the thresholds for the budget and the interest rate which are needed to guarantee that the RJV increases the innovation probability. The conditions under which an RJV increases the innovation probability are thus weaker with higher spillovers, just as they are with higher interest rates:

Proposition 7 (Benefit of RJV increases in the spillover rate)**.** *Fix any σ and ρ. If the innovation probability is strictly larger under the RJV than under R&D competition, then it is also strictly larger for any $\sigma' \geq \sigma$ and $\rho' \geq \rho$.*

As in the case without spillovers, an RJV results in efficiency gains at the investment stage by reducing duplication, and resources can be invested in a larger set of projects. Moreover, whereas spillover effects reduce investment under competitive R&D, this is not the case with an RJV. Thus, the positive effect of R&D cooperation on the innovation probability must be larger with spillovers than without, reflecting the internalization of positive spillovers by the RJV.

1.4.3 Licensing

Like research joint ventures, licensing agreements are an instrument for firms to share the fruits of innovation. The literature has demonstrated the possible benefits and costs of such agreements when R&D efforts are one-dimensional. Here, we show how the possibility of licensing influences R&D project choice in the absence of an RJV and, thereby, the effects of switching to an RJV. In particular, we will show that even when licensing of innovations is possible, RJVs can still lead to an increase in the innovation probability.

We thus extend our benchmark model to allow for licensing of innovations.[28] We suppose that, if only one firm has innovated successfully, it can license the innovation to the competitor with a two-part tariff (L, η), consisting of an output-independent fixed fee L and a variable, output-dependent part η (e.g., royalties).[29] When the unsuccessful firm licenses the innovation, both the innovator and the licensee have the technology state $t_i = I$. However, the incentives of the licensee to compete vigorously are dampened by the variable part of the licensing contract η.[30] This reduction of the intensity of competition increases total industry profits (compared to the situation when both firms independently innovate) by some amount $\Delta \geq 0$.

We assume that the innovator makes a take it or leave it offer, extracting all the rents from the licensee. In particular, the innovator sets the fixed fee L such that the unsuccessful firm earns its outside option π_{0I} and, thus, is indifferent between accepting the contract or not. Therefore, the innovator is willing to license the innovation if her profits with licensing, $2\pi_{II} + \Delta - \pi_{0I}$, are at least as high as her profits without, π_{I0}. Licensing always happens if competition is soft or moderate and sometimes when it is intense.[31] As in

[28] In Appendix 1.A.8, we describe the details of the model. Here, we sketch the main ideas.

[29] As will become clear later, if only simpler licensing contracts were available, our analysis would still apply. See Shapiro (1985) for a discussion of licensing with and without royalties. Fauli-Oller and Sandonis (2003) analyze licensing with fixed fee, royalty and two-part tariff contracts as an alternative to mergers.

[30] For example, royalties increase the licensee's marginal cost and, thus, soften competition. This leads to asymmetric product market competition, although the firms use equal technology.

[31] This is related to the result of Katz and Shapiro (1985) that, in a Cournot setting, a successful innovator will license small innovations, but not large or drastic innovations.

the analysis of spillovers in Section 1.4.2, after replacing the function $\pi_{t_i t_j}$ appropriately, the analysis directly follows Section 1.3. Specifically, we define a function π^L on $\{0, I\} \times \{0, I\}$, which is identical with π except that it takes into account licensing payments when only one firm is successful. The only difference between π^L and π is that $\pi^L_{I0} = \max\{\pi_{I0}, 2\pi_{II} + \Delta - \pi_{0I}\}$. This function captures profits as a function of technology level, but taking into account possible gains from licensing. Using this modified profit function, we derive thresholds θ^L_1 and θ^L_2 by replacing π with π^L in the definitions of θ_1 and θ_2. Crucially, whereas $\theta^L_2 = \theta_2$, $\theta^L_1 \geq \theta_1$, reflecting the potential gains from licensing.

When $2\pi_{II} + \Delta - \pi_{0I} < \pi_{I0}$, the equilibrium under R&D competition is exactly the same as in Lemma 1, because licensing never occurs in this case. When $2\pi_{II} + \Delta - \pi_{0I} \geq \pi_{I0}$, licensing increases the innovation probability in any equilibrium to $\theta^L_1 \geq \theta_1$, as the opportunity to license increases the incentives to explore further projects.

For the comparison with the RJV, we replace the budget threshold $\bar{B}(\rho)$ and the interest threshold $\bar{\rho}$ with thresholds $\bar{B}^L(\rho)$ and $\bar{\rho}^L$ that are based on π^L rather than π, leading to the following modification of Proposition 1.

Proposition 8 (Comparison of R&D competition with licensing and RJV)**.**

(i) Suppose $2\pi_{II} + \Delta - \pi_{0I} \geq \pi_{I0}$. *Then:*
(a) The innovation probability is strictly larger under the RJV than under competition if and only if $B > \bar{B}^L(\rho)$ *and* $\rho > \bar{\rho}^L$.
(b) If the formation of the RJV strictly increases the innovation probability, then it weakly decreases total R&D spending.

(ii) Suppose $2\pi_{II} + \Delta - \pi_{0I} < \pi_{I0}$. *Then the effect of an RJV on the innovation probability is the same as in the absence of a licensing possibility.*

In case (i), firms want to license the innovation. In case (ii), they do not. Importantly, the conditions under which the RJV leads to a higher innovation probability are more rigid than without licensing. This is obvious in the case of soft competition, in which Proposition 1(i) states that an RJV is *always* preferable to R&D competition, while Proposition 8(i) requires that $B > \bar{B}^L(\rho)$ and $\rho > \bar{\rho}^L$. When competition is not soft, the conditions under which an RJV increases the innovation probability are also more restrictive with licensing than without, since $\bar{B}^L(\rho) \geq \bar{B}(\rho)$ and $\bar{\rho}^L \geq \bar{\rho}$ whenever Proposition 8(i) applies. The difference arises because licensing increases innovation incentives under R&D competition, so that there is less to gain from an RJV. Moreover, an RJV that increases the innovation probability weakly decreases total R&D spending, because it invests weakly less than the available budget while both firms invest strictly more than their budget under R&D competition.

To put the results into perspective, we can think of ex-post licensing and RJVs as imperfect substitutes for sharing the fruits of R&D. Nonetheless, the above results show that even when ex-post licensing is possible, an RJV may still lead to a higher innovation probability than R&D competition if financial constraints are sufficiently tight.

1.4.4 Multiple firms

We extend our model by allowing for more than two competing firms. With multiple firms, there are many conceivable ways in which RJVs could be formed, including industry-wide RJVs as well as several competing RJVs. We analyze two illustrative cases. First, we consider a market with three firms that can form an industry-wide RJV. Second, we consider the case of four firms that form two competing RJVs. The analysis is very similar to the benchmark model with two firms. Therefore, we defer details to the Appendix 1.A.9.

Industry-wide RJV

We extend the analysis to the case of three firms, which can form one RJV. Suitably adjusting Assumptions 1 and 2, the analysis and results are analogous to the benchmark model with two firms. The only notable difference is that the R&D competition game now has multiple equilibria in triple cut-off strategies characterized by the three critical values $\theta_3 \leq \theta_2 \leq \theta_1$.[32] However, the innovation probability in any equilibrium is still given by θ_1, the most expensive project in which a single firm can profitably invest relying on external resources. The analysis of the RJV when all firms participate and the resulting comparison between R&D competition and cooperation is qualitatively unchanged. Therefore, we find similar results to Proposition 1: When competition is not too intense, the innovation probability is higher in the RJV; otherwise, this conclusion requires the budget and the external financing costs to be high enough. In the latter case, total R&D-spending in the RJV is lower than under competition.

Multiple RJVs

Next, we consider the formation of multiple RJVs. We consider a market with four firms that form two symmetric RJVs, each with two firms. Therefore, R&D cooperation does not eliminate competition in the innovation stage entirely, but reduces the number of competing agents. Hence, even with an RJV cheap projects are still duplicated. We assume that the budget of an RJV is sufficiently large that it never borrows in equilibrium. Otherwise, the analysis of two competing RJVs turns out to be similar to the R&D competition regime in the baseline model. Analogously to Proposition 1, we find: When competition is relatively soft, then the innovation probability is higher with two RJVs than with R&D competition without additional conditions. Under relatively moderate or intense competition, cooperation on R&D increases the innovation probability only if the budget and the interest rate are sufficiently high. In this case, total R&D-spending with two RJVs is lower than when four firms invest individually.

1.5 Relation to the Literature

Our paper analyzes R&D competition between duopolists who (i) select between different R&D projects and (ii) are financially constrained. It compares

[32] All three firms invest below θ_3, two between θ_3 and θ_2, one between θ_2 and θ_1, and none above θ_1.

their R&D decisions with those of research joint ventures and merged firms. Accordingly, we briefly discuss the relation of our paper to existing treatments of R&D project choice, financially constrained oligopolists, research joint ventures and mergers and innovation.

Innovation project choice: Our model of R&D competition with project choice builds on Letina (2016) who also considers symmetric incumbents. Letina et al. (2021) apply that framework to study the innovation decisions of asymmetric firms (an incumbent and an entrant). These papers neither include financial constraints, nor do they address joint ventures. Contrary to these models, Moraga-González, Motchenkova, and Nevrekar (2022) allow for (two) different types of R&D, but fix the overall spending.[33]

Financially constrained firms: Authors such as Hall and Lerner (2010) and Kerr and Nanda (2015) have stated several reasons why external financing of R&D investments is more costly than for other investments.[34] As a result, internal financing plays a strong role (Czarnitzki and Hottenrott (2011)). Several authors have provided empirical evidence that financial constraints have a negative impact on R&D investment (Mohnen, Palm, Van Der Loeff, and Tiwari (2008), Savignac (2008), Mancusi and Vezzulli (2014), Howell (2017) Krieger, Li, and Papanikolaou (2022), Caggese (2019)).[35] In line with our Propositions 3 and 4, Sovinsky (2022) finds that capital-constrained firms are more likely to join an RJV. While the empirical literature on financially constrained firms is voluminous, the theoretical literature is small.[36] We are not aware of any oligopoly model of financially constrained firms (with or without RJV formation) that choose how much as well as in which projects to invest.

[33]Further, less closely related models of R&D project choice include Gilbert (2019), Bryan and Lemus (2017), Letina and Schmutzler (2019), Bardey, Jullien, and Lozachmeur (2016) and Bavly, Heller, and Schreiber (2022).

[34]Examples are the riskiness of the investments and the difficulty of providing collateral, as physical assets are relatively less important than human capital.

[35]Czarnitzki and Hottenrott (2011) also find that the availability of internal financing has a larger impact on R&D than on capital investment and that cutting-edge innovation projects, like basic research, are more prone to financial constraints in the credit market as they are riskier. These empirical findings suggest that budget-constrained firms can benefit from becoming unconstrained by joining an RJV. Moreover, the empirical results support the relevance of our budget constraint assumption.

[36]One exception is Fumagalli, Motta, and Tarantino (2022), who consider acquisitions of startups that might be financially constrained.

The theory of research joint ventures: Our paper differs from the existing theoretical literature on RJVs in two important dimensions. First, the literature does not allow for different R&D projects. Second, it does not model the role of financial constraints. Instead, it focuses mainly on spillovers. Without RJVs, as in our model, firms invest in R&D to gain a competitive advantage over their rivals. However, if knowledge spillovers to competitors are large enough, such gains are small and firms limit their R&D investments. Accordingly, d'Aspremont and Jacquemin (1988) show in a static, two-stage, duopolistic Cournot model that, with high spillovers, an RJV leads to larger R&D expenditures, more output and higher welfare than R&D competition. By contrast, with low spillovers, welfare under R&D competition is higher than in an RJV, as an RJV would lower total R&D investments.[37] As argued above, we find that, with soft competition (for instance, with sufficiently differentiated price competition), an RJV increases the investment probability (and hence welfare) even with low spillovers. Like our paper, Katz (1986) and Kamien et al. (1992) show how the nature of product market competition affects the comparison between R&D competition and cooperation. Like other authors, such as Amir et al. (2019), both papers also argue that an RJV reduces wasteful effort duplication. However, contrary to our approach, none of these papers explicitly models duplication in the natural setting where firms can select between different projects. Instead, firms can only choose the amount of R&D investment, which is in a strictly positive relation with the R&D outcome (the size or probability of an innovation). In our model, an RJV may well increase the innovation probability while investing less resources because duplication is eliminated. This feature is particularly relevant when there is fundamental uncertainty about the right approach to R&D.[38] The literature has also highlighted important caveats to the claim that research

[37] Suzumura (1992) obtains similar results with more than two firms, and he investigates how the outcomes with and without R&D cooperation relate to the social optimum. Amir, Liu, Machowska, and Resende (2019) show that the gap between market outcome and social optimum increases in the spillover rate. However, subsidies can help to achieve the second-best social optimum.

[38] In broadly related work, Kamien and Zang (2000) allow firms to choose different research approaches, but approaches only differ in their spillover rates, and each approach will succeed with certainty, which is in stark contrast to our model. Other important lines of research include stochastic R&D (Choi (1993)) and absorptive capacity, whereby spillovers are increasing in own R&D (Kamien and Zang (2000)).

cooperation is socially beneficial: RJVs foster product market collusion, which leads to dynamic inefficiency (Grossman and Shapiro (1986), Martin (1996), Jacquemin (1988), Caloghirou, Ioannides, and Vonortas (2003) and Miyagiwa (2009)).[39] Competition authorities who decide on such RJVs have to weigh these risks against the potential benefits, which is difficult given realistic informational constraints (Cassiman, 2000).

The empirics of research joint ventures: Empirical studies support the claims of the theoretical literature. Cassiman and Veugelers (2002) show that Belgian manufacturing firms are more likely to cooperate when spillovers are high. Aschhoff and Schmidt (2008) show that R&D cooperation among competitors reduces production costs. Becker and Dietz (2004) provide empirical evidence that members of an RJV invest more in research than without cooperation, and that they are more likely to obtain new products. Veugelers (1997) also finds that cooperation increases investments, but that this requires absorptive capacity. Further, she concludes that firms are more likely to join an RJV the more they spend on R&D. Röller, Siebert, and Tombak (2007) show that cost-sharing motives are important for RJV formation. Link (1998) provides case-study evidence for efficiencies in a specific RJV. Finally, Duso et al. (2014) and Sovinsky (2022) find empirical evidence that RJVs among competitors are more prone to collusion, which reduces welfare. Thus, horizontal R&D cooperation should come under scrutiny by authorities.

Mergers and innovation: Several authors have recently studied under which circumstances incumbent mergers increase innovation. Federico, Langus, and Valletti (2017, 2018) and Motta and Tarantino (2021) identify negative effects in models with one-dimensional R&D effort; similarly, Letina (2016) and Gilbert (2019) obtain negative effects on R&D diversity in models of project choice. Denicolò and Polo (2018) find positive effects. In Bourreau, Jullien, and Lefouili (2021), both possibilities arise, where the positive effects come from allowing for horizontal rather than only vertical R&D innovations. In our model with project choices of financially constrained

[39] Conversely, Vilasuso and Frascatore (2000) show that, if forming RJVs is costly, firms may form less RJVs than socially optimal; similarly Falvey, Poyago-Theotoky, and Teerasuwannajak (2013).

firms who engage in purely vertical innovations, we similarly find that the effects of the merger can be positive or negative. Contrary to Bourreau et al. (2021), however, the possibility of a positive effect reflects the merged entity's ability to coordinate which projects to invest in and the existence of financial constraints. Moreover, which effect occurs depends on the intensity of product market competition.

1.6 Conclusion

This paper provides a novel theory of research joint ventures for financially constrained firms who can choose the set of research projects that they will pursue. Research joint ventures allow firms to share their R&D budget and to coordinate their R&D investment decisions, while maintaining product market competition.

We find that, if product market competition is sufficiently soft, the RJV will increase the probability of an innovation even when there are no financial constraints. As product market competition increases, a positive innovation effect of the RJV requires that the external funding conditions are sufficiently bad and the budget of the RJV is sufficiently large. In the latter case, the RJV reduces research costs by avoiding duplication – this shows that the relation between R&D spending and R&D success probability need not be positive. Moreover, any RJV that increases the innovation probability also increases expected consumer welfare.

Importantly, the conditions under which the RJV increases the probability of a successful innovation and the conditions under which it is profitable for the participants often coincide; in particular, for soft or intermediate competition, firms always want to form RJVs if they increase the innovation probability. This increases consumer welfare under mild conditions. Nonetheless, we also identify situations under which firms find it profitable to form an innovation-reducing RJV merely because they can coordinate on reducing R&D costs, which is in line with concerns of policy makers.

We obtain qualitatively similar results on the effects of mergers on innovation. More interestingly, we find conditions under which a merger does not lead to a lower innovation probability than an RJV. In such situations, even if the merger has pro-competitive effects on innovation relative to the benchmark of R&D competition, the merger should be prohibited because, contrary to the alternative of an RJV, it results in an adverse effect on product market competition.

1.A Proofs

1.A.1 Proof of Lemma 1

We will first prove an intermediate result.

Lemma 1.A.1. *Any strategy r_i such that $\int_0^1 r_i(\theta)C(\theta)d\theta \leq B$ is dominated.*

Proof. If $\int_0^1 r_i(\theta)C(\theta)d\theta \leq B$, then by Assumption 2 there exists a set $\Theta' \subseteq [0, \theta_2)$ of positive measure, such that $r_i(\theta) = 0$ for all $\theta \in \Theta'$. Consider a strategy r_i', where $r_i'(\theta) = 1$ for all $\theta \in [0, \theta_2)$ and $r_i'(\theta) = r_i(\theta)$ otherwise. We will show that $\mathbb{E}\Pi_i(r_i', r_j) > \mathbb{E}\Pi_i(r_i, r_j)$ for any strategy of the opponent r_j.

Noting that the strategy r_i' requires external financing (while r_i does not), and taking into account that $r_i'(\theta) = r_i(\theta)$ for all $\theta > \theta_2$ then

$$\begin{aligned}
&\mathbb{E}\Pi_i(r_i', r_j) - \mathbb{E}\Pi_i(r_i, r_j) = \\
&\int_0^{\theta_2} (r_i'(\theta) - r_i(\theta)) \Big[(1 - r_j(\theta)) \left[\pi_{IO} - \pi_{OO}\right] + r_j(\theta) \left[\pi_{II} - \pi_{OI}\right]\Big] d\theta \\
&\qquad -(1+\rho) \int_0^1 r_i'(\theta)C(\theta)d\theta + \rho B + \int_0^1 r_i(\theta)C(\theta)d\theta \\
&\geq \int_0^{\theta_2} (r_i'(\theta) - r_i(\theta)) \Big[(1 - r_j(\theta)) \left[\pi_{IO} - \pi_{OO}\right] + r_j(\theta) \left[\pi_{II} - \pi_{OI}\right]\Big] d\theta \\
&\qquad -(1+\rho) \int_0^1 r_i'(\theta)C(\theta)d\theta + (1+\rho) \int_0^1 r_i(\theta)C(\theta)d\theta \\
&= \int_0^{\theta_2} (r_i'(\theta) - r_i(\theta)) \Big[(1 - r_j(\theta)) \left[\pi_{IO} - \pi_{OO}\right] + r_j(\theta) \left[\pi_{II} - \pi_{OI}\right]\Big] d\theta \\
&\qquad -(1+\rho) \int_0^{\theta_2} (r_i'(\theta) - r_i(\theta))C(\theta)d\theta \\
&\geq \int_0^{\theta_2} (r_i'(\theta) - r_i(\theta)) \Big[(\pi_{II} - \pi_{OI}) - (1+\rho)C(\theta)\Big] d\theta \\
&> 0.
\end{aligned}$$

The first inequality follows from the assumption that $\int_0^1 r_i(\theta)C(\theta)d\theta \leq B$, the second from the fact that $\pi_{IO} - \pi_{OO} \geq \pi_{II} - \pi_{OI}$ and $r_i'(\theta) - r_i(\theta) \geq 0$ and the last from the fact that $\pi_{II} - \pi_{OI} > (1+\rho)C(\theta)$ for all $\theta < \theta_2$ and $r_i'(\theta) > r_i(\theta)$ on the set of positive measure Θ'. $\square$

Proof of (i): Take any strategy r_j which corresponds to one of the equilibrium strategies given in Lemma 1. Note that for any fixed r_j, the equilibrium candidate strategy r_i is uniquely determined. Suppose (r_i, r_j) does not constitute an equilibrium. Then, there exists a strategy r_i' such that $\mathbb{E}\Pi_i(r_i', r_j) > \mathbb{E}\Pi_i(r_i, r_j)$. By Assumption 2, all equilibrium candidates satisfy $\int_0^1 r_i(\theta)C(\theta)d\theta > B$. Moreover, by Lemma 1.A.1 we can focus on strategies such that $\int_0^1 r_i'(\theta)C(\theta)d\theta > B$ is satisfied.

Denote the expected total payoff of project θ, conditional on it being correct, as $v_i(\theta, r_i, r_j)$. Then there exists a set $\Theta' \subseteq [0,1)$ with positive measure such that $v_i(\theta, r_i', r_j) > v_i(\theta, r_i, r_j)$ for all $\theta \in \Theta'$, or more explicitly:

$$\begin{aligned}&(1-r_j(\theta))[r_i'(\theta)\pi_{I0} + (1-r_i'(\theta))\pi_{00}] + r_j(\theta)[r_i'(\theta)\pi_{II} + (1-r_i'(\theta))\pi_{0I}]\\&-(1+\rho)C(\theta)r_i'(\theta) > (1-r_j(\theta))[r_i(\theta)\pi_{I0} + (1-r_i(\theta))\pi_{00}]\\&+r_j(\theta)[r_i(\theta)\pi_{II} + (1-r_i(\theta))\pi_{0I}] - (1+\rho)C(\theta)r_i(\theta).\end{aligned} \tag{1.3}$$

If $\theta < \theta_2$ then $r_j(\theta) = 1$ so this inequality simplifies to

$$r_i'(\theta)(\pi_{II} - \pi_{0I} - (1+\rho)C(\theta)) > r_i(\theta)(\pi_{II} - \pi_{0I} - (1+\rho)C(\theta)) \tag{1.4}$$

Since for $\theta < \theta_2$ we have $\pi_{II} - \pi_{0I} - (1+\rho)C(\theta) > 0$ and $r_i(\theta) = 1$, this would imply $r_i'(\theta) > 1$ which is a contradiction.

If $\theta > \theta_1$ then $r_j(\theta) = 0$ so inequality (1.3) simplifies to

$$r_i'(\theta)[\pi_{I0} - \pi_{00} - (1+\rho)C(\theta)] > r_i(\theta)[\pi_{I0} - \pi_{00} - (1+\rho)C(\theta)]. \tag{1.5}$$

Since for $\theta > \theta_1$ we have $\pi_{I0} - \pi_{00} - (1+\rho)C(\theta) < 0$ and $r_i(\theta) = 0$ this would imply $r_i'(\theta) < 0$ which is a contradiction.

Next, consider $\theta \in (\theta_2, \theta_1)$. This case only arises if $\theta_2 < \theta_1$, which immediately implies $\pi_{I0} + \pi_{0I} - \pi_{II} - \pi_{00} > 0$. If $r_j(\theta) = 1$ then, as before, inequality (1.3) simplifies to (1.4). However, now $\pi_{II} - \pi_{0I} - (1+\rho)C(\theta) < 0$ and, for the candidate equilibrium, $r_i(\theta) = 0$. (1.4) would thus require that $r_i'(\theta) < 0$, which is a contradiction. Similarly if $r_j(\theta) = 0$ the inequality (1.3) simplifies to (1.5), but $\theta < \theta_1$ implies $\pi_{I0} - \pi_{00} - (1+\rho)C(\theta) > 0$ and, for the candidate equilibrium, $r_i(\theta) = 1$. (1.5) would thus require that $r_i'(\theta) > 1$, which is a contradiction.

Proof of (ii): Suppose there exist two strategies, r_i and r_j, which constitute an equilibrium, and a set of positive measure $I \subseteq [0,1)$, such that r_i is different from the strategies characterized in the Lemma at all points of the set I. By Lemma 1.A.1 we can focus on strategies such that the budget is binding. Let $I_1 = I \cap (0, \theta_2)$, $I_2 = I \cap (\theta_2, \theta_1)$ and $I_3 = I \cap (\theta_1, 1)$. Note that at least one of the sets I_1, I_2, or I_3 has positive measure.

Define

$$\begin{aligned}\Gamma_i(\theta, r_j) =&\pi_{I0} - \pi_{00} - (1+\rho)C(\theta) \\ &- r_j(\theta)(\pi_{I0} + \pi_{0I} - \pi_{II} - \pi_{00}).\end{aligned}$$

We can express $v_i(\theta, r_i, r_j)$, the expected total payoff of project θ, conditional on it being correct, as

$$v_i(\theta, r_i, r_j) = r_i(\theta)\Gamma_i(\theta, r_j) + (1 - r_j(\theta))\pi_{00} + r_j(\theta)\pi_{0I}.$$

Since $\pi_{I0} + \pi_{0I} - \pi_{II} - \pi_{00} \geq 0$, $\Gamma_i(\theta, r_j)$ is decreasing in $r_j(\theta)$.

Assume first that I_1 has positive measure. Then $r_i(\theta) = 0$ for all $\theta \in I_1$. Since $C(\theta)$ is strictly increasing and $(1+\rho)C(\theta_2) = \pi_{II} - \pi_{0I}$, then $\Gamma_i(\theta, r_j) > 0$ for any r_j. Thus, the best response of firm i is $r_i(\theta) = 1$ for all $\theta \in I_1$, which is a contradiction.

Next, assume I_3 has positive measure. Then $r_i(\theta) = 1$ for all $\theta \in I_3$. But, analogously to before, $\Gamma_i(\theta, r_j) < 0$ for any r_j. Thus, the best response of firm i is $r_i(\theta) = 0$. A contradiction.

Finally, assume I_2 has positive measure, which implies that $r_i(\theta) = r_j(\theta)$ for all $\theta \in I_2$. Suppose first that $r_i(\theta) = 0$ on a set of positive measure $I_2' \subseteq I_2$. Observe that $\Gamma_j(\theta, r_i) > 0$ for all $\theta \in I_2'$. Since this is an equilibrium, it must be that $r_j(\theta) = 1$ for all $\theta \in I_2'$. A contradiction. Next, suppose that $r_i(\theta) = 1$ on a set of positive measure $I_2'' \subseteq I_2$. Observe that $\Gamma_j(\theta, r_i) < 0$ for all $\theta \in I_2''$.

Analogously to the argument above, it must be that $r_j(\theta) = 0$ for all $\theta \in I_2''$, a contradiction. Thus, it cannot be that $r_i(\theta) = r_j(\theta)$ for all $\theta \in I_2$. □

1.A.2 Proof of Lemma 2

We can rewrite the expected total payoff of the RJV as

$$\mathbb{E}\Pi_v(r_v) = 2\pi_{OO} + 2\left[\pi_{II} - \pi_{OO}\right]\int_0^1 r_v(\theta)d\theta \\ - \int_0^1 r_v(\theta)C(\theta)d\theta - \rho \max\left\{\int_0^1 r_v(\theta)C(\theta)d\theta - 2B, 0\right\}$$

where the probability that the RJV discovers the innovation is given by $\int_0^1 r_v(\theta)d\theta$ while $\int_0^1 r_v(\theta)C(\theta)d\theta + \rho \max\left\{\int_0^1 r_v(\theta)C(\theta)d\theta - 2B, 0\right\}$ captures total innovation costs.

Since research projects only differ with respect to investment costs and these costs are increasing in θ, for any fixed probability of innovation $\hat{\theta}$, the RJV optimally chooses a cut-off strategy to obtain this probability: It sets $r_v(\theta) = 1$ for $\theta < \hat{\theta}$ and $r_v(\theta) = 0$ otherwise, so that $\int_0^1 r_v(\theta)C(\theta)d\theta = \int_0^{\hat{\theta}} C(\theta)d\theta$.

The RJV's optimal portfolio can be obtained by maximizing

$$\mathbb{E}\hat{\Pi}_v(\hat{\theta}) = 2\pi_{OO} + 2\left[\pi_{II} - \pi_{OO}\right]\hat{\theta} - \int_0^{\hat{\theta}} C(\theta)d\theta - \rho \max\left\{\int_0^{\hat{\theta}} C(\theta)d\theta - 2B, 0\right\}.$$

Note that

$$\frac{\partial \mathbb{E}\hat{\Pi}_v}{\partial \hat{\theta}} = \begin{cases} 2\left[\pi_{II} - \pi_{OO}\right] - C(\hat{\theta}) & \text{for } \hat{\theta} < \theta^B \\ 2\left[\pi_{II} - \pi_{OO}\right] - (1+\rho)C(\hat{\theta}) & \text{for } \hat{\theta} > \theta^B. \end{cases}$$

Now consider the three cases from the proposition (i.e., whether $\theta^B < \theta^\rho$,

$\theta^B \in [\theta^\rho, \theta^u]$, or $\theta^B > \theta^u$). First, if $\theta^B < \theta^\rho$ then

$$\frac{\partial \mathbb{E}\hat{\Pi}_v}{\partial \hat{\theta}} = \begin{cases} 2\left[\pi_{II} - \pi_{00}\right] - C(\hat{\theta}) > 0 & \text{for } \hat{\theta} < \theta^B \\ 2\left[\pi_{II} - \pi_{00}\right] - (1+\rho)C(\hat{\theta}) > 0 & \text{for } \hat{\theta} \in (\theta^B, \theta^\rho) \\ 2\left[\pi_{II} - \pi_{00}\right] - (1+\rho)C(\hat{\theta}) < 0 & \text{for } \hat{\theta} \in (\theta^\rho, 1). \end{cases}$$

Thus, $\hat{\theta} = \theta^\rho$ maximizes the expected return of the RJV's portfolio. Second, if $\theta^B \in [\theta^\rho, \theta^u]$ then

$$\frac{\partial \mathbb{E}\hat{\Pi}_v}{\partial \hat{\theta}} = \begin{cases} 2\left[\pi_{II} - \pi_{00}\right] - C(\hat{\theta}) > 0 & \text{for } \hat{\theta} < \theta^B \\ 2\left[\pi_{II} - \pi_{00}\right] - (1+\rho)C(\hat{\theta}) < 0 & \text{for } \hat{\theta} > \theta^B, \end{cases}$$

so that $\hat{\theta} = \theta^B$ maximizes the expected return of the RJV's portfolio. Third, if $\theta^B > \theta^u$ then

$$\frac{\partial \mathbb{E}\hat{\Pi}_v}{\partial \hat{\theta}} = \begin{cases} 2\left[\pi_{II} - \pi_{00}\right] - C(\hat{\theta}) > 0 & \text{for } \hat{\theta} < \theta^u \\ 2\left[\pi_{II} - \pi_{00}\right] - C(\hat{\theta}) < 0 & \text{for } \hat{\theta} \in (\theta^u, \theta^B) \\ 2\left[\pi_{II} - \pi_{00}\right] - (1+\rho)C(\hat{\theta}) < 0 & \text{for } \hat{\theta} \in (\theta^B, 1). \end{cases}$$

Thus, $\hat{\theta} = \theta^u$ maximizes the expected return of the RJV's portfolio. □

1.A.3 Proof of Proposition 1

First, we provide a lemma distinguishing the two parts of Proposition 1.

Lemma 1.A.2. $2\pi_{II} > \pi_{I0} + \pi_{00} \Leftrightarrow \theta^\rho > \theta_1$.

Proof.

$$\begin{aligned} 2\pi_{II} &> \pi_{I0} + \pi_{00} \\ 2[\pi_{II} - \pi_{00}] &> \pi_{I0} - \pi_{00} \\ (1+\rho)C(\theta^\rho) &> (1+\rho)C(\theta_1) \\ \theta^\rho &> \theta_1. \end{aligned}$$

□

(i) By Lemma 1.A.2, $2\pi_{II} > \pi_{I0} + \pi_{00}$ implies $\theta^\rho > \theta_1$. By Lemma 2, the probability that the RJV innovates is at least θ^ρ. By Lemma 1, the probability of innovation under competition is θ_1. Therefore, the probability that the innovation will be discovered is strictly larger under the RJV than under competition.

(ii) To prove part (a), we first provide an auxiliary result (Lemma 1.A.3). Using this lemma, we separately show that "if" part follows from Lemma 1.A.4 below and "only if" part from Lemma 1.A.5 below.

Lemma 1.A.3. *Suppose* $2\pi_{II} \leq \pi_{I0} + \pi_{00}$. *Then* $\rho > \bar{\rho} \Leftrightarrow \theta^u > \theta_1$.

Proof. First suppose that $\bar{\rho} < \infty$. Then

$$\begin{aligned} \rho > \bar{\rho} &= \frac{\pi_{I0} - 2\pi_{II} + \pi_{00}}{2[\pi_{II} - \pi_{00}]} \\ 2\rho[\pi_{II} - \pi_{00}] &> \pi_{I0} - 2\pi_{II} + \pi_{00} \\ 2(1+\rho)[\pi_{II} - \pi_{00}] &> \pi_{I0} - \pi_{00} \\ 2[\pi_{II} - \pi_{00}] &> \frac{\pi_{I0} - \pi_{00}}{1+\rho} \\ C(\theta^u) &> C(\theta_1) \\ \theta^u &> \theta_1. \end{aligned}$$

Next suppose $\bar{\rho} = \infty$. Then, clearly $\rho < \bar{\rho}$. Hence, the statement of the lemma holds if and only if $\theta^u \leq \theta_1$ or $2[\pi_{II} - \pi_{00}] \leq \frac{\pi_{I0} - \pi_{00}}{1+\rho}$. As $\bar{\rho} = \infty$ implies $\pi_{II} = \pi_{00}$, this requirement holds. □

Lemma 1.A.4. *Suppose* $2\pi_{II} \leq \pi_{I0} + \pi_{00}$. *If* $B > \bar{B}(\rho)$ *and* $\rho > \bar{\rho}$, *then the probability that the innovation will be discovered is strictly larger under the RJV than under competition.*

Proof. $B > \int_0^{\theta_1} C(\theta)d\theta/2$ implies $\int_0^{\theta^B} C(\theta)d\theta > \int_0^{\theta_1} C(\theta)d\theta$ and therefore $\theta_1 < \theta^B$. Furthermore, by Lemma 1.A.2, $\theta^\rho \leq \theta_1$ so that $\theta^\rho < \theta^B$. Then, either $\theta^B \in (\theta^\rho, \theta^u)$ or $\theta^B \geq \theta^u$. If $\theta^B \in (\theta^\rho, \theta^u)$, then the RJV invests in all projects in the set $(0, \theta^B)$ and discovers the innovation with probability θ^B.

Without the RJV, in any equilibrium, the firms invest in projects in the set $(0, \theta_1)$ and the innovation is discovered with probability θ_1. Since $\theta^B > \theta_1$ it immediately follows that the probability of innovation strictly increases under the RJV.

Next, suppose $\theta^B \geq \theta^u$. Then, the RJV invests in all projects in the set $(0, \theta^u)$ and discovers the innovation with probability θ^u. Since $\rho > \bar{\rho}$ implies $\theta^u > \theta_1$ by Lemma 1.A.3, it follows that the probability of innovation strictly increases under the RJV. □

Lemma 1.A.5. *Suppose $2\pi_{II} \leq \pi_{IO} + \pi_{OO}$. If the probability that the innovation will be discovered is strictly larger under the RJV than under competition, then $B > \bar{B}(\rho)$ and $\rho > \bar{\rho}$.*

Proof. As $2\pi_{II} \leq \pi_{IO} + \pi_{OO}$, Lemma 1.A.2 implies $\theta^\rho \leq \theta_1$. Hence, if the probability that the innovation will be discovered is strictly larger under the RJV than under competition, then $\theta^B > \theta^\rho$ by Lemma 2. Therefore, either $\theta^B \in (\theta^\rho, \theta^u)$ or $\theta^B \geq \theta^u$. If $\theta^B \in (\theta^\rho, \theta^u)$, then, by Lemma 2, the increase in the probability of discovering the innovation under the RJV implies $\theta^B > \theta_1$, so that $\theta^u > \theta^B > \theta_1$. If $\theta^B \geq \theta^u$, then the increase in the probability of discovering the innovation under RJV implies $\theta^u > \theta_1$, so that $\theta^B \geq \theta^u > \theta_1$. In either case, both $\theta^u > \theta_1$ and $\theta^B > \theta_1$.

Note that $\theta^B > \theta_1$ implies

$$\int_0^{\theta^B} C(\theta) d\theta > \int_0^{\theta_1} C(\theta) d\theta.$$

It follows immediately that $B > \int_0^{\theta_1} C(\theta) d\theta / 2 = \bar{B}(\rho)$. Furthermore, $\theta^u > \theta_1$ implies, by Lemma 1.A.3, that $\rho > \bar{\rho}$. □

Finally, we prove part (b) of (ii). With moderate or intense competition, $2\pi_{II} \leq \pi_{IO} + \pi_{OO}$. If the formation of the RJV strictly increases the probability of discovering the innovation, then, by Lemma 1.A.5, $B > \bar{B}(\rho)$ and $\rho > \bar{\rho}$. Using the same argument as in the proof of Lemma 1.A.5, we have that either $\theta^B \in (\theta^\rho, \theta^u)$ or $\theta^B \geq \theta^u$. If $\theta^B \in (\theta^\rho, \theta^u)$, then by Lemma 2, the total costs of the RJV are

$$\int_0^{\theta^B} C(\theta)d\theta = 2B. \tag{1.6}$$

If $\theta^B \geq \theta^u$, then by Lemma 2, the total cost of the RJV are

$$\int_0^{\theta^u} C(\theta)d\theta \leq 2B. \tag{1.7}$$

By Lemma 1, the total costs for equilibrium strategies r_i^* and r_j^* under competition are

$$\begin{aligned}
(1+\rho)\int_0^1 &[r_i^*(\theta)+r_j^*(\theta)]C(\theta)d\theta - 2\rho B \\
&= (1+\rho)\left[2\int_0^{\theta_2} C(\theta)d\theta + \int_{\theta_2}^{\theta_1}[r_i^*(\theta)+r_j^*(\theta)]C(\theta)d\theta\right] - 2\rho B \\
&> (1+\rho)\left[2B + \int_{\theta_2}^{\theta_1}[r_i^*(\theta)+r_j^*(\theta)]C(\theta)d\theta\right] - 2\rho B \\
&= 2B + (1+\rho)\int_{\theta_2}^{\theta_1}[r_i^*(\theta)+r_j^*(\theta)]C(\theta)d\theta \\
&\geq 2B,
\end{aligned}$$

where the first inequality follows from Assumption 2, and the second inequality from $\theta_1 \geq \theta_2$ and $r_i^*(\theta)+r_j^*(\theta) \geq 0$ for any θ.

It immediately follows that the total cost under competition is weakly larger than the total cost under RJV, which proves the proposition. □

1.A.4 Proof of Proposition 2

Denote with $\mathcal{P}_{II}^{com}$ the probability that both firms discover the innovation under competition and with $\mathcal{P}_{I0}^{com}$ the probability that a single firm discovers the innovation under competition. Analogously, let $\mathcal{P}_{II}^{rjv}$ be the probability that the innovation is discovered under the RJV.

The expected consumer surplus is strictly higher under RJV than under competition if

$$\begin{aligned} &\mathcal{P}_{II}^{rjv} CS_{II} + \left[1 - \mathcal{P}_{II}^{rjv}\right] CS_{00} > \\ \mathcal{P}_{II}^{com} CS_{II} + \mathcal{P}_{I0}^{com} CS_{I0} &+ \left[1 - \mathcal{P}_{II}^{com} - \mathcal{P}_{I0}^{com}\right] CS_{00}. \end{aligned} \tag{1.8}$$

We proceed to show that this holds under the assumptions of the proposition. First, observe that by Assumption 3, $CS_{II} > CS_{I0}$, so that

$$\begin{aligned} \left[\mathcal{P}_{II}^{com} + \mathcal{P}_{I0}^{com}\right] CS_{II} &+ \left[1 - \mathcal{P}_{II}^{com} - \mathcal{P}_{I0}^{com}\right] CS_{00} \geq \\ \mathcal{P}_{II}^{com} CS_{II} + \mathcal{P}_{I0}^{com} CS_{I0} &+ \left[1 - \mathcal{P}_{II}^{com} - \mathcal{P}_{I0}^{com}\right] CS_{00}. \end{aligned} \tag{1.9}$$

Second, since the RJV increases the probability of innovation, $\mathcal{P}_{II}^{rjv} > \mathcal{P}_{II}^{com} + \mathcal{P}_{I0}^{com}$ and since by Assumption 3, $CS_{II} > CS_{00}$, it must be that

$$\begin{aligned} &\mathcal{P}_{II}^{rjv} CS_{II} + \left[1 - \mathcal{P}_{II}^{rjv}\right] CS_{00} > \\ \left[\mathcal{P}_{II}^{com} + \mathcal{P}_{I0}^{com}\right] CS_{II} &+ \left[1 - \mathcal{P}_{II}^{com} - \mathcal{P}_{I0}^{com}\right] CS_{00}. \end{aligned} \tag{1.10}$$

Finally, observe that combining inequalities (1.10) and (1.9) gives inequality (1.8), which completes the proof. □

1.A.5 Profitability of RJVs

We now prove the results providing conditions under which firms profit from forming RJVs that increase and decrease the probability of innovation, respectively.

Proof of Proposition 3

Using (1.2), the RJV strictly increases gross profits if and only if

$$\begin{aligned} &2\theta^* \pi_{II} + 2\left(1 - \theta^*\right) \pi_{00} > \\ 2\theta_2 \pi_{II} + \left(\theta_1 - \theta_2\right) &\left(\pi_{I0} + \pi_{0I}\right) + 2\left(1 - \theta_1\right) \pi_{00}, \end{aligned} \tag{1.11}$$

which can be rewritten as

$$2(\theta_1 - \theta_2)\pi_{II} + 2(\theta^* - \theta_1)\pi_{II} > \tag{1.12}$$
$$(\theta_1 - \theta_2)(\pi_{IO} + \pi_{OI}) + 2(\theta^* - \theta_1)\pi_{OO}.$$

(i) By Lemma 1.A.2, soft competition ($2\pi_{II} > \pi_{IO} + \pi_{OO}$), implies $\theta^\rho > \theta_1$ and thus, by Lemma 2, $\theta^* > \theta_1$. Further, observe that $2\pi_{II} > \pi_{IO} + \pi_{OO}$ implies $\pi_{II} > \pi_{OO}$, because $\pi_{II} = \pi_{OO}$ would imply $\pi_{II} > \pi_{IO}$, which contradicts Assumption 1(iii). Thus, under soft competition, inequality (1.12) holds and the RJV strictly increases gross profit.

If $\theta^\rho \leq \theta^B$, Lemma 2 implies that the RJV spends exactly its budget or less; hence an RJV does not increase R&D expenditure and, as it strictly increases gross profits, it must also strictly increase net profits. If instead $\theta^\rho > \theta^B$, then Lemma 2 implies that $\theta^* = \theta^\rho > \theta_1$. Thus, the RJV strictly increases the probability of innovation. By revealed preference, the RJVs profit must be at least as high as if it had chosen $\theta^* = \theta_1$. Even this choice would lead to higher net profits than R&D competition: First, it saves the R&D costs of duplication; second, for those values of θ where total gross profits under the RJV differ from those under R&D competition ($\theta \in (\theta_2, \theta_1)$), total gross profits under the RJV are strictly higher than under R&D competition, as $2\pi_{II} > \pi_{IO} + \pi_{OO} \geq \pi_{IO} + \pi_{OI}$.

(ii) By Proposition 1, moderate competition and $B > \bar{B}(\rho)$ and $\rho > \overline{\rho}$ imply that the RJV weakly reduces cost and that $\theta^* > \theta_1$. Further, $\rho > \bar{\rho}$ implies that $\pi_{II} > \pi_{OO}$. Otherwise, we would have $\bar{\rho} = \infty$ if $\pi_{II} = \pi_{OO}$, which contradicts $\rho > \bar{\rho}$. Hence, since $2\pi_{II} \geq \pi_{IO} + \pi_{OI}$ under moderate competition, together with $\pi_{II} > \pi_{OO}$ and $\theta^* > \theta_1$, inequality (1.12) holds and the RJV strictly increases gross profits. As it does not increase costs, it also increases net profits.

(iii) Suppose first that $\theta^* > \theta_1$. Rearranging (1.11), the requirement that the expected gross profit difference is strictly positive becomes

$$(\theta^* - \theta_1)(2\pi_{II} - 2\pi_{OO}) > (\theta_1 - \theta_2)(\pi_{IO} + \pi_{OI} - 2\pi_{II})$$

As $2\pi_{II} < \pi_{I0} + \pi_{0I}$ (intense competition) implies $\theta_1 > \theta_2$ and the restriction on Ψ can only hold if $\pi_{II} > \pi_{00}$, we can rearrange again to get

$$\frac{\theta^* - \theta_1}{\theta_1 - \theta_2} > \frac{\pi_{I0} + \pi_{0I} - 2\pi_{II}}{2\pi_{I,I} - 2\pi_{00}} = \Psi.$$

Thus, provided $\theta^* > \theta_1$, $\frac{\theta^* - \theta_1}{\theta_1 - \theta_2} > \Psi$ is equivalent with the requirement that the RJV strictly increases expected gross profits. But $\pi_{I0} + \pi_{0I} - 2\pi_{II} > 0$ implies $\Psi > 0$. Using $\theta_1 - \theta_2 > 0$, $\frac{\theta^* - \theta_1}{\theta_1 - \theta_2} > \Psi$ thus implies $\theta^* > \theta_1 + \Psi(\theta_1 - \theta_2) > \theta_1$.

Further, $2\pi_{II} < \pi_{I0} + \pi_{0I} \leq \pi_{I0} + \pi_{00}$ implies that $\theta_1 > \theta^\rho$ and $\Psi > 0$. Hence, $\frac{\min\{\theta^B, \theta^u\} - \theta_1}{\theta_1 - \theta_2} > \Psi$ implies $\theta^B \geq \theta^\rho$. Therefore, using Lemma 2, we obtain that $\theta^* = \min\{\theta^B, \theta^u\}$, so that the RJV is not spending more than its budget and hence not more than the individual firms. Therefore, the RJV strictly increases gross profit and (as it does not increase costs) net profits.

Proof of Proposition 4

We first show that when $2\pi_{II} - (\pi_{I0} + \pi_{00}) = 0$, the RJV leaves the innovation probability unaffected. To see this, first note that, together with condition (ii), Proposition 1 implies that the innovation probability is not strictly higher in the RJV than under R&D competition. Next, $2\pi_{II} - (\pi_{I0} + \pi_{00}) = 0$ implies that $\theta_1 = \theta^\rho$. As $\theta^* \geq \theta^\rho$ by Lemma 2, we obtain $\theta^* \geq \theta_1$, so that the RJV does not have a negative effect on the innovation probability either. All told, there is no effect of the RJV on the innovation probability.

Next, still assuming that $2\pi_{II} - (\pi_{I0} + \pi_{00}) = 0$, total gross profits are weakly higher in the RJV than under competition for $\theta \in (\theta_2, \theta_1)$ because $2\pi_{II} \geq \pi_{I0} + \pi_{0I}$. As gross profits are the same with and without RJV for the remaining realizations of θ, expected total gross profits are at least weakly higher with the RJV than without. Note that by condition (iii), $\theta_2 > 0$, so that the costs with the RJV are strictly lower than the total costs with R&D competition, with the difference being equal to $\int_0^{\theta_2} C(\theta) d\theta$.

Finally, observe that a ceteris paribus increase from π_{I0} to π'_{I0} does not affect θ_2 nor θ^* but increases θ_1 to some $\theta'_1 > \theta^*$. By continuity, there exists some $\hat{\pi}_{I0}$ such that for all $\pi'_{I0} \in (\pi_{I0}, \hat{\pi}_{I0})$, the change in total gross profits under R&D competition is smaller than $\int_0^{\theta_2} C(\theta)d\theta$. For all such π'_{I0}, the RJV is profitable but it decreases the innovation probability from θ'_1 to θ^*.

1.A.6 Examples

Linear Cournot Competition

We now sketch the details for the Cournot example of Section 1.3.6. Using the notation $\alpha = a - c$, it is straightforward to show that, under the assumption that $\alpha > I$ an equilibrium with positive outputs and profits exists for both firms, so that the innovation is non-drastic. The equilibrium profits are given as $\pi_{I0} = \frac{1}{9}\frac{(\alpha+2I)^2}{b}$, $\pi_{II} = \frac{1}{9}\frac{(\alpha+I)^2}{b}$, $\pi_{00} = \frac{1}{9}\frac{\alpha^2}{b}$, $\pi_{0I} = \frac{1}{9}\frac{(\alpha-I)^2}{b}$. These expressions imply that, whenever $\alpha > I$, Assumption 1 holds, as well as the stricter condition that competition is not soft required by Proposition 1(ii). Next, Corollary 1 follows directly from inserting these profit expressions in the term $\overline{\rho}$. Furthermore, the boundary between intense and moderate competition, given by $2\pi_{II} = \pi_{I0} + \pi_{0I}$, can be calculated as $\alpha = 3I/2$.

Differentiated Price Competition

We now add further details for the case of price competition with inverse demand $p_i = 1 - q_i - bq_j$ for $b \in [0, 1)$. We assume that cost differences are not too large $\left(c_i < \frac{2-b-b^2+bc_j}{2-b^2}\right)$ for $i \in \{1, 2\}, j \neq i$. Then standard calculations show that both equilibrium outputs are positive, with equilibrium profit

$$\pi_i = \frac{\left(2 - b - b^2 - \left(2 - b^2\right) c_i + bc_j\right)^2}{\left(4 - b^2\right)^2 \left(1 - b^2\right)}. \tag{1.13}$$

Inserting appropriate values for c_1 and c_2 gives

$$\pi_{I0} = \frac{\left(2 - b - b^2 - \left(2 - b^2\right)(c - I) + bc\right)^2}{\left(4 - b^2\right)^2 \left(1 - b^2\right)}$$

$$\pi_{II} = \frac{\left(2 - b - b^2 - \left(2 - b^2\right)(c - I) + b\left(c - I\right)\right)^2}{\left(4 - b^2\right)^2 \left(1 - b^2\right)}$$

$$\pi_{00} = \frac{\left(2 - b - b^2 - \left(2 - b^2\right)c + bc\right)^2}{\left(4 - b^2\right)^2 \left(1 - b^2\right)}$$

$$\pi_{0I} = \frac{\left(2 - b - b^2 - \left(2 - b^2\right) c + b\left(c - I\right)\right)^2}{\left(4 - b^2\right)^2 \left(1 - b^2\right)}$$

The requirement that all profits be non-negative $\left(c_i \leq \frac{2-b-b^2+bc_j}{2-b^2}\right)$ is most demanding when $c_i = c$ and $c_j = c - I$, in which case it can be guaranteed by assuming

$$I < \frac{b^2 c - 2c - b + bc - b^2 + 2}{b}$$

In Figure 1.4, we set $c = 0.5$ and hence

$$I < \frac{0.5}{b}\left(2 - b^2 - b\right)$$

The assumptions of the paper can easily be verified in this case. We also find expressions for the regions plotted in Figure 1.4. After some rearrangements, the condition that $\pi_{I0} + \pi_{00} \leq 2\pi_{II}$ becomes

$$\left(4 - 8b - 2b^2 + 4b^3 + b^4\right) I \geq 8c + 12b - 4b^2 c + 6b^3 c + 2b^4 c - 12bc + 4b^2 - 6b^3 - 2b^4 - 8$$

For $c = 0.5$, this simplifies to

$$4I - 8bI - 2b^2 I + 4b^3 I + b^4 I \geq -b^4 - 3b^3 + 2b^2 + 6b - 4$$

The condition $\pi_{I0} + \Pi\left(0, I\right) < 2\pi_{II}$ becomes

$$4I - 4b - 8bI - 3b^2 I + 4b^3 I + b^4 I - 3b^2 + 2b^3 + b^4 + 4 \geq 0.$$

1.A.7 Mergers

In this section, we first provide formal statements of the informal claims in the main text; thereafter, we state and prove the central result comparing mergers and research joint ventures.

Optimal R&D portfolio of Merged Entity

We first describe the investment behavior of the merged entity in a similar way as for the RJV (see Lemma 2).

Lemma 1.A.6. *The merged entity chooses a single cut-off strategy with*

$$\hat{\theta} = \begin{cases} \theta_m^\rho & \text{if} \quad \theta^B < \theta_m^\rho \\ \theta^B & \text{if} \quad \theta^B \in [\theta_m^\rho, \theta_m^u] \\ \theta_m^u & \text{if} \quad \theta^B > \theta_m^u. \end{cases}$$

Proof. The proof is entirely analogous to the proof of Lemma 2. We merely have to replace π_{00} with π_0 and π_{II} with π_I in the profit expressions and adjust the critical values. □

Next, we adapt Proposition 1 to the case of mergers.

Proposition 1.A.1 (Comparison of R&D-Competition and Mergers)**.**

(i) Suppose $\pi_I - \pi_0 > \pi_{I0} - \pi_{II}$. Then the innovation probability is strictly larger after the merger than under R&D competition.

(ii) Suppose $\pi_I - \pi_0 \leq \pi_{I0} - \pi_{II}$. Then:
(a) The innovation probability is strictly larger after the merger than in any equilibrium under competition if and only if $B > \bar{B}(\rho)$ and $\rho > \bar{\rho}_m$.
(b) If the merger strictly increases the innovation probability, then it weakly decreases total R&D spending.

The proof is analagous to the proof of Proposition 1.

Proof of Proposition 5: Comparing Mergers and RJVs

We require an auxiliary result, the proof of which is obvious.

Lemma 1.A.7. $2[\pi_{II} - \pi_{00}] \gtreqless \pi_I - \pi_0 \Leftrightarrow \theta^u \gtreqless \theta_m^u \Leftrightarrow \theta^\rho \gtreqless \theta_m^\rho$.

Next, we prove the two statements of the proposition in turn.

(i) Suppose that $2[\pi_{II} - \pi_{00}] > \pi_I - \pi_0$. By Lemma 1.A.7, $\theta_m^\rho < \theta^\rho$ and $\theta_m^u < \theta^u$. This implies that there are two possible orderings of critical values:

$$\theta_m^\rho < \theta^\rho \leq \theta_m^u < \theta^u \tag{1.14}$$

$$\theta_m^\rho < \theta_m^u < \theta^\rho < \theta^u. \tag{1.15}$$

Suppose first that ordering (1.14) holds. If $\theta^B < \theta^\rho_m$, then the RJV invests in the set $(0, \theta^\rho)$ and the merged firm in $(0, \theta^\rho_m)$. If $\theta^B \in [\theta^\rho_m, \theta^\rho)$, then the RJV invests in the set $(0, \theta^\rho)$ and the merged firm in $(0, \theta^B)$. Hence, since $\theta^\rho > \theta^\rho_m$, it follows that the RJV invests in a larger set than the merged firm whenever $\theta^B < \theta^\rho$. If $\theta^B \in [\theta^\rho, \theta^u_m]$, then both invest in the identical set $(0, \theta^B)$. If $\theta^B \in (\theta^u_m, \theta^u)$, then the RJV invests in the set $(0, \theta^B)$ and the merged firm in $(0, \theta^u_m)$. If $\theta^B \geq \theta^u$, then the RJV invests in the set $(0, \theta^u)$, whereas the merged firm still invests in $(0, \theta^u_m)$. Hence, since $\theta^u > \theta^u_m$, it follows that the RJV invests in a larger set than the merged firm whenever $\theta^B > \theta^u_m$.

Now suppose that ordering (1.15) holds. The analysis for $\theta^B < \theta^\rho_m$ and $\theta^B \geq \theta^u$ is unchanged. If $\theta^B \in [\theta^\rho_m, \theta^u_m)$, then the RJV invests in the set $(0, \theta^\rho)$ and the merged firm in $(0, \theta^B)$. If $\theta^B \in [\theta^u_m, \theta^\rho)$, then the RJV still invests in the set $(0, \theta^\rho)$, and the merged firm invests in $(0, \theta^u_m)$. If $\theta^B \in [\theta^\rho, \theta^u)$, then the RJV invests in the set $(0, \theta^B)$ and the merged firm in $(0, \theta^u_m)$. Hence, whenever ordering (1.15) holds, the RJV invests in a larger set than the merged firm.

Next, suppose that $2[\pi_{II} - \pi_{OO}] = \pi_I - \pi_O$. By Lemma 1.A.7, $\theta^u = \theta^u_m$ and $\theta^\rho = \theta^\rho_m$. If $\theta^B < \theta^\rho = \theta^\rho_m$, then both the RJV and the merged firm invest in $(0, \theta^\rho)$. If $\theta^B \in [\theta^\rho, \theta^u)$, then both invest in the set $(0, \theta^B)$. If $\theta^B \geq \theta^u = \theta^u_m$, then both the RJV and the merged firm invest in the set $(0, \theta^u)$. Hence, for any θ^B both the RJV and the merged firm invest in the same set of research projects.

(ii) Suppose that $2[\pi_{II} - \pi_{OO}] < \pi_I - \pi_O$. By Lemma 1.A.7, $\theta^u < \theta^u_m$ and $\theta^\rho < \theta^\rho_m$. This implies that there are two possible orderings of critical values:

$$\theta^\rho < \theta^\rho_m \leq \theta^u < \theta^u_m \tag{1.16}$$

$$\theta^\rho < \theta^u < \theta^\rho_m < \theta^u_m. \tag{1.17}$$

Suppose first that ordering (1.16) holds. If $\theta^B < \theta^\rho$, then the merged firm invests in the set $(0, \theta^\rho_m)$ and the RJV in $(0, \theta^\rho)$. If $\theta^B \in [\theta^\rho, \theta^\rho_m)$, then the merged firm invests in the set $(0, \theta^\rho_m)$ and the RJV in $(0, \theta^B)$. Hence, if $\theta^B < \theta^\rho_m$ the merged firm invests in a larger set than the RJV. If $\theta^B \in [\theta^\rho_m, \theta^u]$, then both invest in the identical set $(0, \theta^B)$. If $\theta^B \in (\theta^u, \theta^u_m)$, then the merged firm invests in the set $(0, \theta^B)$ and the RJV in $(0, \theta^u)$. If $\theta^B \geq \theta^u_m$, then the merged firm invests in the set $(0, \theta^u_m)$ and the RJV still in $(0, \theta^u)$. Hence, whenever $\theta^B > \theta^u$ the merged firm invests in a larger set than the RJV.

Now suppose that ordering (1.17) holds and consider again different values that θ^B can take. The analysis for $\theta^B < \theta^\rho$ and $\theta^B \geq \theta^u_m$ is unchanged. If $\theta^B \in [\theta^\rho, \theta^u)$, then the merged firm invests in the set $(0, \theta^\rho_m)$ and the RJV in $(0, \theta^B)$. If $\theta^B \in [\theta^u, \theta^\rho_m)$, then the merged firm invests in the set $(0, \theta^\rho_m)$ and the RJV in $(0, \theta^u)$. If $\theta^B \in [\theta^\rho_m, \theta^u_m)$, then the merged firm invests in the set $(0, \theta^B)$ and the RJV in $(0, \theta^u)$. Hence, whenever the ordering (1.17) holds, the merged firm invests in a larger set than the RJV. $\square$

1.A.8 Spillovers

No Financial Constraints

As a benchmark, we now consider a model without financial constraints. Instead, we allow for spillovers. Specifically, if a firm has invested successfully in a project and the rival has not, then with probability $\sigma \in [0, 1]$ the rival will obtain access to the innovation. The expected total payoff of firm i, given the strategy of firm j is then

$$\begin{aligned}\mathbb{E}\Pi_i(r_i,r_j) = &\int_0^1 (1 - r_j(\theta)) \left[r_i(\theta)((1-\sigma)\pi_{I0} + \sigma\pi_{II}) + (1 - r_i(\theta))\pi_{00}\right] d\theta + \\ &\int_0^1 r_j(\theta) \left[(r_i(\theta) + \sigma(1 - r_i(\theta)))\pi_{II} + (1-\sigma)(1 - r_i(\theta))\pi_{0I}\right] d\theta - \\ &\int_0^1 r_i(\theta)C(\theta)d\theta.\end{aligned}$$

Compared to the expected total payoff with financial constraints, firms do not have additional costs from borrowing. Moreover, there is now the possibility that a firm obtains the innovation without innovating itself. The equilibrium characterization for R&D competition closely follows the previous analysis.

We first implicitly define critical projects similar to those defined previously.

$$C(\theta_1^{nc}) = (1-\sigma)\pi_{I0} + \sigma\pi_{II} - \pi_{00}$$
$$C(\theta_2^{nc}) = (1-\sigma)[\pi_{II} - \pi_{0I}]$$

The intuition for θ_1^{nc} and θ_2^{nc} is analogous to the one for θ_1 and θ_2, taking into account different payoffs due to potential spillovers. It is straightforward to show that $\theta_1^{nc} \geq \theta_2^{nc}$.

Lemma 1.A.8 (Investment strategies under competition with spillovers).
(i) The research competition game has multiple equilibria. A profile of double cut-off strategies (r_i^*, r_j^*) is an equilibrium if it satisfies (a) $\theta_L = \theta_2^{nc}$ and $\theta_H = \theta_1^{nc}$ and (b) for each $\theta \in (\theta_2^{nc}, \theta_1^{nc})$ either:

$$r_i^*(\theta) = 1 \text{ and } r_j^*(\theta) = 0 \text{ or}$$
$$r_i^*(\theta) = 0 \text{ and } r_j^*(\theta) = 1.$$

(ii) No other equilibria of the research-competition game exist.

Next, we consider the case with RJVs. The analysis is simpler than in the case with financial constraints. The increase in joint profit from a successful innovation is $2\pi_{II} - 2\pi_{00}$. Hence, the RJV invests in all projects up to θ^u, and it does not invest in the remaining ones. We can now prove Proposition 6.

Proof of Proposition 6 When competition is soft, the argument follows as in the case without spillovers (without relying on Assumption 2). When competition is not soft, we need to show that the condition in the proposition is equivalent with the requirement that $\theta_1^{nc} < \theta^u$. This follows from simple rearrangements:

$$\sigma > 1 - \frac{\pi_{II} - \pi_{00}}{\pi_{I0} - \pi_{II}}$$
$$\sigma(\pi_{I0} - \pi_{II}) > \pi_{I0} - 2\pi_{II} + \pi_{00}$$
$$2\pi_{II} - 2\pi_{00} > (1-\sigma)\pi_{I0} + \sigma\pi_{II} - \pi_{00}$$

$$C(\theta^u) > C(\theta_1^{nc}).$$

Financial constraints

We now augment the model with spillovers with financial constraints. The analysis of Sections 1.2 and 1.3 carries over directly if we replace the function π with $\tilde{\pi}$ defined as follows

$$\begin{aligned}
\tilde{\pi}_{00} &\equiv \pi_{00} \\
\tilde{\pi}_{II} &\equiv \pi_{II} \\
\tilde{\pi}_{I0} &\equiv (1-\sigma)\pi_{I0} + \sigma\pi_{II} \\
\tilde{\pi}_{0I} &\equiv (1-\sigma)\pi_{0I} + \sigma\pi_{II}
\end{aligned}$$

Replacing π with $\tilde{\pi}$, we obtain new expressions for expected profits, $\mathbb{E}\tilde{\Pi}_i(r_i, r_j)$, for critical values $\tilde{\theta}_1, \tilde{\theta}_2$, etc. We replace Assumption 2 with

Assumption 5. *$B < \int_0^{\tilde{\theta}_2} C(\theta)d\theta$.*

It is straightforward to show that $\tilde{\theta}_1 \geq \tilde{\theta}_2$. Moreover, if $\theta_1 > 0$, then $\theta_1 > \tilde{\theta}_1$ for all $\sigma > 0$. Hence, spillover reduces the incentives to invest because rivals could also benefit from the innovation. The following result follows directly from replacing π with $\tilde{\pi}$ in Lemma 1 and then inserting the above definitions for $\tilde{\pi}$.

Lemma 1.A.9 (Investment strategies under competition with spillovers and financial constraints)**.**
(i) The research competition game has multiple equilibria. A profile of double-cut off strategies (r_i^*, r_j^*) is an equilibrium if it satisfies (a) $\theta_L = \tilde{\theta}_2$ and $\theta_H = \tilde{\theta}_1$ and (b) for each $\theta \in (\tilde{\theta}_2, \tilde{\theta}_1)$ either:

$$\begin{aligned}
&r_i^*(\theta) = 1 \text{ and } r_j^*(\theta) = 0 \text{ or} \\
&r_i^*(\theta) = 0 \text{ and } r_j^*(\theta) = 1.
\end{aligned}$$

(ii) No other pure-strategy equilibria of the research-competition game exist.

Suppose now that the two firms form an RJV. Since the firms will share a successful innovation, spillovers do not affect innovation behavior under cooperative R&D. Therefore, an RJV still has the critical projects θ^ρ and θ^u and invests according to Lemma 2.

For the comparison between RJV and R&D competition, we replace θ_1 and π in the definitions of $\bar{\rho}$ and $\bar{B}(\rho)$ with $\tilde{\theta}_1$ and $\tilde{\pi}$ to obtain:

$$\tilde{B}(\rho) = \frac{\int_0^{\tilde{\theta}_1} C(\theta) d\theta}{2}$$

$$\tilde{\rho} = \begin{cases} \dfrac{\tilde{\pi}_{I0} - \tilde{\pi}_{II} - (\tilde{\pi}_{II} - \tilde{\pi}_{00})}{2(\tilde{\pi}_{II} - \tilde{\pi}_{00})}, & \text{for } \tilde{\pi}_{II} > \tilde{\pi}_{00} \\ \infty, & \text{for } \tilde{\pi}_{II} = \tilde{\pi}_{00}. \end{cases}$$

It is straightforward to show that $\bar{\rho} > \tilde{\rho}$ and $\bar{B}(\rho) > \tilde{B}(\rho)$. Replacing π with $\tilde{\pi}$ in Proposition 1 and then inserting the values for $\pi_{t_i t_j}$ into the definitions of $\tilde{\pi}_{t_i, t_j}$ immediately shows under which circumstances an RJV increases innovation with spillovers.

Proposition 1.A.2 (Comparison of competition and RJV with spillovers)**.**

(i) Suppose $2\pi_{II} - 2\pi_{00} > \pi_{I0} - \pi_{00} - \sigma(\pi_{I0} - \pi_{II})$. *Then the innovation probability is strictly larger under the RJV than under R&D competition.*

(ii) Suppose $2\pi_{II} - 2\pi_{00} < \pi_{I0} - \pi_{00} - \sigma(\pi_{I0} - \pi_{II})$. *Then:*
(a) The innovation probability is strictly larger under the RJV than under competition if and only if $B > \tilde{B}(\rho)$ *and* $\rho > \tilde{\rho}$.
(b) If the formation of the RJV strictly increases the innovation probability, then it weakly decreases total R&D spending.

Moreover, the conditions on the budget and interest rate that an RJV increases the probability of innovation are weaker with higher spillovers (see Section 1.A.8).

Proof of Proposition 7

First, we note two auxiliary results which are analogous to Lemmas 1.A.2 and 1.A.3, replacing π with $\tilde{\pi}$ and θ_1 with $\tilde{\theta}_1$.

Lemma 1.A.10. $\pi_{II} > (1-\sigma)(\pi_{I0} - \pi_{II}) + \pi_{00} \Leftrightarrow \theta^{\rho} > \tilde{\theta}_1$.

Lemma 1.A.11. $\rho > \tilde{\rho} \Leftrightarrow \theta^{u} > \tilde{\theta}_1$.

Next, we provide a useful monotonicity result:

Lemma 1.A.12. *$\tilde{\theta}_1$ is a weakly decreasing function of σ and ρ. $\tilde{B}$ and $\tilde{\rho}$ are weakly decreasing in σ.*

Proof. Suppose $\sigma' \geq \sigma$ and $\rho' \geq \rho$. Then

$$\frac{\pi_{I0} - \sigma(\pi_{I0} - \pi_{II}) - \pi_{00}}{1+\rho} \geq \frac{\pi_{I0} - \sigma'(\pi_{I0} - \pi_{II}) - \pi_{00}}{1+\rho'}$$
$$\frac{(1-\sigma)\pi_{I0} + \sigma\pi_{II} - \pi_{00}}{1+\rho} \geq \frac{(1-\sigma')\pi_{I0} + \sigma'\pi_{II} - \pi_{00}}{1+\rho'}$$

where the inequality holds since $\pi_{I0} \geq \pi_{II}$. The first result immediately follows. Next, since $C(\theta)$ is a strictly increasing function, $\tilde{B} = \frac{\int_0^{\tilde{\theta}_1} C(\theta)d\theta}{2}$ must also be weakly decreasing in σ. The interest rate cut-off value $\tilde{\rho}$ is decreasing in σ, since

$$\frac{\partial\tilde{\rho}}{\partial\sigma} = \frac{\pi_{II} - \pi_{I0}}{2(\pi_{II} - \pi_{00})} \leq 0,$$

if $\pi_{II} - \pi_{00} > 0$ and zero otherwise. □

To prove Proposition 7, suppose first that we have weak competition in the sense that $\pi_{II} > (1-\sigma)(\pi_{I0} - \pi_{II}) + \pi_{00}$. By Lemma 1.A.10, this implies $\theta^{\rho} > \tilde{\theta}_1$. By Lemma 2, the probability that the RJV innovates is at least θ^{ρ}. By Lemma 1.A.9, the probability of innovation under R&D competition is $\tilde{\theta}_1$, where this expression is decreasing in σ by Lemma 1.A.12. Now, suppose $\pi_{II} \leq (1-\sigma)(\pi_{I0} - \pi_{II}) + \pi_{00}$. By Proposition 1.A.2(ii)(a), a strictly larger innovation probability under the RJV implies $B > \tilde{B}(\rho)$ and $\rho > \tilde{\rho}$. Further, arguing as in the proof of Lemma 1.A.5, if the innovation probability is strictly larger under the RJV than under R&D competition, then $\theta^B > \theta^\rho$. If $\theta^B \in (\theta^\rho, \theta^u)$, the RJV invests in all projects in the set $(0, \theta^B)$ and discovers the innovation with probability θ^B. Since $B > \tilde{B}(\rho) = \int_0^{\tilde{\theta}_1} C(\theta)d\theta/2$, we have $\int_0^{\theta^B} C(\theta)d\theta > \int_0^{\tilde{\theta}_1} C(\theta)d\theta$, which implies $\theta^B > \tilde{\theta}_1$. Without the RJV, in any equilibrium, the firms invest in projects in the set $(0, \tilde{\theta}_1)$ and the innovation is discovered with probability $\tilde{\theta}_1$. As $\tilde{\theta}_1$ is weakly decreasing in ρ and weakly decreasing in σ by Lemma 1.A.12, it immediately follows that, if the probability of innovation is strictly larger under the RJV for any σ and ρ, then this is also true for any $\sigma' \geq \sigma$ and $\rho' \geq \rho$. If $\theta^B \geq \theta^u$, then the RJV invests in all projects in the set $(0, \theta^u)$ and discovers the innovation with probability θ^u. By Lemma 1.A.11, $\rho > \tilde{\rho}$ implies $\theta^u > \tilde{\theta}_1$. It immediately follows that, if the probability of innovation is strictly larger under the RJV for any σ and ρ, then this is also true for $\sigma' \geq \sigma$ and $\rho' \geq \rho$. □

Licensing

We now add some more details to the licensing model sketched in Section 1.4.3, where a successful innovator chooses a two-part tariff licensing contract (L, η) at which the unsuccessful innovator can use the innovation. The buyer accepts any contract that yields at least the outside option of π_{0I}. In equilibrium, the innovator extracts all rents and sets a fixed fee L such that the unsuccessful firm earns π_{0I}. Therefore, the single innovator receives the total market surplus net of the outside option, $2\pi_{II} + \Delta - \pi_{0I}$. We spell out the profit function π^L as

$$\begin{aligned}
&\pi^L_{00} \equiv \pi_{00} \\
&\pi^L_{II} \equiv \pi_{II} \\
&\pi^L_{I0} = \max\{\pi_{I0}, 2\pi_{II} + \Delta - \pi_{0I}\} \\
&\pi^L_{0I} = \pi_{0I}.
\end{aligned}$$

Replacing π with π^L, we obtain a new expression for expected profits,

$\mathbb{E}\Pi_i^L(r_i, r_j)$, and critical values θ_1^L, θ_2^L, etc. We maintain Assumption 2. It is straightforward to show that $\theta_1^L \geq \theta_1$ and $\theta_2^L = \theta_2$. The next result follows directly from replacing π with π^L in Lemma 1 and then inserting the above definitions for π^L.

Lemma 1.A.13 (Investment strategies under competition with licensing)**.** *Suppose that* $2\pi_{II} + \Delta - \pi_{0I} \geq \pi_{I0}$. *Then:*
(i) The research competition game with licensing has multiple equilibria. A profile of double-cut off strategies (r_i^*, r_j^*) *is an equilibrium if it satisfies (a)* $\theta_L = \theta_2$ *and* $\theta_H = \theta_1^L$ *and (b) for each* $\theta \in (\theta_2, \theta_1^L)$ *either:*

$$r_i^*(\theta) = 1 \textit{ and } r_j^*(\theta) = 0 \textit{ or}$$
$$r_i^*(\theta) = 0 \textit{ and } r_j^*(\theta) = 1.$$

(ii) No other equilibria of the research-competition game exist.

The analysis of the RJV is unchanged; it invests according to Lemma 2. Define the budget threshold $\bar{B}^L(\rho)$ and the interest threshold $\bar{\rho}^L$ as

$$\bar{\rho}^L = \begin{cases} \dfrac{\pi_{00} + \Delta - \pi_{0I}}{2(\pi_{II} - \pi_{00})}, & \text{for } \pi_{II} > \pi_{00} \\ \infty, & \text{for } \pi_{II} = \pi_{00}. \end{cases}$$

$$\bar{B}^L(\rho) = \frac{\int_0^{\theta_1^L} C(\theta) d\theta}{2}$$

With this notation in place, it is straightforward to see how Proposition 8 directly follows by reformulation of Proposition 1 with π replaced by π^L.

1.A.9 Multiple firms

Industry-wide RJV

We extend the model to three ex-ante symmetric firms ($i \in \{1, 2, 3\}$). The product market profits of firm i are now given in the reduced form $\pi_{t_i t_j t_k}$ for $j, k \neq i$, $j \neq k$. We suppose profits are symmetric: $\pi_{0I0} = \pi_{00I}$ and

$\pi_{II0} = \pi_{I0I}$. That is, only the number of successful rivals matters. We adjust the assumptions on the product market profits accordingly.

Assumption 6 (Regularity of market profit functions)**.**

(i) Profits are non-negative: $\pi_{t_i t_j t_k} \geq 0$ *for all* t_i, t_j *and* t_k.

(ii) Innovation increases profits: $\pi_{III} \geq \pi_{000}$.

(iii) Competitor innovation reduces profits: $\pi_{000} \geq \pi_{0I0} \geq \pi_{0II}$.

(iv) Competitor innovations reduce the value of own innovations: $\pi_{I00} - \pi_{000} \geq \pi_{II0} - \pi_{0I0} \geq \pi_{III} - \pi_{0II}$.

We obtain cut-off values $\theta_3 \leq \theta_2 \leq \theta_1$ from

$$\begin{aligned}(1+\rho)C(\theta_1) &= \pi_{I00} - \pi_{000}\\ (1+\rho)C(\theta_2) &= \pi_{II0} - \pi_{0I0}\\ (1+\rho)C(\theta_3) &= \pi_{III} - \pi_{0II}.\end{aligned}$$

After appropriately modifying Assumption 2, we find that all equilibria have a triple cut-off structure with all firms investing in $[0, \theta_3)$, two firms in (θ_3, θ_2), one firm in (θ_2, θ_1) and no firm investing in $(\theta_1, 1)$

Now we suppose that all three firms form an RJV. Let θ^B be defined as the solution to $\int_0^{\theta^B} C(\theta)d\theta = 3B$ if $\int_0^1 C(\theta)d\theta > 3B$ and $\theta^B = 1$ otherwise. Next, let θ^u and θ^ρ be the solutions to the following equations

$$\begin{aligned}(1+\rho)C(\theta^\rho) &= 3[\pi_{III} - \pi_{000}]\\ C(\theta^u) &= 3[\pi_{III} - \pi_{000}].\end{aligned}$$

Using these cut-off values, the RJV follows a cut-off strategy as in Lemma 2. Defining soft competition by the requirement that $3\pi_{III} > \pi_{I00} + 2\pi_{000}$ and adjusting the budget cut-off value $\bar{B}$ and the interest rate cut-off value $\bar{\rho}$ appropriately, we finally obtain conditions under which an RJV increases the probability of innovation, which are analogous to those in Proposition 1.

Multiple RJVs

Next, we extend the model to four ex-ante symmetric firms ($i \in \{1, 2, 3, 4\}$). We write the product market profits of firm i facing competitors j, k and l as $\pi_{t_i t_j t_k t_\ell}$. As in the previous subsection, we assume that profits depend only on the own technology and the number of competitors with the new technology, not on their identity. Further, we impose the regularity conditions that profits are non-negative, weakly increasing in own innovation and that the positive effect of own innovation decreases in the number of competitors with access to the new technology.

We again adjust Assumption 2 so that firms want to borrow externally under R&D competition. Unsurprisingly, it turns out that, under R&D competition these equilibria have four cut-off values $\theta_4 \leq \theta_3 \leq \theta_2 \leq \theta_1$, defined in the by now familiar way.

Now we suppose that two RJVs are formed, each consisting of two firms. Thus, instead of four firms, we have two competing RJVs $\{v_1, v_2\}$, each with budget $2B$. Let θ^B be defined as the solution to $\int_0^{\theta^B} C(\theta)d\theta = 2B$ if $\int_0^1 C(\theta)d\theta > 2B$ and $\theta^B = 1$ otherwise. To find the cutoff-values, consider the equations

$$\begin{aligned}
(1+\rho)C(\theta_1^\rho) &= 2[\pi_{IIOO} - \pi_{OOOO}] \\
(1+\rho)C(\theta_2^\rho) &= 2[\pi_{IIII} - \pi_{OOII}] \\
C(\theta_1^u) &= 2[\pi_{IIOO} - \pi_{OOOO}] \\
C(\theta_2^u) &= 2[\pi_{IIII} - \pi_{OOII}].
\end{aligned}$$

The interpretation is the same as with one RJV. We restrict our analysis to the case in which the budget of an RJV is sufficiently large such that no RJV borrows in equilibrium.

Assumption 7. *$2B > \int_0^{\theta_1^\rho} C(\theta)d\theta$.*

The assumptions imply $\theta^B > \theta_1^\rho \geq \theta_2^\rho$. How θ^B relates to the two values $\theta_2^u \leq \theta_1^u$ will determine the optimal portfolio of an RJV. The research competition game turns out to have multiple equilibria with double cut-offs (and no other equilibria).

The proof follows a similar structure as in Lemma 1, but we have to distinguish between the three cases $\theta^B < \theta_2^u$, $\theta^B \in [\theta_2^u, \theta_1^u)$ and $\theta_1^u \leq \theta^B$. Further, Assumption 7 implies $(1+\rho)C(\theta) > 2[\pi_{II00} - \pi_{0000}]$ for any $\theta > \theta^B$. Thus, it is never optimal to invest more than the available budget for an RJV. Defining soft competition by the requirement that $\pi_{II00} - \pi_{0000} > \pi_{I000} - \pi_{II00}$ and adjusting the budget cut-off value $\bar{B}$ and the interest rate cut-off value $\bar{\rho}$ appropriately, we finally obtain conditions under which the formation of two RJVs increases the probability of innovation, which are analogous to those in Proposition 1.

1.A.10 Sources for RJV Examples

In the Introduction, we mentioned several actual research joint ventures. More information about these ventures can be found at the following links, which are listed in the order in which the RJV appeared in text. All links were last accessed on June 28, 2022 and are archived on https://web.archive.org.

- https://www.cbo.gov/publication/57126
- https://www.bdo.co.uk/en-gb/news/2021/top-20-global-carmakers-spend-another-71-7bn-as-electric-vehicle-rollout-gathers-pace
- https://group-media.mercedes-benz.com/marsMediaSite/de/instance/ko.xhtml?oid=42917172
- https://www.saftbatteries.com/media-resources/press-releases/psa-a-total-automotive-cells-company
- https://www.bp.com/en/global/corporate/news-and-insights/press-releases/paving-the-way-for-sustainable-mobility-bp-bmw-daimler-announce-bp-third-shareholder-of-dcs.html
- https://www.reuters.com/business/autos-transportation/renault-nissan-mitsubishi-alliance-say-deepen-cooperations-ev-production-2022-01-27/

- https://www.basf.com/global/de/media/news-releases /2021/05/p-21-215.html
- https://www.volvogroup.com/en/news-and-media/news /2020/apr/news-3640568.html
- https://www.washingtonpost.com/technology/2022/05/18 /solid-state-batteries-electric-vehicles-race/
- https://www.forbes.com/sites/greggardner/2020/02/03 /toyota-and-panasonic-launch-joint-ev-battery-venture/
- https://www.volkswagen-newsroom.com/en/press-releases /enel-x-and-volkswagen-team-up-for-electric-mobility-in-italy-7315
- https://www.media.stellantis.com/em-en/fca-archive/press /fiat-chrysler-automobiles-and-engie-eps-plan-to-join-forces -in-a-jv-creating-a-leading-company-in-the-e-mobility-sector

Chapter 2

Drip Pricing and After-sales with Behavioral Consumers

joint with Christian Zihlmann

2.1 Introduction

Airlines charge fees for optional services such as seat allocation, while online sellers offer additional products during the purchasing process. Banks charge late and overdraft fees, and electronic suppliers promote extended warranties. Decoupling a product or service into a base good and an extra item has become a common business practice, possibly because it increases the consumer's willingness to pay ("WTP") for the add-on (Morwitz, Greenleaf, and Johnson, 1998). Not surprisingly, add-on selling and drip pricing[1] have drawn the attention of competition authorities and raised concerns among policymakers.[2] In October 2022, the Biden-Harris administration announced an initiative to tackle potential issues arising from such practices (The White House, 2022).

We analyze a model of add-on selling in which some consumers relate the add-on price to the price of the base product. For them, spending \$10 on shipping and handling feels less significant when the product price is \$100, as opposed to \$20. In the former case, we may perceive that opting for home delivery won't put a dent in the wallet and accept it. In the latter case, we may decline the extra service and instead pick up the product in-store. Experiments provide causal evidence that consumers are more likely to buy a queue-skipping voucher because the price of the ski pass was higher (Erat and Bhaskaran, 2012) and that they will put more effort into redeeming a \$5 discount for a \$25 radio than for a \$500 TV (Bushong, Rabin, and Schwartzstein, 2021; Thaler, 1980) because the relative gain is larger for the cheap product.

We integrate such consumer behavior into a model of drip and add-on pricing, which features both classical and behavioral consumers. We propose that behavioral consumers' WTP for the add-on increases as the base good

[1] Drip pricing is the sequential presentation of prices and is defined as "[...] a pricing technique in which firms advertise only part of a product's price and reveal other charges later as the customer goes through the buying process. The additional charges can be mandatory charges [...] or fees for optional upgrades and add-ons" (Federal Trade Commission, 2012). In this work, we focus on the latter case of optional upgrades and add-ons.

[2] See, for example, the British Competition Market Authority (2022) and the US Federal Trade Commission (2022).

price rises, a behavior that we microfound with well-documented phenomena such as salience, relative thinking, proportional thinking, mental accounting, and anchoring. The WTP of classical consumers is independent of the base good.

Our findings show that firms face an incentive to increase the prices of their *base goods* in the presence of behavioral consumers. This creates an endogenous base good price floor as not all add-on revenues are redistributed. Our analysis further reveals that the effect of behavioral consumers on the surplus of classical consumers is non-monotonic, which distinguishes our insights from previous studies in the literature (Ellison, 2005; Gabaix and Laibson, 2006). The presence of behavioral consumers can either benefit or harm classical consumers, depending on the proportion of behavioral consumers in the market.

This non-monotonicity has significant implications beyond the distribution of the surplus between the two consumer types. Specifically, policy interventions such as consumer education may have unintended consequences as they can either increase or decrease consumer surplus, depending on the pre-existing equilibrium in the economy. This makes it exceptionally challenging to predict the effects of policy measures in practice. Our analysis underscores the complexity of regulating an after-sales market with heterogeneous consumers.

In our market, two firms compete in prices with horizontally differentiated base products. Consumers first purchase a base good from a seller. Subsequently, they are presented with the seller's offer for an optional ancillary product, which can be rejected at zero costs. The seller enjoys monopoly power in the add-on market. Although there could be multiple sources of market power in the add-on market,[3] this assumption will be relaxed later to show that our results generalize to sequential buying settings with competition in the after-sales market.

Consumers are heterogeneous. Some consumers behave according to traditional theory. For those classical consumers, the WTP for the add-

[3]See, for example, Diamond (1971); Ellison (2005); Gabaix and Laibson (2006); Shapiro (1994).

on is naturally independent of the price of the previous base good purchase. In contrast, behavioral consumers compare the add-on to the more expensive base good, resulting in temporarily increased demand for the add-on. Our reduced-form model accommodates this behavior without relying on a specific psychological mechanism, and we show that the previously mentioned behavioral micro-foundations all satisfy our model's assumptions. Therefore, behavioral consumers' WTP for the add-on increases with the price of the base good, leading to an add-on WTP that is strictly greater than that of classical consumers.

Firms face a mixed population of consumers and must decide on their pricing strategy. They know the distribution of consumer types but cannot identify an individual's type. Firms can potentially exploit behavioral consumers by increasing the add-on price to their higher WTP. While classical consumers will not buy the add-on anymore, firms can extract a higher mark-up from behavioral consumers. This strategy is referred to as the *exploiting strategy*. Alternatively, firms can choose not to adapt their pricing strategy, resulting in the add-on price being equal to the WTP of the classical consumers. In this case, all base good buyers accept the additional offer, the *non-exploiting strategy*.

The equilibrium strategy is determined by the proportion of behavioral consumers in the market. With a low share, the exploiting strategy is not optimal as the loss from classical consumers not buying the add-on exceeds the extra revenue from selling it at a higher price to behavioral consumers. The add-on is priced as if all consumers were classical, resulting in the same outcome as in our benchmark economy, which consists of classical consumers only.

Exploiting behavioral consumers' higher WTP in the aftermarket is optimal when the behavioral type becomes sufficiently frequent. Crucially, pursuing this strategy also affects the optimal price of the base good: due to the behavioral mechanism, increasing the base good price allows firms to set an even higher add-on price. Simultaneously, however, demand in the base good market decreases, and with it, also in the after-sales market. The two opposing effects arising from this trade-off determine the optimal prices and the distributional effects between the two types of consumers. Indeed, the optimal base good price may be either higher or lower than the benchmark price, depending on the proportion of behavioral consumers.

We demonstrate that there exist (unique) equilibria in which classical consumers are worse off. When the share of behavioral consumers is intermediate, such that it is optimal to exploit, their presence creates a negative externality on classical consumers because base good prices are higher than in the benchmark. Yet, for a large proportion of behavioral consumers, the effect is reversed and they cross-subsidize classical consumers who benefit from a reduced base good price. When the share is low, then firms price like in the benchmark economy. Thus, only a few behavioral consumers do not affect the market outcome and the surplus of classical consumers.

We find similar results in a setting with a monopolist in the base good market. In general, the non-monotonicity result holds as long as firms have some market power. Only perfect competition in the base good market provides complete protection for the classical consumer. Due to the (perfectly) competitive pressure, firms cannot increase the price of the base good, and the mark-ups from the add-on market must be completely passed on to the base good market. As a result, classical consumers are better off, benefiting from the cross-subsidization of behavioral consumers. Thus, not only imperfect competition, but also perfect competition may raise important distributional concerns.

Recently, in the US airline industry, legislation was proposed to state all prices upfront, including those of optional add-ons such as seat allocations (Department of Transportation, 2022). We analyze the effect of such legislation, which, in essence, is an education policy that aims to reduce the share of behavioral consumers. Our findings suggest that disclosing prices upfront can either enhance or reduce total consumer surplus, depending on the effectiveness of the intervention and the ex-ante equilibrium of the market. Thus, education can counter-intuitively harm both types of consumers.

Similarly, a government that enacts an exogenous price floor on the base good to prevent loss leading, often also referred to predatory pricing, can actually harm both consumers. This is the case if the floor is set above the endogenous price floor created by behavioral consumers: the imposed price floor makes the base good more expensive. As a consequence, firms can take even greater advantage of the behavioral consumer, while foregoing some of the competitive pressure to redistribute the mark-up from the after-sales market to the base good market.

2.1.1 Literature Review

Our work is related to the model proposed by Ellison (2005), which assumes that some consumers have an exogenously higher WTP for an add-on product than others. As such, our mechanism shares some similarities with Ellison's model, as firms in both cases must decide whether to offer the add-on to all consumers or only to those with a higher WTP. A critical difference in our model is that the heterogeneous WTP arises endogenously from consumers' behavior, which can be microfounded by a set of common behavioral mechanisms.[4] This endogenous heterogeneity in the WTP for the add-on leads to a distinct prediction: firms face an incentive to increase the price of the base good. This, in turn, results in the finding that behavioral consumers have non-monotonic effects on classical consumers. This non-monotonicity is similar to Armstrong (2015)'s results derived from a model in which consumers differ in whether they consider the add-on price at the time of the base good purchase or not. Related is also Inderst and Obradovits (2023)'s model of drip pricing in which consumers have context-dependent preferences, resulting in competition to increase the welfare loss due to a distortion in product choice.

This paper also contributes to the recent debate around drip pricing in economics (Kosfeld and Schüwer, 2016), marketing science (see Ahmetoglu, Furnham, and Fagan, 2014, for a review) and antitrust (see Greenleaf,

[4]Although our model explains heterogeneous valuations of add-ons in an endogenous way, it does not exclude the existence of exogenous differences in valuations. Our framework can be extended to account for situations where consumers are heterogeneous due to exogenous circumstances, such as different marginal utilities of income.

Johnson, Morwitz, and Shalev, 2016, for a review). A decoupled good consisting of a base and add-on product can increase demand because consumers underweight the add-on price (Brown, Hossain, and Morgan, 2010; Ellison and Ellison, 2009; Hossain and Morgan, 2006; Santana, Dallas, and Morwitz, 2020). In a large field experiment on StubHub.com, Blake, Moshary, Sweeney, and Tadelis (2021) find that drip pricing increases demand in quantity and quality (see also Dertwinkel-Kalt, Köster, and Sutter, 2020), and drip pricing has been shown to reduce consumer surplus in experimental markets (Huck and Wallace, 2015; Rasch, Thöne, and Wenzel, 2020). Moreover, there is empirical evidence that add-on purchases are more frequent if base good prices are higher (Erat and Bhaskaran, 2012; Xia and Monroe, 2004). Karle, Kerzenmacher, Schumacher, and Verboven (2022) provide experimental evidence suggesting that consumers tend to search less when product prices are high and behave, due to relative thinking, as if they were less price sensitive. This finding is closely related to our assumption that consumers have a higher WTP for add-ons when the price of the base product is higher.

Further, this paper relates to the literature on pricing in multi-good settings and loss leading (see Armstrong and Vickers, 2012, for a review). In these models, typically, firms enjoy ex-post monopoly power over the add-on, allowing them to extract high margins from those after-sales products (see, for example, Holton, 1957). However, perfect competition forces firms to redistribute those rents to the base good, which must be sold as a loss-leader to attract consumers ex-ante (Shapiro, 1994). Loss-leading is often seen as a predatory practice that exploits consumers and reduces welfare (Chen and Rey, 2012). For this reason, the issue has gauged the interest of researchers and antitrust agencies alike. For example, 22 U.S. states prohibit the sale of goods below costs, and loss-leading is banned in several countries in the European Union.[5]

In our model, however, a law that enacts an exogenous price floor on the base good reduces the surplus of the remaining consumers. Thus,

[5] See `https://www.aeaweb.org/research/loss-leading-bans-retail-competition`.

banning loss-leading may actually be detrimental for consumers. This result contributes to recent evidence that points towards the potential negative effects of such bans due to other reasons, such as for example a smaller product choice (Johnson, 2017).

Moreover, our insights contribute to the literature on behavioral industrial organization (see Heidhues and Kőszegi, 2018, for a review), particularly to the literature investigating heterogeneous consumer populations (Heidhues and Kőszegi, 2017). In Gabaix and Laibson (2006), some consumers anticipate the add-on, while others are myopic and (need to) purchase the add-on. The authors find a monotonic effect: Classical consumers are always better off when behavioral consumers enter the economy. Related is also Rasch et al. (2020), who theoretically and experimentally investigated drip pricing with mandatory add-on purchases. However, in our model, there are no surprise charges: consumers voluntarily purchase the add-on or can reject the seller's offer at no cost. Likewise, our article relates to Michel (2017), where consumers overestimate the value of optional warranties due to underestimating the costs of returning faulty products. But our model is distinct in that it considers consumers who differ in whether they evaluate the price of the add-on in reference to the price of the base good, a behavioral foundation that has not yet been introduced in after-sales market models, and is applicable to a variety of add-ons beyond warranties.

Finally, our analysis shows that the presence of behavioral consumers can jeopardize the surplus of classical consumers. This result, together with the insights of Ellison (2005) and Armstrong and Chen (2009), refines insights from the previous literature that usually finds the presence of behavioral consumers does not adversely affect classical consumers, and may even benefit them—the infamous reverse Robin Hood effect (e.g., Gabaix and Laibson, 2006).[6] Usually, the behavioral consumer subsidizes the classical consumer, and thus, the discussion has centered on whether to protect behavioral consumers from their own mistakes. Our findings raise the question of whether one should protect the perfectly acting classical consumer from the

[6] The issue also received considerable attention in the popular press, see for example `https://thehill.com/opinion/finance/580513-reverse-robin-hood-is-real`.

mistakes and biases of others.

This chapter proceeds as follows. Section 2.2 defines the model set-up. Section 2.3 provides the equilibrium analysis. Section 2.4 analyzes policy implications. Consumer behavior is microfounded in Section 2.5 and Section 2.6 provides further results. Section 2.7 concludes. All proofs are presented in appendices.

2.2 Model

We consider a market with products that feature add-on components. After purchasing a base product, the consumer is subsequently confronted with the offer for an ancillary product (or service).[7] Formally, we suppose that two firms $j \in \{1, 2\}$ compete in prices with differentiated base goods, which are imperfect substitutes. Each firm offers a base good at price $p_{1,j}$. There is a continuum of consumers. Firm j faces a weakly concave demand function $D_j(p_{1,j}, p_{1,-j})$, which is twice continuously differentiable, strictly decreasing in its own price and $\lim_{p_{1,j}\to\infty} D_j(\cdot) = 0$. We suppose that the base good demand is (i) supermodular, (ii) the own-price elasticity is stronger than the cross-price elasticity, and (iii) satisfies $\left|\frac{\partial D_j^2(\cdot)}{\partial p_{1,j}^2}\right| \geq \frac{\partial D_j^2(\cdot)}{\partial p_{1,j}\partial p_{1,-j}}$.[8] The last assumption implies that the decrease (increase) of demand is higher when only one firm increases (decreases) prices than when both change prices.[9] To ease notation, we will suppress the firm index j when not necessary. Thus, a single subscript indicates whether it is the price of the base good or add-on.

Once a consumer purchased the base good(s), firms offer one unit of an additional good (or service) per base good sold at price $p_{2,j}$. The add-on demand for firm j is thus bounded from above by $D_j(p_{1,j}, p_{1,-j})$. Consumers are locked-in in the aftermarket, which implies monopolistic power for firms.

[7]Our model also applies to tentative purchases of the base product, assuming that consumers search little once tentatively committed to a base good, as demonstrated by Rasch et al. (2020).

[8]These assumptions are merely used for traceability of asymmetric strategies and are not needed when focusing on symmetric equilibria only.

[9]For example, the linear demand function derived in Singh and Vives (1984) satisfies these assumptions. In Section 2.6.2, we show that our results also hold for unit demand Ã la Hotelling (1929). In Section 2.3.6, we consider the simpler case of a monopolist, which yields similar results with much less structure on the demand function.

For simplicity, we suppose that add-ons are homogeneous across firms and marginal costs of production for both goods are normalized to zero. The WTP for the add-on of consumer i is given in reduced form by the expression

$$W(v_2, \tilde{\Delta}), \text{ where } \tilde{\Delta} = \beta_i \Delta(p_2, p_1)$$

$W(v_2, \tilde{\Delta})$ is strictly increasing in both arguments, non-negative, weakly concave and twice continuously differentiable. All consumers receive a positive gross consumption utility of v_2 for the add-on purchase. Additionally, some consumers may compare the price of the add-on with the price of the base good, leading to a temporarily increased add-on demand. Our model introduces this behavior in a reduced form with $\tilde{\Delta} = \beta_i \Delta(p_2, p_1)$, which is strictly decreasing in p_2 and strictly increasing in the reference price p_1, where $\Delta(p_2, p_1)$ captures the (relative) difference in prices with $\Delta(p_2, p_1) = 0$ when $p_2 = p_1$. In Section 2.5, we formally show that such behavior can be micro-founded by salience, relative thinking, proportional thinking, mental accounting or anchoring and adjustment.

The parameter β_i captures the strength of the behavioral mechanism, where $\beta_i = 0$ characterizes a classical consumer who is not subject to any of the above-discussed behaviors. The classical consumer's WTP for the add-on is independent of the base good price and constant, $W(v_2, 0) = W(v_2)$.

A $\beta_i > 0$ characterizes a behavioral consumer who puts the add-on price p_2 in relation to the price of the base good p_1. The argument $\Delta(p_2, p_1)$ captures this behavioral mechanism. It follows that behavioral consumers have a higher WTP for the add-on than classical consumers when $p_1 > p_2$.[10] In the following, we focus on the case of a more expensive base good than the add-on, such that $\Delta > 0$ in any equilibrium.[11] In Appendix 2.C.7, we

[10]Note that the mechanism could also work through a price rather than WTP distortion. That is, behavioral consumers may misperceive the price of the add-on and perceive it as cheaper than it actually is. This affects the incentive constraint to buy the add-on similarly to a WTP distortion and thus, does not change our analysis. Further, we do not specify a utility function for behavioral consumers. For the moment, we remain agnostic about whether the behavioral mechanism increases the utility of behavioral consumers or not since our main results do not depend on a specification. That is whether behavioral consumers receive a utility of $W(v_2, \tilde{\Delta})$ or $W(v_2)$ when consuming the add-on. Proposition 2.3.2 deals with the welfare of behavioral consumers. The proof shows that the result is independent of the welfare specification. We will define a specific utility function in Section 2.4 when studying policy implications.

[11]We rule out the corner solution $p_1 = p_2$, which implies $\Delta = 0$ and thus, $W(v_2) = W(v_2, \tilde{\Delta})$. Therefore, we focus on interior solutions and consider only equilibria with $p_1 > p_2$.

investigate the case in which the base good costs less than the add-on.

We analyze an economy that potentially consists of both types, classical and behavioral consumers, i.e., $\beta_i \in \{0, \beta\}$ with $i = \{c, b\}$ and $\beta \in (0, 1]$.[12] The share of behavioral consumers in the market is denoted with $\alpha \in [0, 1]$. Firms know the distribution of the types but cannot identify an individual's type. It follows that firms cannot price discriminate and need to offer the same prices p_1 and p_2 to all consumers. The timing of the game is as follows:

- Period 0: Firms choose the prices p_1 and p_2 simultaneously.
- Period 1: Consumers observe the base good price p_1, choose a seller, and buy the base good(s).
- Period 2: Each firm offers an add-on to its base good consumers. Consumers observe the add-on offer and either accept or reject it.

We assume that the add-on does not affect consumer choice in the base good market. They choose a firm only because of the surplus provided by the base good. This assumption is reasonable in a number of settings. For example, the add-on price may be truly unobservable at the time of the base good purchase: many firms reveal prices of add-ons only after a (tentative) base good purchase, a practice known as drip pricing (Competition Market Authority, 2022). Firms may not need to commit to the add-on price ex ante or firms may not advertise and shroud add-on prices (see Gabaix and Laibson, 2006; Gamp, 2015; Spiegler, 2006, 2016).[13] Finally, add-on prices may be too expensive to learn ex ante before arriving at a point of sale (Ellison, 2005; Heidhues, Johnen, and Kőszegi, 2021).[14]

[12]Since β reflects the strength of the behavioral mechanism of an individual, it is restricted to the unit interval.

[13]When prices are unobservable, in line with Gabaix and Laibson (2006), we suppose consumers form Bayesian posteriors about the add-on with the beliefs that firms set monopolistic add-on prices since they are profit-maximizing. When firms play symmetric strategies, the rational expectations are identical across firms and, thus, do not affect consumer's choice problem. For asymmetric strategies, firms set different prices for the base good, implying different prices for the add-on as one firm will exploit behavioral consumers' higher WTP. However, either firm could set a higher base good price. Thus, we suppose consumers cannot infer a firm's strategy from observing $p_{1,j}$ and expect with equal probability that firm j exploits when observing $p_{1,j} \neq p_{1,-j}$.

[14]See Heidhues and Kőszegi (2018); Spiegler (2011) for a review of behavioral models in add-on pricing.

2.3 Equilibrium Analysis

We solve the game for Nash equilibria in pure strategies.

2.3.1 Aftermarket

In period 2, after the purchase of the base good, consumers with WTP $W(v_2, \tilde{\Delta})$ can buy an add-on at price p_2. Classical consumers ($\beta = 0$) buy the add-on when $W(v_2) \geq p_2$. Behavioral consumers ($\beta \in (0, 1]$) buy when $W(v_2, \tilde{\Delta}) \geq p_2$. Therefore, the demand for the add-on of firm j is given by

$$Q_j(p_{2,j}, D_j(p_{1,j}, p_{1,-j})) = \begin{cases} D_j(p_{1,j}, p_{1,-j}) & \text{if } p_{2,j} \leq W(v_2), \\ \alpha D_j(p_{1,j}, p_{1,-j}) & \text{if } W(v_2) < p_{2,j} \leq W(v_2, \tilde{\Delta}), \\ 0 & \text{if } p_{2,j} > W(v_2, \tilde{\Delta}). \end{cases}$$

Observe that the add-on demand also depends indirectly on p_1, because only base good buyers can purchase the add-on.

2.3.2 Firms' Problem

The profit function of firm j is given by

$$\pi_j(p_{1,j}, p_{1,-j}, p_{2,j}) = p_{1,j} D_j(p_{1,j}, p_{1,-j}) + p_{2,j} Q_j(p_{2,j}, D_j(\cdot)). \tag{2.1}$$

Firms have monopolistic power in the aftermarket and extract the entire rent of one of the two consumer groups by making them indifferent. Two possible prices emerge in equilibrium, implicitly defined by $p_2^* \in \{W(v_2), W(v_2, \tilde{\Delta})\}$. If $p_2^* = W(v_2)$, firms do not exploit behavioral consumers and all consumers accept the additional offer. We refer to this as the *non-exploiting strategy.* If $p_2^* = W(v_2, \tilde{\Delta})$, however, firms do exploit the behavioral consumer and price the add-on at behavioral consumers' WTP. Consequently, classical consumers do not accept the add-on offer. We refer to this as the *exploiting strategy.*

Selecting one strategy determines $p_{2,j}$ and $Q_j(p_{2,j}, D_j(\cdot))$ in Equation (2.1). Given a chosen strategy, firm j maximizes its profits by choosing the base good price $p_{1,j}$, which yields the implicitly defined best response functions for any $p_{1,-j}$. Depending on the rival's actions, we obtain either (i) symmetric non-exploiting prices p_1^n and profits π^n, (ii) symmetric exploiting prices $p_1^e(\alpha)$ and profits π^e or (iii) an asymmetric outcome, where the non-exploiting firm sets $\tilde{p}_1^n(\alpha)$ and gets $\tilde{\pi}^n$, and the exploiting firm sets $\tilde{p}_1^e(\alpha)$ and receives $\tilde{\pi}^e$, where

$$\begin{aligned}
\pi^n &= \pi(p_1^n, p_1^n, W(v_2)) \\
\pi^e &= \pi(p_1^e(\alpha), p_1^e(\alpha), W(v_2, \tilde{\Delta})) \\
\tilde{\pi}^n &= \pi\left(\tilde{p}_1^n(\alpha), \tilde{p}_1^e(\alpha), W(v_2)\right) \\
\tilde{\pi}^e &= \pi(\tilde{p}_1^e(\alpha), \tilde{p}_1^n(\alpha), W(v_2, \tilde{\Delta})).
\end{aligned}$$

The full derivation of all prices and profits are characterized in Appendix 2.A. Lemma 2.A.1 shows that the exploiting profits π^e and $\tilde{\pi}^e$ are strictly increasing in α. This is because behavioral consumers—who can be exploited by firms—become more frequent when α increases. The symmetric non-exploiting profit π^n is independent of the share of behavioral consumers because firms set the add-on price to the WTP of the classical consumer, implying that all consumers accept the add-on offer. The asymmetric non-exploiting profit $\tilde{\pi}^n$ is either increasing or decreasing in α.[15]

2.3.3 Equilibrium

The formal equilibrium derivation is provided in Appendix 2.A.4. The share of behavioral consumers α crucially determines which equilibrium arises. When behavioral consumers are particularly frequent, then both firms exploit in equilibrium by choosing $p_1^* = p_1^e(\alpha)$, $p_2^* = W(v_2, \tilde{\Delta})$. When the share of behavioral consumers is low, then neither firm exploits and both set $p_1^* = p_1^n$, $p_2^* = W(v_2)$ in equilibrium. For a wide range of α, the *symmetric non-exploiting equilibrium* and the *symmetric exploiting equilibrium*, respectively, are unique. Only for an intermediate share of behavioral consumers multiple equilibria exists. Either the best response implies to do the same as the

[15]See Appendix 2.A for details and explanation.

rival and both, the symmetric non-exploiting and symmetric exploiting equilibrium exist, or the best response is to do the opposite and multiple asymmetric equilibria arise.

The equilibrium structure is intuitive. Firms face a trade-off between a higher aftermarket demand versus a higher mark-up.[16] When the share of behavioral consumers is low, the demand effect dominates. The income from selling a high-priced add-on to only a few behavioral consumers cannot compensate for the demand loss arising from classical consumers who decline the additional offer. Accordingly, firms do not exploit and set the add-on price at the classical consumers' WTP $W(v_2)$. When the share of behavioral consumers is large, both firms exploit behavioral consumers by setting $p_2^* = W(v_2, \tilde{\Delta})$. The demand loss from not serving classical consumers in the aftermarket is (over)compensated by the higher add-on mark-up because sufficiently many behavioral consumers are in the population.

2.3.4 The base good price

In classical aftersales models, the only incentive firms face is to lower the base good price in order to lock in more consumers (Diamond, 1971). An important consequence of introducing consumers following our behavioral pattern is that it gives firms a countervailing incentive to increase the base good price. A more expensive base good increases behavioral consumers' WTP for the add-on. This allows firms to extract a higher mark-up in the add-on market, increasing the value of the aftermarket. Yet, a higher base good price leads to a lower demand in the base good market, which also implies less demand for the add-on. Hence, firms that exploit in the aftermarket face a trade-off when setting the optimal base good price $p_1^e(\alpha)$ or $\tilde{p}_1^e(\alpha)$, and consequently also $\tilde{p}_1^n(\alpha)$, since base good prices are strategic complements. This trade-off is captured by the relationship between the two

[16]This is reminiscent of Ellison (2005), but as we show in our analysis, firms face an incentive to increase the base good price in our model. Proposition 2.3.1 shows that there exist exploiting equilibria, in which the base good is more expensive than in a non-exploiting equilibrium.

semi-elasticities

$$\epsilon_D = \frac{-\partial D(\cdot)/\partial p_{1,j}}{D(\cdot)} \quad \text{and} \quad \epsilon_W = \frac{\partial W(v_2, \tilde{\Delta})/\partial p_{1,j}}{W(v_2, \tilde{\Delta})}.$$

The base good demand semi-elasticity, ϵ_D, denotes the demand effect of a price change in the base good market and, thus, the amount of consumers in the add-on market. The second semi-elasticity, ϵ_W, captures how strongly the add-on WTP of behavioral consumers reacts to a change in the reference price. Depending on which effect dominates, the optimal base good price is either a decreasing or increasing function in the share of behavioral consumers α. When $\epsilon_D > \epsilon_W$, the demand effect is stronger and the optimal prices $p_1^e(\alpha)$, $\tilde{p}_1^e(\alpha)$ and $\tilde{p}_1^n(\alpha)$ are decreasing in α. In this case, it is profitable to attract and lock-in more consumers by lowering the base good price. In contrast, when $\epsilon_D < \epsilon_W$, the optimal base good prices are increasing in the share of behavioral consumers. This is the case when the base good demand is relatively inelastic and a price change has little effect on the sold quantity of base goods. We show in Lemma 2.A.2 in Appendix 2.A that the order of the semi-elasticities, ϵ_D and ϵ_W, is monotonic in α and thus, also the price functions.[17]

Crucially, this implies that the base good price in the symmetric exploiting equilibrium and asymmetric equilibrium depends on the share of behavioral consumers in the population. As a consequence, their presence affects classical consumers in the base good market.

To analyze this effect, we consider a benchmark economy consisting of classical consumers only ($\alpha = 0$). The benchmark base good price is given by p_1^b. In equilibrium, firms sell the add-on to all consumers at $p_2 = W(v_2)$. Thus, the benchmark outcome is identical to the symmetric non-exploiting equilibrium, and we obtain $p_1^b = p_1^n$, which implies that in any symmetric non-exploiting equilibrium, firms price as if there were only classical consumers. It is immediate to see that a low share of behavioral consumers has no effect on the market outcome and the surplus of classical consumers.

[17]That is for specific functions $D(\cdot)$ and $W(\cdot)$, it is either $\epsilon_D \geq \epsilon_W$ for all α or $\epsilon_D \leq \epsilon_W$ for all α.

Whether the base good in a symmetric exploiting or asymmetric equilibrium is cheaper or more expensive than in the benchmark depends crucially on α and whether ϵ_D or ϵ_W is stronger. We define the implicit price threshold

$$\bar{\alpha}_p = \begin{cases} \dfrac{W(v_2)}{W(v_2,\tilde{\Delta}) + \dfrac{\frac{\partial W(v_2,\tilde{\Delta})}{\partial p_1} D(p_1^*, p_1^*)}{\frac{\partial D(p_1^*,p_1^*)}{\partial p_1}}}, & \text{for } \epsilon_D \neq \epsilon_W \\ \infty, & \text{for } \epsilon_D = \epsilon_W, \end{cases}$$

where $p_1^* \in \{p_1^n, p_1^e(\bar{\alpha}_p), \tilde{p}_1^n(\bar{\alpha}_p), \tilde{p}_1^e(\bar{\alpha}_p)\}$. When $\alpha = \bar{\alpha}_p$, then the base good costs the same in any equilibrium, $p_1^n = p_1^b = p_1^e(\bar{\alpha}_p) = \tilde{p}_1^n(\bar{\alpha}_p) = \tilde{p}_1^e(\bar{\alpha}_p)$. Observe that, since $D(p_{1,j}, p_{1,-j})$ is decreasing in $p_{1,j}$, the denominator of $\bar{\alpha}_p$ is not necessarily positive, but depends on the relationship of the semi-elasticities. The price threshold $\bar{\alpha}_p$ is positive when $\epsilon_D > \epsilon_W$ and negative when $\epsilon_D < \epsilon_W$.[18] Lemma 2.3.1 captures when the base good is cheaper or more expensive than in the benchmark economy.

Lemma 2.3.1.

(i) Suppose $\epsilon_D > \epsilon_W$. If $\alpha \in (\min\{\bar{\alpha}, \hat{\alpha}\}, \bar{\alpha}_p)$, then the base good is more expensive in any symmetric exploiting or asymmetric equilibrium than in the benchmark. If $\alpha > \bar{\alpha}_p$, then the base good is cheaper in any symmetric exploiting equilibrium.

(ii) Suppose $\epsilon_D < \epsilon_W$. The base good is more expensive in any symmetric exploiting or asymmetric equilibrium than in the benchmark.

For $\epsilon_D > \epsilon_W$, the price functions $p_1^e(\alpha)$, $\tilde{p}_1^n(\alpha)$, and $\tilde{p}_1^e(\alpha)$ are all decreasing in α. Hence, when the share of behavioral consumers is sufficiently low ($\alpha < \bar{\alpha}_p$), then the base good, offered by an exploiting firm (and by the non-exploiting firm in the asymmetric case), is more expensive compared to the benchmark case. Note that symmetric exploiting and asymmetric equilibria exist only when $\alpha > \min\{\bar{\alpha}, \hat{\alpha}\}$.[19] Otherwise, for a large share $\alpha > \bar{\alpha}_p$,

[18]The inequality $W(v_2,\tilde{\Delta}) > \frac{\partial W(v_2,\tilde{\Delta})}{\partial p_1} \frac{D(p_1^*,p_1^*)}{\frac{\partial D(p_1^*,p_1^*)}{\partial p_1}}$ can be rearranged to $\epsilon_D > \epsilon_W$.

[19]We define the profit thresholds $\bar{\alpha}$ and $\hat{\alpha}$ in Appendix 2.A.4. They are necessary to formalize the equilibrium characterization.

the base good is cheaper in any symmetric exploiting equilibrium.[20] For $\epsilon_D < \epsilon_W$, the price functions are increasing in α, and the price threshold $\bar{\alpha}_p$ is negative. Hence, for any share of behavioral consumers, the base good of an exploiting firm (and of the non-exploiting firm in the asymmetric case) is more expensive than in the benchmark economy.

2.3.5 The Surplus of Classical Consumers

We turn now to the central part of our analysis and main result. Combining the results from Lemma 2.3.1 and the equilibrium characterization (Lemma 2.A.3) identifies that the presence of behavioral consumers has non-monotonic effects on classical consumers. Importantly, when $\epsilon_D > \epsilon_W$, then the price threshold is always larger than the profit thresholds. That is $\bar{\alpha}_p > \max\{\bar{\alpha}, \hat{\alpha}\}$. Hence, there exists an interval, in which α is such that a symmetric exploiting or asymmetric equilibrium exists and the base good price in these equilibria is larger than in the benchmark economy.[21] When $\epsilon_D < \epsilon_W$, then the base good is always more expensive in a symmetric exploiting or asymmetric equilibrium. Proposition 2.3.1 states the conditions when classical consumers benefit or are harmed by the presence of behavioral consumers.

Proposition 2.3.1 (The effect on the surplus of classical consumers)**.**

(a) Behavioral consumers do not affect the market in any symmetric non-exploiting equilibrium.

(b) Suppose $\epsilon_D > \epsilon_W$. Then the presence of behavioral consumers: (i) harms classical consumers in any symmetric exploiting equilibrium if $\alpha < \bar{\alpha}_p$ and benefits them otherwise, (ii) harms classical consumers in any asymmetric equilibrium.

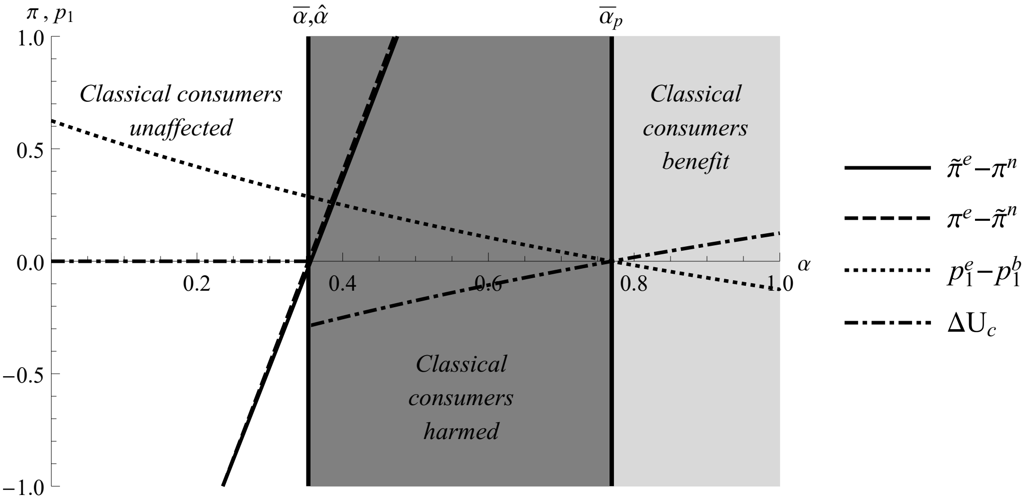
(c) Suppose $\epsilon_D < \epsilon_W$. Then the presence of behavioral consumers harms classical consumers in any symmetric exploiting or asymmetric

[20] Asymmetric equilibria do not exist when $\alpha > \bar{\alpha}_p$.

[21] When $\alpha > \max\{\bar{\alpha}, \hat{\alpha}\}$, then the unique symmetric exploiting equilibrium exists. When $\hat{\alpha} < \alpha < \bar{\alpha}$, then the asymmetric equilibria exist. When $\bar{\alpha} < \alpha < \hat{\alpha}$, then the multiple symmetric equilibria exist.

equilibrium.

Figure 2.1: Proposition 1 (a) with Symmetric Prices.



The figure depicts Proposition 2.3.1 (a) with $\epsilon_D > \epsilon_W$ and $\bar{\alpha} < \hat{\alpha}$, using $W(v_2, \tilde{\Delta}) = v_2\beta(1 + p_1 - p_2)$ with $\beta = 1$. The base good demand function is adopted from Singh and Vives (1984): $D_j(p_{1,j}, p_{1,-j}) = \frac{v_1}{1+d} - \frac{p_{1,j}}{1-d^2} + \frac{dp_{1,-j}}{1-d^2}$. The parameter specifications are $v_1 = 9$, $v_2 = 1$, $d = 0.4$ and marginal costs $c = 2.6$.
Notation:
π^n, π^e: symmetric non-exploiting and exploiting profits;
$\tilde{\pi}^n, \tilde{\pi}^e$: profits in asymmetric outcomes;
p_1^e, p_1^b: symmetric exploiting and benchmark economy prices;
$\hat{\alpha}$: profit threshold such that $\pi^n = \tilde{\pi}^e$; $\bar{\alpha}$: profit threshold such that $\tilde{\pi}^n = \pi^e$;
ΔU_c: classical consumer surplus change.

When the share of behavioral consumers is low, like in case (a), there is no effect on classical consumers because no firm exploits, which leads to the same outcome as in the benchmark economy. This is independent of how elastic demand is. The presence of behavioral consumers, however, may harm classical consumers once there are sufficient behavioral buyers in the economy. For a relatively elastic base good demand ($\epsilon_D > \epsilon_W$), this is the case when the share of behavioral consumers is intermediate, such that exploiting is optimal, but the base good is more expensive than in the benchmark economy. Then, the presence of behavioral consumers harms classical consumers because they have to pay more for the desired base good than when all consumers would be classical and not subject to a behavioral mechanism. Though, since the optimal base good price is a decreasing function for $\epsilon_D > \epsilon_W$, classical consumers are better off when the share is large ($\alpha > \bar{\alpha}_p$) because the base good is cheaper than in the benchmark economy. These three different outcomes of Proposition 2.3.1 (a) and (b) are depicted in Figure 2.1. For a relatively inelastic demand for the base good ($\epsilon_D < \epsilon_W$), classical consumers are always harmed when at least one firm exploits in equilibrium because then the base good is always more expensive than in the benchmark.

Our finding implies that consumers who behave perfectly according to classical economic theory, do not always benefit from behavioral or "naive" consumers. In our model, classical consumers may be harmed by the presence of behavioral consumers. This has major policy implications, which we will discuss in more detail in Section 2.4.

So far, we have focused only on the surplus of classical consumers. Let us now consider total consumer surplus. The results in Proposition 2.3.2 state how exploitation by firms affects the surplus of behavioral consumers and total consumer welfare. The variable $CS^E(\alpha)$, which is increasing in α when $\epsilon_D > \epsilon_W$ and decreasing otherwise, captures the consumer surplus in the base good market when at least one firm exploits in equilibrium. Similarly, CS^{NE} denotes the consumer surplus in the base good market when no firm exploits.

Proposition 2.3.2 (Consumer surplus)**.**

(a) Suppose $\epsilon_D > \epsilon_W$

  (i) Behavioral consumers are worse off when exploited.

  (ii) Suppose at least one firm exploits in equilibrium. Then, the total consumer surplus is strictly lower when $\alpha D(\cdot)[W(v_2, \tilde{\Delta}) - W(v_2)] > CS^E(\alpha) - CS^{NE}$*. The condition is always satisfied for* $\alpha \leq \bar{\alpha}_p$*.*

(b) Suppose $\epsilon_D < \epsilon_W$*. Then behavioral consumer and total consumer surplus is strictly lower when at least one firm exploits in equilibrium.*

First, behavioral consumers always have a lower surplus in an exploiting equilibrium than in a symmetric non-exploiting equilibrium. This is independent of whether the behavioral mechanism increases the received add-on utility or not. Thus, behavioral consumers are worse off, when firms exploit them. The effect on total consumer surplus is mixed when $\epsilon_D > \epsilon_W$. Exploitation can decrease the base good price below the price in a symmetric non-exploiting (or benchmark) equilibrium. This leads to an increased demand for the base good, but it also implies that more behavioral consumers are exploited in the aftermarket. This trade-off is captured by the inequality $\alpha D(\cdot)[W(v_2, \tilde{\Delta}) - W(v_2)] > CS^E(\alpha) - CS^{NE}$.[22] On the left-hand side, we have the surplus effect of behavioral consumers buying the add-on, where $W(v_2, \tilde{\Delta}) - W(v_2)$ is the price difference for the add-on, which is strictly positive. The fraction α of the population pays a strictly larger price in an equilibrium with exploitation. This means, depending on the welfare specification, behavioral consumers either pay a price above their true WTP or lose some surplus from the add-on compared to the non-exploiting case. The right-hand side displays the change in consumer surplus in the base good market, which can be positive or negative. When $\alpha \leq \bar{\alpha}_p$, by Lemma 2.3.1, the base good is more expensive when at least one firm exploits compared to a symmetric non-exploiting (or benchmark) equilibrium. This implies less demand, and it must be that $CS^E(\alpha) \leq CS^{NE}$. Hence, when classical consumers are harmed according to Proposition 2.3.1, then the total consumer surplus is always lower because of exploitation. For $\alpha > \bar{\alpha}_p$, the effect of exploitation is unclear on consumer surplus. In this case, the base good is cheaper, which increases the surplus in the base good market $CS^E(\alpha) > CS^{NE}$. Therefore, it depends on whether the positive effect in the base good market dominates the negative impact in the after-sales market. When $\epsilon_D < \epsilon_W$, all consumers are worse off when firms exploit, and thus, total consumer surplus is also lower.

[22] Note that we need to adjust the condition slightly for asymmetric equilibria since not all behavioral consumers are exploited. However, since asymmetric equilibria exist only for $\alpha < \bar{\alpha}_p$, the adjusted condition is always satisfied.

2.3.6 Monopoly and Perfect Competition

In this section, we discuss the outcomes of the two extreme cases of competition, monopoly and perfect competition. The findings in Proposition 2.3.1 and 2.3.2 are robust to a monopoly setting, while perfect competition eliminates the harmful effect on classical consumers. We defer the details of the formal analysis and results to Appendix 2.C.1. With only one firm in the base good market, the analysis is identical to the case of imperfect competition, but we need to impose fewer assumptions on the demand function.[23] The findings are similar to the two-firm case, except that asymmetric equilibria do not exist. This is non-trivial since the cross-subsidization result usually vanishes with monopolistic competition in the existing literature (Heidhues and Kőszegi, 2018).

Under perfect price competition, that is when base goods are perfect substitutes, then only the positive effect of behavioral consumers on classical consumers survives. The intuitive reason is that due to competitive pressure, firms cannot increase the base good price above the benchmark level. Otherwise, firms would face zero demand. Thus, when firms exploit in equilibrium, then the base good price must be strictly lower, which benefits classical consumers. Hence, under perfect competition, classical consumers can never be harmed by the presence of behavioral consumers, which resembles the findings of Gabaix and Laibson (2006).

2.4 Policy Implications

We apply our main results stated in Proposition 2.3.1 and 2.3.2, and analyze how different policies affect consumer welfare. First, we consider a policy that educates behavioral consumers and thus reduces their frequency in the population, such as for example revealing all prices (separately) up-front. Second, we analyze the effect of a price floor regulation on the base good,

[23] When base goods are perfectly differentiated, then each firm is a monopolist in its respective base good market. Compared to the imperfect competition case, we need much less structure on the base good demand function. We simply impose that $D(p_1)$ is strictly decreasing, twice continuously differentiable, $\lim_{p_1 \to \infty} D(p_1) = 0$ and satisfies $D(p_1)D''(p_1) < 2D'(p_1)^2$, which, for instance, holds for log-concave but also CES demand functions.

which is a common tool used by policymakers to prevent loss-leading and predatory pricing. Lastly, we discuss the impact of a price cap on the add-on.

2.4.1 Educating behavioral consumers

The Department of Transportation (2022) proposed a rule to require airlines to reveal the full price of a ticket up-front, including ancillary services such as checked baggage. This leads to so-called partitioned pricing, where all prices are shown up-front but still separately for the base good and ancillary services. This is in contrast to drip pricing, where prices are presented sequentially, and all-inclusive pricing, where consumers see one total price. Studies in marketing science give us indications on the effect of partitioned pricing: Compared to drip pricing, partitioned pricing might reduce consumer demand for the add-on, however, consumer demand may still be higher with partitioned prices compared to all-inclusive prices (Morwitz et al., 1998; Robbert and Roth, 2014). Therefore, we assume that such a policy intervention, as proposed by the Department of Transportation (2022), reduces the frequency of behavioral consumers, because it educates some consumers who become classical consumers after the intervention.

Suppose there exists an instrument or technology for policymakers to reduce the share of behavioral consumers in the population, and that the ex-ante share of behavioral consumers is such that the unique symmetric exploiting equilibrium exists, $\alpha > \max\{\bar{\alpha}, \hat{\alpha}\}$.[24] To characterize the policy impact, we need to distinguish between effective and ineffective instruments. An *effective* policy leads to a sufficiently large reduction of behavioral consumers such that after the intervention, exploiting is not optimal anymore for either firm. Hence, the measure results in the unique symmetric non-exploiting equilibrium. In contrast, an *ineffective* policy reduces the share of behavioral consumers only by a bit such that both firms still exploit in equilibrium.

[24] The intuition for ex-ante asymmetric equilibria is similar.

Definition 4.

(i) An effective education policy reduces α sufficiently large such that the symmetric non-exploiting equilibrium emerges ex-post.

(ii) An ineffective education policy reduces α only by a little such that both firms still exploit ex-post.

Further, we need to distinguish whether classical consumers benefit or are harmed by the presence of behavioral consumers prior to the policy implementation. According to Proposition 2.3.1, we call any equilibrium in which classical consumers benefit, *beneficial equilibrium.* Otherwise, when classical consumers are harmed by the presence of behavioral consumers, we have a *harmful equilibrium.* To study the surplus effect on consumers who are educated, we need to specify a utility function for behavioral consumers. We suppose that the behavioral mechanism does not yield utility and define $U_b = v_1 - p_1 + W(v_2) - p_2$, where v_1 is the gross utility generated by the base good. We demonstrate in the proof of Proposition 2.4.1 that when the other case applies, and add-on utility is increased, $\tilde{U}_b = v_1 - p_1 + W(v_2, \tilde{\Delta}) - p_2$, then the policy effect on educated consumers is identical to that of classical consumers. For consumers who remain behavioral ex-post, the policy impact is independent of the welfare specification.

Proposition 2.4.1 (Education)**.**

(a) Suppose $\epsilon_D > \epsilon_W$.

(i) Any ineffective education policy makes behavioral and classical consumers worse off. Educated consumers benefit if and only if $W(v_2, \tilde{\Delta}) - W(v_2) > p_1^e(\alpha'') - p_1^e(\alpha')$.

(ii) Any effective education policy benefits behavioral and educated consumers. Classical consumers benefit when the ex-ante equilibrium was harmful and are worse off if the ex-ante equilibrium was beneficial.

(b) When $\epsilon_D < \epsilon_W$, all consumers benefit from any policy, increasing total consumer surplus.

When the base good demand is relatively elastic as in Proposition 2.4.1 (a), then the optimal base good price under symmetric exploiting $p_1^e(\alpha)$ is decreasing in α. Since an ineffective policy reduces the number of behavioral consumers only by a little, the intervention does not change firms' behavior but strictly increases $p_1^e(\alpha)$. Hence, firms still exploit in equilibrium, and consumers must pay more for the base good, which makes any uneducated consumer clearly worse off. In addition, the add-on also becomes more expensive. Educated consumers do not buy the overpriced add-on anymore ex-post, which affects their surplus positively by $W(v_2, \tilde{\Delta}) - W(v_2)$. But whether the overall policy effect for them is positive depends on how strongly the price increase in the base good market is, $p_1^e(\alpha'') - p_1^e(\alpha')$, where α' is the share of behavioral consumers before the intervention and α'' denotes the share ex-post. This implies that even consumers who get educated could be worse off.

In contrast, an effective policy affects a firm's behavior, and there is no exploitation ex-post. The intervention results in the symmetric non-exploiting equilibrium. This always benefits behavioral and educated consumers, even if the base good becomes more expensive. This is directly related to Proposition 2.3.2. The benefit of not being exploited dominates any negative effects in the base good market.

For classical consumers, it depends on whether the ex-ante equilibrium was harmful or beneficial. In the former case, classical consumers benefit from an effective policy because the base good is cheaper ex-post. In the latter case, they are hurt, because an effective policy prevents the cross-subsidization from behavioral consumers, which results in a more expensive base good ex-post.[25]

When the base good demand is relatively inelastic as in Proposition 2.4.1 (b), then the optimal base good price under symmetric exploiting $p_1^e(\alpha)$ is increasing in α. This implies that any reduction in the share of behavioral consumers lowers the base good price, which benefits all consumers. Hence, any policy is beneficial and increases consumer surplus.

[25]The add-on surplus for classical consumers is still zero ex-post.

The results in Proposition 2.4.1 provide important insights for policymakers. Not every education policy improves the welfare of consumers. On the contrary, they may even hurt consumers. When demand effects are relatively strong, then policymakers need to be careful with imposing regulations and interventions as they may worsen the situation for consumers. Even the educated consumer can be worse off, when the base good price reacts strongly to an ineffective policy.

2.4.2 Price Floor

Loss-leading is a controversial practice that raises concerns over anti-competitive effects. For that reason, predatory pricing is banned in many US States and some European countries.[26] Policymakers impose a price floor on goods by prohibiting pricing below costs with the aim of protecting consumers. The literature finds mixed results on the effectiveness of this policy (e.g., Chen and Rey, 2012; Johnson, 2017). In our model, a binding price floor yields negative effects for most consumers, while it is ambiguous whether consumers who could benefit really do so.

We impose a price floor on the base good that does not affect the benchmark economy, $\underline{p}_1 \leq p_1^b$. This is binding only in an exploiting equilibrium when the base good is cheaper than in the benchmark economy, which requires $\epsilon_D > \epsilon_W$.[27] This is the case when the share of behavioral consumers is sufficiently large, $\alpha > \bar{\alpha}_p$. Then, firms want to decrease base good prices to attract more behavioral consumers who are willing to buy the overpriced add-on and, thus, can be exploited.

Proposition 2.4.2 (Price floor)**.** *Suppose $\epsilon_D > \epsilon_W$ and $\underline{p}_1 \leq p_1^b$. A binding price floor $\underline{p}_1$ (i) increases the add-on price and (ii) reduces the base good demand. Classical and remaining behavioral consumers in the market are strictly worse off by a binding regulation. The effect on behavioral consumers who left the market is ambiguous.*

[26] See for example https://www.aeaweb.org/research/loss-leading-bans-retail-competition.

[27] When $\epsilon_D < \epsilon_W$, the base good price in an exploiting equilibrium is always larger than in the benchmark and a price floor is never binding.

Since firms must offer the base good at a higher price than in equilibrium, the demand for the base good declines and the add-on becomes more expensive with a binding price floor. This is because an exploiting firm sets $p_2 = W(v_2, \tilde{\Delta})$ and $W(v_2, \tilde{\Delta})$ is strictly increasing in p_1. This clearly harms consumers that remain in the market as the prices for both goods increase. Additionally, classical consumers who dropped out of the market because of the regulation are worse off. Without a price floor, they would buy the base good and obtain a positive surplus. Given that U_b applies, the only potential positive effect is that some behavioral consumers leave the market and do not buy the overpriced add-on.[28] But, similar to classical consumers, they also lose a positive surplus from the base good. Thus, the overall effect is ambiguous. Therefore, regulators should be careful with prohibiting predatory pricing in markets, which are likely to have behavioral consumers.

2.4.3 Add-on price cap

In his 2023 State of the Union speech, Biden called for a \$8 cap on credit card late fees (The White House, 2023), with the intention of extending such a policy to add-on fees offered by concert and sports promoters.

We discuss the impact of this regulation on the outcome of our model intuitively. First, with a price cap $\bar{p}_2 < W(v_2)$, only the non-exploiting equilibrium exists because firms cannot set the add-on price above the WTP of classical consumers.[29] Thus, a price cap prevents the exploitation of behavioral consumers completely. But this comes at efficiency costs. Since firms redistribute revenues from add-on selling, the price cap reduces those earnings, which increases the base good price. This, in turn, lowers the base-good demand. Therefore, a price cap $\bar{p}_2 < W(v_2)$ solves the problem of exploitation, but the effect on consumer surplus is unclear as fewer base goods are sold at a higher price.

[28] If $\tilde{U}_b$ applies, then the positive effect does not exist, and behavioral consumers who left are strictly worse off.

[29] When $\bar{p}_2 = W(v_2)$, the price cap leads to the identical outcome of the non-exploiting (or benchmark) equilibrium without a price cap.

When $W(v_2) < \bar{p}_2 < W(v_2, \tilde{\Delta})$, exploiting equilibria may still exist ex-post. Similar to an education policy, we need to distinguish between inefficient and efficient price caps. Analogous to Definition 4, with an ineffective price cap, firms still exploit behavioral consumers after the regulation, while an effective price cap prevents exploitation and the non-exploiting equilibrium emerges ex-post. Whether a price cap is effective depends on how strongly it affects the add-on revenue. An effective price cap limits exploitation sufficiently enough such that selling the add-on to all consumers is more profitable. The impact and intuition of an effective price cap is similar to the result in Proposition 2.4.1 (a ii). It benefits classical consumers when the ex-ante equilibrium was harmful and hurts them when the ex-ante equilibrium was beneficial. Depending on the ex-ante equilibrium, classical consumers pay a higher or lower base good price ex-post, while still receiving zero surplus from the add-on. Behavioral consumers are always better off when a regulation prevents exploitation as shown in Proposition 2.3.2. Thus, when the ex-ante equilibrium was harmful, an effective price cap regulation unambiguously increases consumer surplus. When the ex-ante equilibrium was beneficial, the impact on consumer surplus is unclear.

When the price cap is ineffective, then firms still exploit behavioral consumers and set a higher base good price because they earn less add-on revenue. This clearly harms classical consumers. Behavioral consumers enjoy a lower add-on price, but also must pay more for the base good. Thus, the impact on them is ambiguous.

In general, a price cap on the add-on limits the extent to which firms can exploit behavioral consumers. However, it is likely that the regulation will lead to more expensive base goods. The implications are similar to an education policy. However, instead of reducing the share of behavioral consumers directly, a price cap tackles the revenue that firms can make from exploitation, which increases the profit thresholds. In other words, a larger share of behavioral consumers is required that exploitation is profitable. Graphically, a price cap $W(v_2) < \bar{p}_2 < W(v_2, \tilde{\Delta})$ implies that the lines denoted with $\bar{\alpha}$ and $\hat{\alpha}$ in Figure 2.1 shift to the right and the region where classical consumers are unaffected increases.

Our discussion shows that a price regulation in the add-on market leads to non-trivial effects. It has a mixed impact on individual welfare, and it is unclear whether consumer surplus increases or decreases. A formal analysis, however, is beyond the scope of this paper, and we encourage future research to investigate on this topic.

2.5 Microfoundations

For the behavioral consumer, the base good price increases the WTP for the add-on, as empirically documented in a range of studies (Chatterjee, 2010; Erat and Bhaskaran, 2012; Xia and Monroe, 2004). In this section, we discuss several mechanisms that may microfound such behavior, which is captured in our reduced-form model through the function $W(v_2, \tilde{\Delta})$.

Relative thinking. Relative thinking has been shown to be an important determinant in individual decision-making (Jacowitz and Kahneman, 1995; Thaler, 1980). In Bushong et al. (2021), 48% of participants are willing to accept a 30-minute drive to save \$25 for a \$1000 laptop, while 73% of participants are willing to do so to save the same monetary amount when shopping for \$100 headphones. Somerville (2022) experimentally shows that more than two-thirds of the participants are better characterized as relative thinkers than as standard utility maximizers.

In Bushong et al. (2021), consumers put a relative weight $w(\Delta_k)$ on each consumption dimension $k = v, p$; where $\Delta_k = \max k_s - \min k_s$ for $s = 1, 2$ and $w(\Delta_p)$ is a differentiable and decreasing function on $(0, \infty)$. Adapting the model to our setting, the behavioral consumer is a relative thinker regarding the price dimension. To focus on this channel, we set $w(\Delta_v) = 1$. Behavioral consumers' incentive constraint for the purchase of the add-on can be written as $v_2 - w(\Delta_p)p_2 \geq 0$. We employ the parameterized example of their model: $w(\Delta_p) = (1-\rho) + \frac{\rho}{\Delta_p + \xi}$ where $\rho \in [0, 1)$ and $\xi \in (0, \infty)$. Rearranging the incentive constraint yields $W(v_2, \tilde{\Delta}) = \frac{v_2}{(1-\rho) + \frac{\rho}{\Delta_p + \xi}} \geq p_2$. Therefore, whenever the base good is more expensive than the add-on, the model of Bushong et al.

(2021) satisfies our assumptions on $W(v_2, \tilde{\Delta})$.[30]

Somerville (2022) provides a similar parameterized function for relative thinking. The incentive constraint is characterized by $v_2 - (\Delta_p)^y \cdot p_2 \geq 0$ with $y \in (-1, 0)$ and $\Delta_p = \max p_k - \min p_k$ for $k = 1, 2$. Rearranging yields $W(v_2, \tilde{\Delta}) = \frac{v_2}{(\Delta_p)^y} \geq p_2$. Again, whenever $p_1 > p_2$, the micro-foundation of Somerville (2022) satisfies our assumption of consumer behavior.

Proportional thinking. Closely related is proportional thinking (Thaler, 1980). In Tversky and Kahneman's (1981) famous jacket-calculator example, a person is willing to exert more effort to save \$10 when the relative amount of money saved is higher (see also the replications by Frisch, 1993; Mowen and Mowen, 1986; Ranyard and Abdel-Nabi, 1993). Azar (2011) shows that consumers are willing to pay more for the same constant improvement in quality when the good's price is higher. In a field experiment, Blake et al. (2021) document a lower proportional price boosts add-on sales.[31] Formally, a behavioral consumer perceives the add-on prices as $\frac{p_2}{p_1}$ (or $\frac{p_2}{1+p_1-p_2}$), which implies an incentive constraint of $v_2 - \frac{p_2}{p_1} \geq 0$. Rearranging shows that the WTP $W(v_2, \tilde{\Delta})$ with $\Delta = p_1$ (or $\Delta = 1 + p_1 - p_2$) is increasing in p_1.[32]

Salience. Consumers may devote more attention to product attributes that are more salient. For example, it is documented that consumers underreact to taxes when those are not salient (Chetty, Looney, and Kroft, 2009; Feldman and Ruffle, 2015; Taubinsky and Rees-Jones, 2018). Also, when prices become less salient, demand substantially increases (Finkelstein, 2009; Sexton, 2015). In a large field experiment on StubHub.com, Blake et al. (2021) show that drip pricing strategies increase demand due to the additional fee appearing less salient for consumers (see also Brown et al., 2010; Dertwinkel-Kalt et al., 2020; Hossain and Morgan, 2006).

Bordalo, Gennaioli, and Shleifer (2022) formalize salience theory. In their

[30] Note that $\frac{\partial W(v_2,\tilde{\Delta})}{\partial p_1} > 0$ and $\frac{\partial^2 W(v_2,\tilde{\Delta})}{\partial p_1^2} < 0$ when $p_1 > p_2$ and $\frac{\partial W(v_2,\tilde{\Delta})}{\partial v_2} > 0$ and $\frac{\partial^2 W(v_2,\tilde{\Delta})}{\partial v_2^2} = 0$.

[31] "An old selling trick is to quote a low price for a stripped-down model and then coax the consumer into a more expensive version in a series of increments each of which seems small relative to the entire purchase" (Thaler, 1980, p. 51).

[32] See also Azar (2007) who develops a model of add-on pricing with mixed consumers in which behavioral consumers' add-on WTP is given by $w(P_L) = dP_L^{\alpha\beta}$, where d is a constant capturing utility, P_L the price of the base good, $\alpha \in [0, 1]$ captures the extent of proportional thinking of a consumer, and $\beta \in [0, 1]$ reflects the extent of relative thinking inherent in a certain decision context. Setting $d = v_2$, $P_L = p_1$, $\alpha = \beta_i$ and $\beta = 1$, leads directly to our reduced form $W(v_2, p_1) = v_2 p_1^{\beta_i}$ with $\beta_i \in \{0, \beta\}$.

model, the surplus function for behavioral consumers is $\hat{V} = \sum_k w_k \pi_k a_k$ for a good with k attributes, where w_k is the weighting function capturing bottom-up attention to salient attributes, π_k is the decision weight attached to attribute k, and a_k denotes the attribute's value (see also Bordalo, Gennaioli, and Shleifer, 2012, 2013, 2020). In our case, four attributes exist, $k = \{p_1, p_2, v_1, v_2\}$. We suppose that salience happens through contrast effects, namely between the prices of the base good and add-on. To accommodate our model, salient thinking does not affect the attention weight to quality, $w_{v_1} = w_{v_2} = 1$. This deviates from the traditional theory, which typically considers the purchase between two substitutes when either a product's quality or price is salient. Thus, we encourage future research to study how salience on quality affects add-on selling. In our setup, it is the choice of buying the add-on or not, given the (tentative) purchase of the base good. We suppose that a more expensive base good captures the attention of consumers, who then underweight the add-on's price.[33] For the ease of exposition, we assume $\pi_{v_2} = 1$ and $\pi_{p_2} = -1$.

Contrast between the prices is measured by the salience function $\sigma(a_k, \bar{p}) = \frac{|a_k - \bar{p}|}{|a_k + \bar{p}|}$, where $a_k \in \{p_1, p_2\}$ and $\bar{p} = \frac{p_1 + p_2}{2}$, satisfying ordering and diminishing sensitivity properties. Observe that $p_1 > p_2 \Leftrightarrow \sigma(p_1, \bar{p}) > \sigma(p_2, \bar{p})$, implying that p_1 is more salient when the base good is more expensive. This distorts the weighting function accordingly to $w_k = w(\sigma_k; \sigma_{-k})$. Importantly, according to Bordalo et al. (2022), w_k is increasing in the salience of k, σ_k, and decreases in other attributes $-k$ salience, σ_{-k}. Thus, increasing p_1 makes the base good price more salient, with the consequence of p_2 becoming less salient. This, in turn, decreases $w_{p_2} = w(\sigma_{p_2}; \sigma_{p_1})$ and thus, behavioral consumers put less weight on the add-on price. A behavioral consumer buys the add-on when $v_2 - w_{p_2} p_2 > 0$. Rearranging the incentive constraint yields $W(v_2, \tilde{\Delta}) = \frac{v_2}{w_{p_2}} \geq p_2$, where w_{p_2} is decreasing in p_1 and increasing in p_2. Hence, $W(v_2, \tilde{\Delta})$ is increasing in both arguments. The diminishing sensitivity property of the salience function $\sigma(a_k, \bar{p})$ corresponds to our concavity assumption.

[33]Given that consumers observe the add-on offer only after the (tentative) purchase, we suppose that salience does not affect the base good market.

Mental accounting. Because consumers are mental accountants " [...] sellers have a distinct advantage in selling something if its cost can be added on to another larger purchase" (Thaler, 1985, p. 209). See also Erat and Bhaskaran (2012); Moon, Keasey, and Duxbury (1999); Ranyard and Abdel-Nabi (1993).

The transaction utility theory from Thaler (1985) is a two-stage process. First, there is a judgment process, where consumers evaluate potential transactions. The total utility is defined as $w(z, p, p^*) = v(\bar{p} - p) + v(-p : -p^*)$, where $\bar{p}$ is the valuation for a good z with price p, reference price p^*, and $v(\cdot)$ is a concave function. The term $v(\bar{p} - p)$ captures the acquisition utility, which is simply the net utility accrued by the trade and corresponds to the add-on net utility of classical consumers.[34] The transaction utility (or reference outcome) is captured by $v(-p : -p^*)$, which depends on the add-on price and the reference price. Note that $v(-p : -p) = 0$, $v(-p : -p^*) > 0$ when $p < p^*$, and $v(-p : -p^*)$ is increasing in p^*. Intuitively, when the reference price exceeds the market price, then it affects the value of good z positively. The size of the effect depends on the difference between p and p^*. Second, there is a decision process, where consumers (dis-)approve each potential transaction. A behavioral consumer will buy a good z if $\frac{w(z,p,p^*)}{p} > k$, where k is a constant. We interpret $k = 0$ as the outside option of not buying the add-on. Supposing $v(\bar{p} - p) = W(\bar{p}) - p$ and setting $\bar{p} = v_2$, $p = p_2$, and $p^* = p_1$ leads to the incentive constraint $\frac{W(v_2) - p_2 + v(-p_2 : -p_1)}{p_2} \geq 0$. Assuming $p_1 > p_2$, then $W(v_2, v(-p_2 : -p_1)) = W(v_2) + v(-p_2 : -p_1) \geq p_2$ implies $\frac{\partial v(-p_2 : -p_1)}{\partial p_1} > 0$, and, since $v(\cdot)$ is concave, the assumptions on $W(v_2, \tilde{\Delta})$ with $\Delta = v(-p_2 : -p_1)$ are satisfied. Therefore, consumers subject to mental thinking can be characterized as behavioral consumers in our model.[35]

[34] We use directly the notation $v(\bar{p} - p)$ instead of $v(\bar{p}, -p)$, since Thaler (1985) argues that acquisition utility will generally be coded as integrated outcome.

[35] Our reduced-form model also accommodates Erat and Bhaskaran (2012), who provide a mental accounting model in the context of add-on selling. The behavioral mechanism is defined as a mental book value $BV = p - V$, where p is the paid base good price and V is the cumulative benefit a consumer has obtained so far from using the base good, which increases over time. Thus, BV is maximal just after the base good purchase occurred. Further, a consumer buys the add-on if and only if $p_A \leq u_A + \gamma u_A BV$. Setting $p = p_1$, $p_A = p_2$ and $u_A = v_2$ translates immediately to our reduced form incentive constraint $W(v_2, BV(p_1)) = v_2(1 + \gamma BV(p_1)) \geq p_2$, where $BV(p_1)$ is strictly increasing in p_1.

Reference point dependence and anchoring-and-adjustment. A large amount of experimental evidence documents the importance of reference points in individual decision-making, starting with Jacowitz and Kahneman (1995); Kahneman and Tversky (1979); Tversky and Kahneman (1974, 1981). Arbitrary high anchors have been shown to increase the WTP for a variety of goods (Alevy, Landry, and List, 2015; Ariely, Loewenstein, and Prelec, 2003; Bergman, Ellingsen, Johannesson, and Svensson, 2010; Fudenberg, Levine, and Maniadis, 2012; Ioannidis, Offerman, and Sloof, 2020; Maniadis, Tufano, and List, 2014; Yoon, Fong, and Dimoka, 2019): "For example, consumers may use a heuristic called 'anchoring and adjustment', in which case consumers will anchor on the base price and insufficiently adjust for the surcharge" (Office of Fair Trading, 2013, p. 8).[36] It is also documented that the price observed in previous market periods affects subsequent bids of market participants (Beggs and Graddy, 2009; Ferraro, Messer, Shukla, and Weigel, 2021; Tufano, 2010). Therefore, we argue that anchoring and adjustment is a suitable explanation for our reduced form function $W(v_2, \tilde{\Delta})$. Formally, we incorporate the distance between p_2 and the reference price p_1 as the behavioral mechanism into the incentive constraint, $u - p_i + \gamma(\tilde{p} - p_i) \geq 0$, where $\gamma(\cdot)$ captures loss aversion (Wenner, 2015). Setting $u = v_2$, $p_i = p_2$ and $\tilde{p} = p_1$ yields immediately $W(v_2, \tilde{\Delta}) = v_2 + \gamma(\Delta) \geq p_2$ with $\Delta = p_1 - p_2$.

2.6 Further Results

2.6.1 After-sales Competition (Sequential Buying)

We relax the lock-in assumption and allow for competition in the after-sales market. We suppose the same setup as in the baseline model, but a fraction $\rho \in (0, 1)$ of base good buyers search for the cheapest add-on, while the fraction $(1 - \rho)$ stays loyal and purchases the add-on from the same company. Firms know the distribution of loyal consumers but cannot price discriminate. They choose prices p_1 and p_2 simultaneously and can commit to add-on

[36]See Furnham and Boo (2011) for a literature review on the heuristic.

prices.[37]

In equilibrium, firms still choose between the non-exploiting strategy ($p_2 \leq W(v_2)$) and the exploiting strategy ($W(v_2) < p_2 \leq W(v_2, \tilde{\Delta})$), but mix over the choice of add-on prices. To see this, consider the symmetric equilibria given in Lemma 2.A.3. Since searching consumers buy the add-on from the cheapest seller, a firm can profitably deviate by setting a slightly lower price and capturing all non-loyal customers. The other extreme, marginal cost pricing and earning zero after-sales profits, is also not optimal. Firms can always just sell the add-on to the loyal consumers at the WTP of either classical or behavioral consumers and make positive after-sales profits. Thus, there is mixing in add-on prices, where firms must be indifferent between mixing, and potentially attracting some consumers from the rival, or setting $p_2 \in \{W(v_2), W(v_2, \tilde{\Delta})\}$, and sell the add-on to only loyal (non-) classical consumers. This result resembles the findings of Baye and Morgan (2001).

Therefore, the expected profit from the add-on market of a non-exploiting firm is given by $(1-\rho)D_j(\cdot)W(v_2)$. Similarly, an exploiting firm expects to earn $\alpha(1-\rho)D_j(\cdot)W(v_2, \tilde{\Delta})$ in the aftermarket. These can be substituted into the profit function (2.1), which simplifies the maximization problem greatly, and we can proceed as in the baseline model. Note that the introduction of searching consumers is simply a rescaling and does not change the analysis qualitatively. Relaxing the lock-in assumption lowers not only the (expected) profits of an exploiting firm but also of non-exploiting firms. Hence, the equilibrium characterization is identical to Lemma 2.A.3 with following exception for p_2^*:

Lemma 2.6.1.

(i) A non-exploiting firm draws an add-on price p_2^n from a continuous and atomless price distribution $F^n(p_2^n)$ with $p_2^n \in (\underline{p}_2^n, W(v_2))$.

(ii) An exploiting firm draws an add-on price p_2^e from a continuous and atomless price distribution $F^e(p_2^e)$ with $p_2^e \in (\max\{\underline{p}_2^e, W(v_2)\}, W(v_2, \tilde{\Delta}))$.

[37] This assumption is merely for simplicity. The results do not change when p_1 and p_2 are chosen sequentially. In any equilibria, firms mix their choice of p_2, independently of simultaneous or sequential price setting.

Importantly, the semi-elasticities ϵ_D and ϵ_W, and the price threshold $\bar{\alpha}_p$ are (qualitatively) unchanged and we still have $\bar{\alpha}_p > \max\{\bar{\alpha}, \hat{\alpha}\}$. Thus, the results from Lemma 2.3.1 and Proposition 2.3.1 follow immediately. Therefore, our central finding that the presence of behavioral consumers can harm classical consumers does not rely on the lock-in assumption.

2.6.2 Unit Demand

We apply our framework to a model with unit base good demand and horizontal differentiation, which is commonly used in the literature (e.g., Armstrong and Vickers, 2012; Ellison, 2005; Gabaix and Laibson, 2006; Heidhues and Kőszegi, 2017). The full analysis is provided in Appendix 2.C.5.

We find similar results as in the baseline model with imperfect competition. We show that the equilibrium is alike as characterized in Lemma 2.A.3 and the optimal base good price p_1 behaves similar to Lemma 2.3.1. Crucially, the main findings stated in Proposition 2.3.1 hold and are not affected by the different demand structure. When firms exploit behavioral consumers, this can benefit or harm classical consumers.

2.7 Conclusion

We study an after-sales market with behavioral consumers who are subject to an effect that temporarily increases the WTP for the add-on. We show that several well-known and studied mechanisms, such as relative thinking, proportional thinking, salience, mental accounting, or anchoring-and-adjustment, can motivate our reduced-form model. When confronted with such behavior, firms face an incentive to increase the base good price.

We provide a novel result in the context of add-on selling and drip pricing: when firms have market power in the base good market, then behavioral consumers exert non-monotonic effects on the surplus of classical consumers. The direction of the impact depends on the proportion of behavioral consumers in the population. A relatively equal mixture of the two types in the population turns classical consumers worse off than in the benchmark economy. These findings prompt the question of whether it is necessary to shield rational consumers from the errors and biases of others. However, when the proportion of behavioral consumers is substantial, it can lead to an increase in the surplus of classical consumers but at the expense of behavioral consumers, which can result in significant distributional effects.

We use our model to assess the impact of some potential policies. Education policies that reduce the share of the behavioral consumer can increase consumer surplus but could also decrease it. Similarly, an exogenous price floor imposed in the base good market, originally implemented with the intention of preventing loss leading, harms all consumers remaining in the market. While classical consumers' surplus is not jeopardized by the behavioral consumer in a perfectly competitive market, competition comes along with a distributional effect in which the behavioral type subsidizes the classical consumer.

Our model and findings could also be relevant in labor markets, organizational settings and for gift-exchange (Akerlof, 1982; Azar, 2019; Fehr, Kirchsteiger, and Riedl, 1993; Hart and Moore, 2008). Experimental evidence suggests that low wages relative to past wages decrease the labor supply (Bracha, Gneezy, and Loewenstein, 2015). Similarly, the flat wage agreed-upon in a contract might represent the base good and serve as reference. Consequently, workers might differently perceive and reciprocate the very same \$1000 year-end bonus. A similar effect may play a role in meal allowances or other additional employee benefits.

Our model comes along with some limitations. First, we presume that firms offer first a base good and then an add-on, but have no choice in offering a bundle product. Variations of our model could investigate firms' optimal pricing strategies and its consequences on welfare in product design, such as Apffelstaedt and Mechtenberg (2021) do with context-sensitive consumers. Second, our model does not capture a setting with mandatory add-ons, such as unavoidable surcharge fee's (Rasch et al., 2020). Future research could address this and extend our model to accommodate mandatory surcharges. Third, we assume that the WTP of behavioral consumers for the add-on is monotonously increasing in the price of the base good. A disproportionately high-priced add-on, however, may be perceived as unfair (Herz and Taubinsky, 2017; Rabin, 1993; Robbert and Roth, 2014). Fairness effects may impose an upper limit for the add-on price behavioral consumers accept. Finally, it would be valuable if future research could provide causal empirical evidence for our theoretical implications.

2.A Auxiliary Results

To ease notation, we denote $\frac{\partial D_j(p_{1,j}, p_{1,-j})}{\partial p_{1,j}} = D'_j(p_{1,j}, p_{1,-j})$ and $\frac{\partial^2 D_j(p_{1,j}, p_{1,-j})}{\partial p^2_{1,j}} = D''_j(p_{1,j}, p_{1,-j})$.

2.A.1 Non-exploiting strategy

Suppose firm j does not exploit and sets $p_{2,j} = W(v_2)$. This implies $Q_j(p_{2,j}, D_j(\cdot)) = D_j(p_{1,j}, p_{1,-j})$ and the profit function (2.1) reduces to

$$\pi_j(p_{1,j}, p_{1,-j}, W(v_2)) = \left[p_{1,j} + W(v_2)\right] D_j(p_{1,j}, p_{1,-j}). \tag{2.2}$$

Note that the optimization problem in the benchmark economy ($\alpha = 0$) is identical to (2.2). Maximizing this expression with respect to $p_{1,j}$ yields the first-order condition

$$\begin{aligned} D_j(p_{1,j}, p_{1,-j}) + D'_j(p_{1,j}, p_{1,-j})[p_{1,j} + W(v_2)] = 0 \\ \Leftrightarrow \quad p_{1,j} = \frac{-D_j(p_{1,j}, p_{1,-j})}{D'_j(p_{1,j}, p_{1,-j})} - W(v_2). \end{aligned}$$

Substituting $p_{1,j}$ in expression (2.2) leads to

$$\pi_j(p_{1,j}, p_{1,-j}, W(v_2)) = \frac{-D_j(p_{1,j}, p_{1,-j})^2}{D'_j(p_{1,j}, p_{1,-j})}.$$

Whether firm j sets p_1^n or $\tilde{p}_1^n(\alpha)$ depends on the action of firm $-j$. First, suppose firm $-j$ does not exploit. Then, both firms set

$$p_1^n = \frac{-D(p_1^n, p_1^n)}{D'(p_1^n, p_1^n)} - W(v_2)$$

and obtain

$$\pi^n = \pi(p_1^n, p_1^n, W(v_2)) = \frac{-D(p_1^n, p_1^n)^2}{D'(p_1^n, p_1^n)}.$$

Observe that neither p_1^n nor π^n depend on α. Therefore, the symmetric non-exploiting outcome is independent of the share of behavioral consumers. Further, the benchmark outcome ($\alpha = 0$) is identical since it has the same maximization problem. That is $p_1^n = p_1^b$ and $\pi^n = \pi^b$. Now suppose firm $-j$ exploits and sets $\tilde{p}_1^e(\alpha)$. Then, firm j sets

$$\tilde{p}_1^n(\alpha) = \frac{-D(\tilde{p}_1^n(\alpha), \tilde{p}_1^e(\alpha))}{D'(\tilde{p}_1^n(\alpha), \tilde{p}_1^e(\alpha))} - W(v_2)$$

and obtains

$$\tilde{\pi}^n = \pi\left(\tilde{p}_1^n(\alpha), \tilde{p}_1^e(\alpha), W(v_2)\right) = \frac{-D(\tilde{p}_1^n(\alpha), \tilde{p}_1^e(\alpha))^2}{D'(\tilde{p}_1^n(\alpha), \tilde{p}_1^e(\alpha))}.$$

As we show later, $\tilde{p}_1^n(\alpha)$ and thus, $\tilde{\pi}^n$, depend on α because $\tilde{p}_1^e(\alpha)$ does and base good prices are strategic complements.

2.A.2 Exploiting strategy

Suppose firm j exploits and sets $p_{2,j} = W(v_2, \tilde{\Delta})$. This implies $Q_j(p_{2,j}, D_j(\cdot)) = \alpha D_j(p_{1,j}, p_{1,-j})$ and the profit function (2.1) reduces to

$$\pi_j(p_{1,j}, p_{1,-j}, W(v_2, \tilde{\Delta})) = \left[p_{1,j} + \alpha W(v_2, \tilde{\Delta})\right] D_j(p_{1,j}, p_{1,-j}). \tag{2.3}$$

Maximizing this expression with respect to $p_{1,j}$ yields the first-order condition

$$[1 + \alpha W'(v_2, \tilde{\Delta})] D_j(p_{1,j}, p_{1,-j}) + D_j'(p_{1,j}, p_{1,-j})[p_{1,j} + \alpha W(v_2, \tilde{\Delta})] = 0$$

$$\Leftrightarrow \quad p_{1,j} = \frac{-[1+\alpha W'(v_2,\tilde{\Delta})]D_j(p_{1,j},p_{1,-j})}{D'_j(p_{1,j},p_{1,-j})} - \alpha W(v_2,\tilde{\Delta}),$$

where $W'(v_2,\tilde{\Delta}) = \frac{\partial W(v_2,\tilde{\Delta})}{\partial p_{1,j}}$. Substituting $p_{1,j}$ in expression (2.3) leads to

$$\pi_j(p_{1,j},p_{1,-j},W(v_2,\tilde{\Delta})) = \frac{-[1+\alpha W'(v_2,\tilde{\Delta})]D_j(p_{1,j},p_{1,-j})^2}{D'_j(p_{1,j},p_{1,-j})}.$$

Whether firm j sets $p_1^e(\alpha)$ or $\tilde{p}_1^e(\alpha)$ depends on the action of firm $-j$. First, suppose firm $-j$ exploits. Then, both firms set

$$p_1^e(\alpha) = \frac{-[1+\alpha W'(v_2,\tilde{\Delta})]D(p_1^e(\alpha),p_1^e(\alpha))}{D'(p_1^e(\alpha),p_1^e(\alpha))} - \alpha W(v_2,\tilde{\Delta})$$

and obtain

$$\pi^e = \pi(p_1^e(\alpha),p_1^e(\alpha),W(v_2,\tilde{\Delta})) = \frac{-[1+\alpha W'(v_2,\tilde{\Delta})]D(p_1^e(\alpha),p_1^e(\alpha))^2}{D'(p_1^e(\alpha),p_1^e(\alpha))}.$$

Now suppose firm $-j$ does not exploit and sets $\tilde{p}_1^n(\alpha)$. Then, firm j sets

$$\tilde{p}_1^e(\alpha) = \frac{-[1+\alpha W'(v_2,\tilde{\Delta})]D(\tilde{p}_1^e(\alpha),\tilde{p}_1^n(\alpha))}{D'(\tilde{p}_1^e(\alpha),\tilde{p}_1^n(\alpha))} - \alpha W(v_2,\tilde{\Delta})$$

and obtains

$$\tilde{\pi}^e = \pi\left(\tilde{p}_1^e(\alpha),\tilde{p}_1^n(\alpha),W(v_2,\tilde{\Delta})\right) = \frac{-[1+\alpha W'(v_2,\tilde{\Delta})]D(\tilde{p}_1^e(\alpha),\tilde{p}_1^n(\alpha))^2}{D'(\tilde{p}_1^e(\alpha),\tilde{p}_1^n(\alpha))}.$$

2.A.3 Derivatives

It is crucial for our analysis to understand how the base-good prices and profits react to changes in α. Recall that $D_j(p_{1,j},p_{1,-j})$ is a concave function and is strictly decreasing in $p_{1,j}$, which implies $D'(\cdot) < 0$ and $D''(\cdot) \leq 0$. Further, we have $\frac{\partial D_j(p_{1,j},p_{1,-j})}{\partial p_{1,-j}} \geq 0$ and $\frac{\partial^2 D_j(p_{1,j},p_{1,-j})}{\partial p_{1,j}\partial p_{1,-j}} \geq 0$, since base goods are strategic complements and demand is supermodular. Further, a stronger own price elasticity implies $\left|D'_j(p_{1,j},p_{1,-j})\right| > \frac{\partial D_j(p_{1,j},p_{1,-j})}{\partial p_{1,-j}}$. Finally, since $W(v_2,\tilde{\Delta})$ is strictly increasing in all arguments and concave, we have

$W'(v_2, \tilde{\Delta}) = \frac{\partial W(v_2,\tilde{\Delta})}{\partial p_{1,j}} > 0$ and $W''(v_2, \Delta) = \frac{\partial^2 W(v_2,\tilde{\Delta})}{\partial p_{1,j}^2} \leq 0$. Given the assumptions on $D(\cdot)$ and $W(v_2, \tilde{\Delta})$, Lemma 2.A.1 characterizes how profits and prices react to a change of α.

Lemma 2.A.1.

(a) p_1^n and π^n are constant in α.

(b) π^e and $\tilde{\pi}^e$ are strictly increasing in α.

(c) $p_1^e(\alpha)$, $\tilde{p}_1^e(\alpha)$, $\tilde{p}_1^n(\alpha)$ and $\tilde{\pi}^n$ are (i) strictly decreasing in α if $\epsilon_D > \epsilon_W$, (ii) strictly increasing in α if $\epsilon_D < \epsilon_W$, and (iii) constant in α if $\epsilon_D = \epsilon_W$.

Proof. (a)

$$\frac{\partial p_1^n}{\partial \alpha} = \left[-1 + \frac{D(p_1^n, p_1^n)D''(p_1^n, p_1^n)}{D'(p_1^n, p_1^n)^2}\right]\frac{\partial p_1^n}{\partial \alpha}$$
$$\Leftrightarrow \quad \frac{\partial p_1^n}{\partial \alpha}\left[2 - \frac{D(p_1^n, p_1^n)D''(p_1^n, p_1^n)}{D'(p_1^n, p_1^n)^2}\right] = 0$$

Since $\frac{D(\cdot)D''(\cdot)}{D'(\cdot)^2} \leq 0$, it must be that $\frac{\partial p_1^n}{\partial \alpha} = 0$.

$$\frac{\partial \pi_1^n}{\partial \alpha} = \frac{-2D(p_1^n, p_1^n)D'(p_1^n, p_1^n)^2 + D(p_1^n, p_1^n)^2 D''(p_1^n, p_1^n)}{D'(p_1^n, p_1^n)^2}\underbrace{\frac{\partial p_1^n}{\partial \alpha}}_{=0} = 0.$$

(b) We characterize first $\frac{\partial p_1^e(\alpha)}{\partial \alpha}$, before we can derive $\frac{\partial \pi_1^e}{\partial \alpha}$.

$$\begin{aligned}
&\frac{\partial p_1^e(\alpha)}{\partial \alpha} = \\
&-\left[1 + 2\alpha W'(v_2, \tilde{\Delta}) + \frac{\alpha W''(v_2, \Delta)D(p_1^e(\alpha), p_1^e(\alpha))}{D'(p_1^e(\alpha), p_1^e(\alpha))}\right]\frac{\partial p_1^e(\alpha)}{\partial \alpha} \\
&+ (1 + \alpha W'(v_2, \tilde{\Delta}))\frac{D(p_1^e(\alpha), p_1^e(\alpha))D''(p_1^e(\alpha), p_1^e(\alpha))}{D'(p_1^e(\alpha), p_1^e(\alpha))^2}\frac{\partial p_1^e(\alpha)}{\partial \alpha} \\
&- W(v_2, \tilde{\Delta}) - \frac{W'(v_2, \tilde{\Delta})D(p_1^e(\alpha), p_1^e(\alpha))}{D'(p_1^e(\alpha), p_1^e(\alpha))} \\
&\Leftrightarrow \quad \frac{\partial p_1^e(\alpha)}{\partial \alpha} =
\end{aligned}$$

$$\frac{-W(v_2,\tilde{\Delta})-\frac{W'(v_2,\tilde{\Delta})D(p_1^e(\alpha),p_1^e(\alpha))}{D'(p_1^e(\alpha),p_1^e(\alpha))}}{(1+\alpha W'(v_2,\tilde{\Delta}))\left(2-\frac{D(p_1^e(\alpha),p_1^e(\alpha))D''(p_1^e(\alpha),p_1^e(\alpha))}{D'(p_1^e(\alpha),p_1^e(\alpha))^2}\right)+\frac{\alpha W''(v_2,\Delta)D(p_1^e(\alpha),p_1^e(\alpha))}{D'(p_1^e(\alpha),p_1^e(\alpha))}}$$

$$\begin{aligned}\frac{\partial\pi_1^e}{\partial\alpha} &= D(p_1^e(\alpha),p_1^e(\alpha))\Bigg[-\frac{W'(v_2,\tilde{\Delta})D(p_1^e(\alpha),p_1^e(\alpha))}{D'(p_1^e(\alpha),p_1^e(\alpha))}\\ &-\frac{\partial p_1^e(\alpha)}{\partial\alpha}\Bigg[(1+\alpha W'(v_2,\tilde{\Delta}))\left(2-\frac{D(p_1^e(\alpha),p_1^e(\alpha))D''(p_1^e(\alpha),p_1^e(\alpha))}{D'(p_1^e(\alpha),p_1^e(\alpha))^2}\right)\\ &+\frac{\alpha W''(v_2,\Delta)D(p_1^e(\alpha),p_1^e(\alpha))}{D'(p_1^e(\alpha),p_1^e(\alpha))}\Bigg]\Bigg] = D(p_1^e(\alpha),p_1^e(\alpha))W(v_2,\tilde{\Delta})>0,\end{aligned}$$

where the second equality follows from substituting $\frac{\partial p_1^e(\alpha)}{\partial\alpha}$. Before we derive $\frac{\partial\tilde{\pi}_1^e}{\partial\alpha}$, we need to characterize $\frac{\partial\tilde{p}_1^n(\alpha)}{\partial\alpha}$ and $\frac{\partial\tilde{p}_1^e(\alpha)}{\partial\alpha}$:

$$\frac{\partial\tilde{p}_1^n(\alpha)}{\partial\alpha}=\underbrace{\left[\frac{\frac{D(\tilde{p}_1^n(\alpha),\tilde{p}_1^e(\alpha))\frac{\partial^2 D(\tilde{p}_1^n(\alpha),\tilde{p}_1^e(\alpha))}{\partial\tilde{p}_1^n(\alpha)\partial\tilde{p}_1^e(\alpha)}}{D'(\tilde{p}_1^n(\alpha),\tilde{p}_1^e(\alpha))^2}-\frac{\frac{\partial D(\tilde{p}_1^n(\alpha),\tilde{p}_1^e(\alpha))}{\partial\tilde{p}_1^e(\alpha)}}{D'(\tilde{p}_1^n(\alpha),\tilde{p}_1^e(\alpha))}}{2-\frac{D(\tilde{p}_1^n(\alpha),\tilde{p}_1^e(\alpha))D''(\tilde{p}_1^n(\alpha),\tilde{p}_1^e(\alpha))}{D'(\tilde{p}_1^n(\alpha),\tilde{p}_1^e(\alpha))^2}}\right]}_{=A}\frac{\partial\tilde{p}_1^e(\alpha)}{\partial\alpha}.$$

Note that $A\geq 0$ since $D(\cdot)$ is concave, supermodular, strictly decreasing in the first argument and increasing in the second argument. Taking the derivative of $\tilde{p}_1^e(\alpha)$ with respect to α yields

$$\begin{aligned}&\frac{\partial\tilde{p}_1^e(\alpha)}{\partial\alpha}\Bigg[(1+\alpha W'(v_2,\tilde{\Delta}))\left[2-\frac{D(\tilde{p}_1^e(\alpha),\tilde{p}_1^n(\alpha))D''(\tilde{p}_1^e(\alpha),\tilde{p}_1^n(\alpha))}{D'(\tilde{p}_1^e(\alpha),\tilde{p}_1^n(\alpha))^2}\right]\\ &+\frac{\alpha W''(v_2,\Delta)D(\tilde{p}_1^e(\alpha),\tilde{p}_1^n(\alpha))}{D'(\tilde{p}_1^e(\alpha),\tilde{p}_1^n(\alpha))}\Bigg]=\\ &-W(v_2,\tilde{\Delta})-\frac{W'(v_2,\tilde{\Delta})D(\tilde{p}_1^e(\alpha),\tilde{p}_1^n(\alpha))}{D'(\tilde{p}_1^e(\alpha),\tilde{p}_1^n(\alpha))}+(1+\alpha W'(v_2,\tilde{\Delta}))\\ &\cdot\underbrace{\left[\frac{D(\tilde{p}_1^e(\alpha),\tilde{p}_1^n(\alpha))\frac{\partial^2 D(\tilde{p}_1^e(\alpha),\tilde{p}_1^n(\alpha))}{\partial\tilde{p}_1^e(\alpha)\partial\tilde{p}_1^n(\alpha)}}{D'(\tilde{p}_1^e(\alpha),\tilde{p}_1^n(\alpha))^2}-\frac{\frac{\partial D(\tilde{p}_1^e(\alpha),\tilde{p}_1^n(\alpha))}{\partial\tilde{p}_1^n(\alpha)}}{D'(\tilde{p}_1^e(\alpha),\tilde{p}_1^n(\alpha))}\right]}_{=B}\frac{\partial\tilde{p}_1^n(\alpha)}{\partial\alpha}.\end{aligned}$$

Note that $B \geq 0$. Substituting $\frac{\partial \tilde{p}_1^n(\alpha)}{\partial \alpha}$ yields

$$\frac{\partial \tilde{p}_1^e(\alpha)}{\partial \alpha} = \frac{-W(v_2, \tilde{\Delta}) - \frac{W'(v_2,\tilde{\Delta})D(\tilde{p}_1^e(\alpha),\tilde{p}_1^n(\alpha))}{D'(\tilde{p}_1^e(\alpha),\tilde{p}_1^n(\alpha))}}{(1 + \alpha W'(v_2, \tilde{\Delta}))\left[2 - \frac{D(\tilde{p}_1^e(\alpha),\tilde{p}_1^n(\alpha))D''(\tilde{p}_1^e(\alpha),\tilde{p}_1^n(\alpha))}{D'(\tilde{p}_1^e(\alpha),\tilde{p}_1^n(\alpha))^2} - AB\right] + \frac{\alpha W''(v_2,\Delta)D(\tilde{p}_1^e(\alpha),\tilde{p}_1^n(\alpha))}{D'(\tilde{p}_1^e(\alpha),\tilde{p}_1^n(\alpha))}}$$

Now we can take the derivative of $\tilde{\pi}_1^e$ with respect to α.

$$\begin{aligned}
\frac{\partial \tilde{\pi}_1^e}{\partial \alpha} &= D(\tilde{p}_1^e(\alpha), \tilde{p}_1^n(\alpha))\Bigg[- \frac{\partial \tilde{p}_1^e(\alpha)}{\partial \alpha}\Bigg[(1 + \alpha W'(v_2, \tilde{\Delta})) \\
&\left[2 - \frac{D(\tilde{p}_1^e(\alpha), \tilde{p}_1^n(\alpha))D''(\tilde{p}_1^e(\alpha), \tilde{p}_1^n(\alpha))}{D'(\tilde{p}_1^e(\alpha), \tilde{p}_1^n(\alpha))^2}\right] + \frac{\alpha W''(v_2, \Delta)D(\tilde{p}_1^e(\alpha), \tilde{p}_1^n(\alpha))}{D'(\tilde{p}_1^e(\alpha), \tilde{p}_1^n(\alpha))}\Bigg] \\
&+ \frac{\partial \tilde{p}_1^n(\alpha)}{\partial \alpha}(1 + \alpha W'(v_2, \tilde{\Delta})) \\
&\underbrace{\left[\frac{D(\tilde{p}_1^e(\alpha), \tilde{p}_1^n(\alpha))\frac{\partial^2 D(\tilde{p}_1^e(\alpha),\tilde{p}_1^n(\alpha))}{\partial \tilde{p}_1^e(\alpha)\partial \tilde{p}_1^n(\alpha)}}{D'(\tilde{p}_1^e(\alpha), \tilde{p}_1^n(\alpha))^2} - \frac{2\frac{\partial D(\tilde{p}_1^e(\alpha),\tilde{p}_1^n(\alpha))}{\partial \tilde{p}_1^n(\alpha)}}{D'(\tilde{p}_1^e(\alpha), \tilde{p}_1^n(\alpha))}\right]}_{=C} \\
&- \frac{W'(v_2, \tilde{\Delta})D(\tilde{p}_1^e(\alpha), \tilde{p}_1^n(\alpha))}{D'(\tilde{p}_1^e(\alpha), \tilde{p}_1^n(\alpha))}\Bigg] \\
&= D(\tilde{p}_1^e(\alpha), \tilde{p}_1^n(\alpha))\Bigg[- \frac{W'(v_2, \tilde{\Delta})D(\tilde{p}_1^e(\alpha), \tilde{p}_1^n(\alpha))}{D'(\tilde{p}_1^e(\alpha), \tilde{p}_1^n(\alpha))} \\
&- \frac{\partial \tilde{p}_1^e(\alpha)}{\partial \alpha}\Bigg[(1 + \alpha W'(v_2, \tilde{\Delta}))\left[2 - \frac{D(\tilde{p}_1^e(\alpha), \tilde{p}_1^n(\alpha))D''(\tilde{p}_1^e(\alpha), \tilde{p}_1^n(\alpha))}{D'(\tilde{p}_1^e(\alpha), \tilde{p}_1^n(\alpha))^2} - AC\right] \\
&+ \frac{\alpha W''(v_2, \Delta)D(\tilde{p}_1^e(\alpha), \tilde{p}_1^n(\alpha))}{D'(\tilde{p}_1^e(\alpha), \tilde{p}_1^n(\alpha))}\Bigg]\Bigg] \\
&= D(\tilde{p}_1^e(\alpha), \tilde{p}_1^n(\alpha))\Bigg[\left(W(v_2, \tilde{\Delta}) + \frac{W'(v_2, \tilde{\Delta})D(\tilde{p}_1^e(\alpha), \tilde{p}_1^n(\alpha))}{D'(\tilde{p}_1^e(\alpha), \tilde{p}_1^n(\alpha))}\right) \cdot \\
&\frac{(1 + \alpha W'(v_2, \tilde{\Delta}))\left[2 - \frac{D(\tilde{p}_1^e(\alpha),\tilde{p}_1^n(\alpha))D''(\tilde{p}_1^e(\alpha),\tilde{p}_1^n(\alpha))}{D'(\tilde{p}_1^e(\alpha),\tilde{p}_1^n(\alpha))^2} - AC\right] + \frac{\alpha W''(v_2,\Delta)D(\tilde{p}_1^e(\alpha),\tilde{p}_1^n(\alpha))}{D'(\tilde{p}_1^e(\alpha),\tilde{p}_1^n(\alpha))}}{(1 + \alpha W'(v_2, \tilde{\Delta}))\left[2 - \frac{D(\tilde{p}_1^e(\alpha),\tilde{p}_1^n(\alpha))D''(\tilde{p}_1^e(\alpha),\tilde{p}_1^n(\alpha))}{D'(\tilde{p}_1^e(\alpha),\tilde{p}_1^n(\alpha))^2} - AB\right] + \frac{\alpha W''(v_2,\Delta)D(\tilde{p}_1^e(\alpha),\tilde{p}_1^n(\alpha))}{D'(\tilde{p}_1^e(\alpha),\tilde{p}_1^n(\alpha))}} \\
&- \frac{W'(v_2, \tilde{\Delta})D(\tilde{p}_1^e(\alpha), \tilde{p}_1^n(\alpha))}{D'(\tilde{p}_1^e(\alpha), \tilde{p}_1^n(\alpha))}\Bigg] > 0
\end{aligned}$$

The inequality follows since $2 - \frac{D(\tilde{p}_1^e(\alpha),\tilde{p}_1^n(\alpha))D''(\tilde{p}_1^e(\alpha),\tilde{p}_1^n(\alpha))}{D'(\tilde{p}_1^e(\alpha),\tilde{p}_1^n(\alpha))^2} - AC > 0$ because $\left|\frac{\partial D_j^2(\cdot)}{\partial p_{1,j}^2}\right| \geq \frac{\partial D_j^2(\cdot)}{\partial p_{1,j}\partial p_{1,-j}}$ and $\left|D_j'(p_{1,j},p_{1,-j})\right| > \left|\frac{\partial D_j(p_{1,j},p_{1,-j})}{\partial p_{1,-j}}\right|$, and $B < C$ implies

$$2 - \frac{D(\tilde{p}_1^e(\alpha),\tilde{p}_1^n(\alpha))D''(\tilde{p}_1^e(\alpha),\tilde{p}_1^n(\alpha))}{D'(\tilde{p}_1^e(\alpha),\tilde{p}_1^n(\alpha))^2} - AB > 0.$$

Thus, the fraction on the second last line is strictly positive. Since $\frac{W'(v_2,\tilde{\Delta})D(\tilde{p}_1^e(\alpha),\tilde{p}_1^n(\alpha))}{D'(\tilde{p}_1^e(\alpha),\tilde{p}_1^n(\alpha))} < 0$, the result follows.

(c) First, observe that, given our assumptions, the denominators of $\frac{\partial p_1^e(\alpha)}{\partial \alpha}$ and $\frac{\partial \tilde{p}_1^e(\alpha)}{\partial \alpha}$ are strictly positive. Whether the nominators are positive or negative depends on whether ϵ_D or ϵ_W dominates.

$$\begin{aligned}
\frac{\partial p_1^e(\alpha)}{\partial \alpha} &\leq 0 \\
-W(v_2,\tilde{\Delta}) - \frac{W'(v_2,\tilde{\Delta})D(p_1^e(\alpha),p_1^e(\alpha))}{D'(p_1^e(\alpha),p_1^e(\alpha))} &\leq 0 \\
-\frac{D'(p_1^e(\alpha),p_1^e(\alpha))}{D(p_1^e(\alpha),p_1^e(\alpha))} &\geq \frac{W'(v_2,\tilde{\Delta})}{W(v_2,\tilde{\Delta})} \\
\epsilon_D &\geq \epsilon_W.
\end{aligned}$$

Hence, it follows $\epsilon_D \geq \epsilon_W \Leftrightarrow \frac{\partial p_1^e(\alpha)}{\partial \alpha} \leq 0$. Thus, $p_1^e(\alpha)$ is strictly decreasing if $\epsilon_D > \epsilon_W$, strictly increasing if $\epsilon_D < \epsilon_W$ and constant if $\epsilon_D = \epsilon_W$.

The argument for $\frac{\partial \tilde{p}_1^e(\alpha)}{\partial \alpha}$ and thus, $\tilde{p}_1^e(\alpha)$, is analogous. Observe that $\frac{\partial \tilde{p}_1^n(\alpha)}{\partial \alpha} \leq 0$ if $\frac{\partial \tilde{p}_1^e(\alpha)}{\partial \alpha} \leq 0$ and strictly positive otherwise. Thus, the result for $\tilde{p}_1^n(\alpha)$ follows immediately.

Finally,

$$\frac{\partial \tilde{\pi}_1^n}{\partial \alpha} = \underbrace{-D(\tilde{p}_1^n(\alpha),\tilde{p}_1^e(\alpha))\frac{\frac{\partial D(\tilde{p}_1^n(\alpha),\tilde{p}_1^e(\alpha))}{\partial \tilde{p}_1^e(\alpha)}}{D'(\tilde{p}_1^n(\alpha),\tilde{p}_1^e(\alpha))}}_{>0}\frac{\partial \tilde{p}_1^e(\alpha)}{\partial \alpha}.$$

Hence, $\tilde{\pi}_1^n$ is strictly decreasing if $\epsilon_D > \epsilon_W$, strictly increasing if $\epsilon_D < \epsilon_W$

and constant if $\epsilon_D = \epsilon_W$.

□

Lemma 2.A.2 shows that the prices $p_1^e(\alpha)$, $\tilde{p}_1^e(\alpha)$ and $\tilde{p}_1^n(\alpha)$, and thus $\tilde{\pi}_1^n$, are monotonic in α. That is, for a given $D(\cdot)$ and $W(v_2, \tilde{\Delta})$, the price functions are either increasing or decreasing for all $\alpha \in [0, 1]$.

Lemma 2.A.2. *Fix $D(\cdot)$ and $W(v_2, \tilde{\Delta})$. The base-good prices $p_1^e(\alpha)$, $\tilde{p}_1^e(\alpha)$ and $\tilde{p}_1^n(\alpha)$ are monotonic in the share of behavioral consumers α.*

Proof. We provide the proof for $p_1^e(\alpha)$. The argument for $\tilde{p}_1^e(\alpha)$ and $\tilde{p}_1^n(\alpha)$ are analogous. Observe that

$$\frac{\partial \epsilon_D}{\partial p_1^e(\alpha)} = \frac{-D''(p_1^e(\alpha), p_1^e(\alpha))D(p_1^e(\alpha), p_1^e(\alpha)) + D'(p_1^e(\alpha), p_1^e(\alpha))^2}{D(p_1^e(\alpha), p_1^e(\alpha))^2} > 0$$

since $D(p_{1,j}, p_{1,-j})$ is concave, and

$$\frac{\partial \epsilon_W}{\partial p_1^e(\alpha)} = \frac{W''(v_2, \Delta)W(v_2, \tilde{\Delta}) - W'(v_2, \tilde{\Delta})^2}{W(v_2, \tilde{\Delta})^2} < 0$$

since $W(v_2, \tilde{\Delta})$ is concave.

First, suppose $\epsilon_D > \epsilon_W$ at an initial share of behavioral consumers $\alpha_0 \in [0, 1]$. By Lemma 2.A.1, we have $\frac{\partial p_1^e(\alpha)}{\partial \alpha} < 0$. This implies, since $\frac{\partial \epsilon_D}{\partial p_1^e(\alpha)} > 0$ and $\frac{\partial \epsilon_W}{\partial p_1^e(\alpha)} < 0$, that ϵ_D and ϵ_W are converging for $\alpha > \alpha_0$ and diverging for $\alpha < \alpha_0$.

Since ϵ_D and ϵ_W are converging for an increasing α, there exists a threshold value $\tilde{\alpha} > \alpha_0$ such that $\epsilon_D = \epsilon_W$. Note that $\tilde{\alpha} > 1$ is possible. By Lemma 2.A.1, we have $\frac{\partial p_1^e(\alpha)}{\partial \alpha} = 0$ when $\epsilon_D = \epsilon_W$. Hence, a further increase $\alpha > \tilde{\alpha}$ does not change the optimal base-good price $p_1^e(\alpha)$. But then it must be $\epsilon_D = \epsilon_W$ for all $\alpha \geq \tilde{\alpha}$ and thus, $p_1^e(\alpha)$ is constant in α for all $\alpha \geq \tilde{\alpha}$ and strictly decreasing in α for all $\alpha \in [\alpha_0, \tilde{\alpha})$.

Since ϵ_D and ϵ_W are diverging for a decreasing α, $p_1^e(\alpha)$ is a strictly decreasing function for all $\alpha \in [0, \alpha_0]$. Hence, $p_1^e(\alpha)$ is strictly decreasing in the domain $\alpha \in [0, \tilde{\alpha})$ and constant in α for all $\alpha \geq \tilde{\alpha}$, which implies that $p_1^e(\alpha)$ is monotonic for $\alpha \in [0, 1]$ if $\epsilon_D > \epsilon_W$ at α_0.

Now, suppose that $\epsilon_D < \epsilon_W$ at an initial share of behavioral consumers $\alpha_0 \in [0,1]$. By Lemma 2.A.1, we have $\frac{\partial p_1^e(\alpha)}{\partial \alpha} > 0$. This implies again that ϵ_D and ϵ_W are converging for $\alpha > \alpha_0$ and diverging for $\alpha < \alpha_0$. Thus, we can apply the same argument as above. This implies that $p_1^e(\alpha)$ is a strictly increasing function for all $\alpha \in [0, \tilde{\alpha})$ and constant in α for all $\alpha \geq \tilde{\alpha}$, which implies that $p_1^e(\alpha)$ is monotonic for $\alpha \in [0,1]$ if $\epsilon_D < \epsilon_W$ at α_0.

Observe that the argument does not depend on the specific value of α_0 and the statements are true for any $\alpha_0 \in [0,1]$. Therefore, $p_1^e(\alpha)$ must be monotonic in α. The argument for $\tilde{p}_1^e(\alpha)$ and $\tilde{p}_1^n(\alpha)$ follows immediately by replacing $p_1^e(\alpha)$. $\square$

2.A.4 Equilibrium

Lemma 2.A.3 characterizes the Nash equilibria in pure strategies. We define the unique, implicit profit threshold $\hat{\alpha}$ such that $\pi^n = \tilde{\pi}^e$ when $\alpha = \hat{\alpha}$. When $\epsilon_D > \epsilon_W$, we can also define the unique, implicit threshold $\bar{\alpha}$ such that $\tilde{\pi}^n = \pi^e$ when $\alpha = \bar{\alpha}$.[38]

Lemma 2.A.3 (Equilibrium)**.**

(a) Suppose $\epsilon_D > \epsilon_W$.

- *(i) If* $\alpha < \min\{\bar{\alpha}, \hat{\alpha}\}$, *then both firms do not exploit and set* $p_1^* = p_1^n$ *and* $p_2^* = W(v_2)$.
- *(ii) If* $\alpha > \max\{\bar{\alpha}, \hat{\alpha}\}$, *then both firms exploit and set* $p_1^* = p_1^e(\alpha)$ *and* $p_2^* = W(v_2, \tilde{\Delta})$.
- *(iii) If* $\bar{\alpha} < \alpha < \hat{\alpha}$, *then either both firms do not exploit or both firms exploit.*
- *(iv) If* $\hat{\alpha} < \alpha < \bar{\alpha}$, *then firm* j *does not exploit and sets* $p_{1,j}^* = \tilde{p}_1^n(\alpha)$ *and* $p_{2,j}^* = W(v_2)$, *and firm* $-j$ *exploits and sets* $p_{1,-j}^* = \tilde{p}_1^e(\alpha)$ *and* $p_{2,-j}^* = W(v_2, \tilde{\Delta})$.

(b) Suppose $\epsilon_D < \epsilon_W$.

[38] When $\epsilon_D < \epsilon_W$, then both profits, $\tilde{\pi}^n$ and π^e, are strictly increasing in α.

(i) If $\alpha < \hat{\alpha}$, then only symmetric equilibria exist.

(ii) If $\alpha > \hat{\alpha}$, then symmetric exploiting and asymmetric equilibria exist.

In the case of (a), $\epsilon_D > \epsilon_W$, the symmetric non-exploiting equilibrium in (i) and the symmetric exploiting equilibrium in (ii) are unique.[39] In (iii), the best response of a firm is to do the same as the rival, and in (iv), the best response is to do the opposite.[40] Thus, for intermediate values of α, we observe either multiple symmetric equilibria or multiple asymmetric equilibria.

In part (b), when $\epsilon_D < \epsilon_W$, we observe a similar pattern of equilibria, but we cannot characterize when a unique symmetric equilibrium emerges. For a low share of behavioral consumers, (i), either both firms do not exploit (when $\tilde{\pi}^n > \pi^e$) or there exists multiple symmetric equilibria like in case $(aiii)$. For a large α, (ii), either both firms exploit, or an asymmetric outcome emerges like in case (aiv).

2.A.5 Proof of Lemma 2.A.3

We will first prove two intermediate result.

Lemma 2.A.4 (Unique thresholds)**.**

(i) The critical threshold $\hat{\alpha}$ is the unique solution to $\pi^n = \tilde{\pi}^e$ and $\alpha < \hat{\alpha} \Leftrightarrow \pi^n > \tilde{\pi}^e$.

(ii) Suppose $\epsilon_D > \epsilon_W$. The critical threshold $\bar{\alpha}$ is the unique solution to $\tilde{\pi}^n = \pi^e$ and $\alpha < \bar{\alpha} \Leftrightarrow \tilde{\pi}^n > \pi^e$.

Proof. (i) By Lemma 2.A.1, π^n is constant in α and $\tilde{\pi}^e$ is strictly increasing in α. Thus, there exists a unique solution solved for α such that $\pi^n = \tilde{\pi}^e$ and $\alpha < \hat{\alpha} \Leftrightarrow \pi^n > \tilde{\pi}^e$.

[39] If $\pi^n > \tilde{\pi}^e$ and $\tilde{\pi}^n > \pi^e$, non-exploiting is the dominant strategy for both firms. Similarly, if $\pi^n < \tilde{\pi}^e$ and $\tilde{\pi}^n < \pi^e$, then exploiting is the dominant strategy.

[40] Lemma 2.A.3 $(a)(iii)$ also applies, when $\bar{\alpha} = \alpha < \hat{\alpha}$ or $\bar{\alpha} < \alpha = \hat{\alpha}$. When $\hat{\alpha} = \alpha < \bar{\alpha}$, then, next to the asymmetrica equilibria, there exist also the symmetric non-exploiting equilibrium. Similarly, when $\hat{\alpha} < \alpha = \bar{\alpha}$, then, next to the asymmetrica equilibria, there exist also the symmetric exploiting equilibrium. In the special case of $\alpha = \hat{\alpha} = \bar{\alpha}$, any strategy is optimal since $\pi^n = \pi^e = \tilde{\pi}^n = \tilde{\pi}^e$.

(ii) When $\epsilon_D > \epsilon_W$, then, by Lemma 2.A.1, π^e is strictly increasing in α and $\tilde{\pi}^n$ is decreasing in α. Thus, there exists a unique solution solved for α such that $\tilde{\pi}^n = \pi^e$ and $\alpha < \bar{\alpha} \Leftrightarrow \tilde{\pi}^n > \pi^e$.

□

Lemma 2.A.5 (Dominant strategies)**.**

(i) Non-exploiting is the dominant strategy for both firms if $\pi^n > \tilde{\pi}^e$ and $\tilde{\pi}^n > \pi^e$.

(ii) Exploiting is the dominant strategy for both firms if $\pi^n < \tilde{\pi}^e$ and $\tilde{\pi}^n < \pi^e$.

Proof. (i) First, suppose that firm $-j$ does not exploit. The best response of firm j is to not exploit since $\pi^n > \tilde{\pi}^e$. Now suppose that firm $-j$ does exploit. The best response of firm j is to not exploit since $\tilde{\pi}^n > \pi^e$. Hence, in any case, the best response is to not exploit and thus, the dominant strategy. The best response of firm $-j$ is similarly.

(ii) First, suppose that firm $-j$ does not exploit. The best response of firm j is to exploit since $\pi^n < \tilde{\pi}^e$. Now suppose that firm $-j$ does exploit. The best response of firm j is to exploit since $\tilde{\pi}^n < \pi^e$. Hence, in any case, the best response is to exploit and thus the dominant strategy. The best response of firm $-j$ is similarly.

□

Now, we can proof the statements in Lemma 2.A.3.

(a) (i) By Lemma 2.A.4, we have $\pi^n > \tilde{\pi}^e$ and $\tilde{\pi}^n > \pi^e$ if $\alpha < \min\{\bar{\alpha}, \hat{\alpha}\}$. Hence, by Lemma 2.A.5, it is optimal for both firms to not exploit behavioral consumers, and set $p_1^* = p_1^n$ and $p_2^* = W(v_2)$.

(ii) By Lemma 2.A.4, we have $\pi^n < \tilde{\pi}^e$ and $\tilde{\pi}^n < \pi^e$ if $\alpha > \max\{\bar{\alpha}, \hat{\alpha}\}$. Hence, by Lemma 2.A.5, it is optimal for both firms to exploit behavioral consumers, and set $p_1^* = p_1^e(\alpha)$ and $p_2^* = W(v_2, \tilde{\Delta})$.

(iii) By Lemma 2.A.4, we have $\pi^n > \tilde{\pi}^e$ and $\tilde{\pi}^n < \pi^e$ if $\bar{\alpha} < \alpha < \hat{\alpha}$. Suppose that firm $-j$ does not exploit. The best response of firm j is to not exploit since $\pi^n > \tilde{\pi}^e$. Now suppose that firm $-j$ does exploit. The best response of firm j is to exploit since $\tilde{\pi}^n < \pi^e$. Hence, the best response of firm j is to do the same as firm $-j$. The best response of firm $-j$ is similarly. Thus, there exists two Nash equilibria in pure strategies {(not exploit, not exploit),(exploit,exploit)}.

(iv) By Lemma 2.A.4, we have $\pi^n < \tilde{\pi}^e$ and $\tilde{\pi}^n > \pi^e$ if $\hat{\alpha} < \alpha < \bar{\alpha}$. Suppose that firm $-j$ does not exploit. The best response of firm j is to exploit since $\pi^n < \tilde{\pi}^e$. Now suppose that firm $-j$ does exploit. The best response of firm j is to not exploit since $\tilde{\pi}^n > \pi^e$. Hence, the best response of firm j is to do the opposite as firm $-j$. The best response of firm $-j$ is similarly. Thus, there exists two Nash equilibria in pure strategies {(not exploit, exploit),(exploit, not exploit)}.

(b) (i) By Lemma 2.A.4, we have $\pi^n > \tilde{\pi}^e$. If $\tilde{\pi}^n > \pi^e$, then by Lemma 2.A.5, it is optimal for both firms to not exploit behavioral consumers. Thus, the unique symmetric non-exploiting equilibrium emerges. Otherwise, if $\tilde{\pi}^n < \pi^e$, case $(a)(iii)$ arises and the best response of firm j is to do the same as firm $-j$. Thus, multiple symmetric equilibria emerge. In either case, only symmetric equilibria exist.

(ii) By Lemma 2.A.4, we have $\pi^n < \tilde{\pi}^e$. If $\tilde{\pi}^n < \pi^e$, then by Lemma 2.A.5, it is optimal for both firms to exploit behavioral consumers. Thus, the unique symmetric exploiting equilibrium emerges. Otherwise, if $\tilde{\pi}^n > \pi^e$, case $(a)(iv)$ arises and the best response of firm j is to do the opposite as firm $-j$. Thus, multiple asymmetric equilibria emerge. The symmetric non-exploiting equilibrium does not exist if $\alpha > \hat{\alpha}$.

□

2.B Proofs Main Results

2.B.1 Proof of Lemma 2.3.1

By Lemma 2.A.1 and 2.A.2, there exists a unique solution $\bar{\alpha}_{p^k} \in \mathbb{R}$ to $p_1^n = p_1^k(\alpha)$ solved for α, where $p_1^k \in \{p_1^e, \tilde{p}_1^n, \tilde{p}_1^e\}$. First, suppose $\alpha = \bar{\alpha}_{\tilde{p}_1^n}$. Then, from Section 2.A,

$$\begin{aligned} p_1^n &= \tilde{p}_1^n(\bar{\alpha}_{\tilde{p}_1^n}) \\ \Leftrightarrow \quad \frac{-D(p_1^n, p_1^n)}{D'(p_1^n, p_1^n)} &= \frac{-D(\tilde{p}_1^n(\bar{\alpha}_{\tilde{p}_1^n}), \tilde{p}_1^e(\bar{\alpha}_{\tilde{p}_1^n}))}{D'(\tilde{p}_1^n(\bar{\alpha}_{\tilde{p}_1^n}), \tilde{p}_1^e(\bar{\alpha}_{\tilde{p}_1^n}))} \\ \Leftrightarrow \quad \frac{-D(p_1^n, p_1^n)}{D'(p_1^n, p_1^n)} &= \frac{-D(p_1^n, \tilde{p}_1^e(\bar{\alpha}_{\tilde{p}_1^n}))}{D'(p_1^n, \tilde{p}_1^e(\bar{\alpha}_{\tilde{p}_1^n}))} \end{aligned} \tag{2.4}$$

Consider the equation $\pi^n = \tilde{\pi}^n$, which must have a solution by Lemma 2.A.1.

$$\begin{aligned} \pi^n &= \tilde{\pi}^n \\ \Leftrightarrow \quad \frac{-D(p_1^n, p_1^n)^2}{D'(p_1^n, p_1^n)} &= \frac{-D(\tilde{p}_1^n(\bar{\alpha}_{\tilde{p}_1^n}), \tilde{p}_1^e(\bar{\alpha}_{\tilde{p}_1^n}))^2}{D'(\tilde{p}_1^n(\bar{\alpha}_{\tilde{p}_1^n}), \tilde{p}_1^e(\bar{\alpha}_{\tilde{p}_1^n}))} \\ \Leftrightarrow \quad \frac{-D(p_1^n, p_1^n)^2}{D'(p_1^n, p_1^n)} &= \frac{-D(p_1^n, \tilde{p}_1^e(\bar{\alpha}_{\tilde{p}_1^n}))^2}{D'(p_1^n, \tilde{p}_1^e(\bar{\alpha}_{\tilde{p}_1^n}))} \\ \Leftrightarrow \quad D(p_1^n, p_1^n) &= D(p_1^n, \tilde{p}_1^e(\bar{\alpha}_{\tilde{p}_1^n})), \end{aligned}$$

where the last equality follows from Equation (2.4), which holds only if $p_1^n = \tilde{p}_1^e(\bar{\alpha}_{\tilde{p}_1^n})$. Thus, when $\alpha = \bar{\alpha}_{\tilde{p}_1^n}$, it must be $p_1^n = \tilde{p}_1^e(\bar{\alpha}_{\tilde{p}_1^n})$. Since this must be unique, we have $\bar{\alpha}_{\tilde{p}_1^n} = \bar{\alpha}_{\tilde{p}_1^e}$, where

$$\begin{aligned} p_1^n &= \tilde{p}_1^e(\bar{\alpha}_{\tilde{p}_1^e}) \\ \frac{-D(p_1^n, p_1^n)}{D'(p_1^n, p_1^n)} - W(v_2) &= \frac{-[1 + \bar{\alpha}_{\tilde{p}_1^e} W'(v_2, \tilde{\Delta})] D(\tilde{p}_1^e(\bar{\alpha}_{\tilde{p}_1^e}), \tilde{p}_1^n(\bar{\alpha}_{\tilde{p}_1^e}))}{D'(\tilde{p}_1^e(\bar{\alpha}_{\tilde{p}_1^e}), \tilde{p}_1^n(\bar{\alpha}_{\tilde{p}_1^e}))} - \bar{\alpha}_{\tilde{p}_1^e} W(v_2, \tilde{\Delta}) \\ \bar{\alpha}_{\tilde{p}_1^e} &= \frac{W(v_2)}{W(v_2, \tilde{\Delta}) + \frac{W'(v_2, \tilde{\Delta}) D(p_1^n, p_1^n)}{D'(p_1^n, p_1^n)}} = \bar{\alpha}_{\tilde{p}_1^n}, \end{aligned}$$

and $\tilde{\Delta} = \beta_i \Delta(p_2, \tilde{p}_1^e(\bar{\alpha}_{\tilde{p}_1^e})) = \beta_i \Delta(p_2, p_1^n)$.

Now, suppose $\alpha = \bar{\alpha}_{p_1^e}$. Then,

$$
\begin{aligned}
p_1^n &= p_1^e(\bar{\alpha}_{p_1^e}) \\
\frac{-D(p_1^n, p_1^n)}{D'(p_1^n, p_1^n)} - W(v_2) &= \\
&\frac{-[1 + \bar{\alpha}_{p_1^e} W'(v_2, \tilde{\Delta})] D(p_1^e(\bar{\alpha}_{p_1^e}), p_1^e(\bar{\alpha}_{p_1^e}))}{D'(p_1^e(\bar{\alpha}_{p_1^e}), p_1^e(\bar{\alpha}_{p_1^e}))} - \bar{\alpha}_{p_1^e} W(v_2, \tilde{\Delta}) \\
\bar{\alpha}_{p_1^e} &= \frac{W(v_2)}{W(v_2, \tilde{\Delta}) + \frac{W'(v_2, \tilde{\Delta}) D(p_1^n, p_1^n)}{D'(p_1^n, p_1^n)}},
\end{aligned}
$$

and $\tilde{\Delta} = \beta_i \Delta(p_2, p_1^e(\bar{\alpha}_{p_1^e})) = \beta_i \Delta(p_2, p_1^n)$. Observe that $\bar{\alpha}_{p_1^e} = \bar{\alpha}_{\tilde{p}_1^n} = \bar{\alpha}_{\tilde{p}_1^e}$. Thus, we can define a single threshold

$$
\bar{\alpha}_p = \begin{cases} \dfrac{W(v_2)}{W(v_2, \tilde{\Delta}) + \frac{W'(v_2, \tilde{\Delta}) D(p_1^*, p_1^*)}{D'(p_1^*, p_1^*)}}, & \text{for } \epsilon_D \neq \epsilon_W \\ \infty, & \text{for } \epsilon_D = \epsilon_W, \end{cases}
$$

where $p_1^* \in \{p_1^n, p_1^e(\bar{\alpha}_p), \tilde{p}_1^n(\bar{\alpha}_p), \tilde{p}_1^e(\bar{\alpha}_p)\}$. Further, in the benchmark economy ($\alpha = 0$), only the non-exploiting strategy is possible, which implies $p_1^b = p_1^n$. Hence, $p_1^b = p_1^n = p_1^e(\bar{\alpha}_p) = \tilde{p}_1^n(\bar{\alpha}_p) = \tilde{p}_1^e(\bar{\alpha}_p)$.

(i) $\epsilon_D > \epsilon_W$ implies $\bar{\alpha}_p > 0$. Further, by Lemma 2.A.1, the prices $p_1^e(\alpha)$, $\tilde{p}_1^n(\alpha)$ and $\tilde{p}_1^e(\alpha)$ are decreasing in α when $\epsilon_D > \epsilon_W$ and $p_1^b = p_1^n$ are constant in α. Hence, for any $\alpha \in (\min\{\bar{\alpha}, \hat{\alpha}\}, \bar{\alpha}_p)$, it follows $p_1^k > p_1^b$, and for any $\alpha > \bar{\alpha}_p$ it follows $p_1^k < p_1^b$.

(ii) $\epsilon_D < \epsilon_W$ implies $\bar{\alpha}_p < 0$. By Lemma 2.A.1, the prices $p_1^e(\alpha)$, $\tilde{p}_1^n(\alpha)$ and $\tilde{p}_1^e(\alpha)$ are increasing in α when $\epsilon_D < \epsilon_W$ and $p_1^b = p_1^n$ are constant in α. Hence, for any $\alpha > 0 > \bar{\alpha}_p$, it follows $p_1^k > p_1^b$.

□

2.B.2 Proof of Proposition 2.3.1

Lemma 2.B.1. *Suppose $\epsilon_D > \epsilon_W$. The price threshold is larger than any profit threshold,* $\max\{\hat{\alpha}, \bar{\alpha}\} < \bar{\alpha}_p$.

Proof. Suppose $\alpha = \bar{\alpha}_p$. Hence, $p_1^n = p_1^e(\bar{\alpha}_p) = \tilde{p}_1^n(\bar{\alpha}_p) = \tilde{p}_1^e(\bar{\alpha}_p)$. It follows

$$\pi^n = \frac{-D(p_1^n, p_1^n)^2}{D'(p_1^n, p_1^n)} < \frac{-[1 + \alpha W'(v_2, \tilde{\Delta})]D(\tilde{p}_1^e(\bar{\alpha}_p), \tilde{p}_1^n(\bar{\alpha}_p))^2}{D'(\tilde{p}_1^e(\bar{\alpha}_p), \tilde{p}_1^n(\bar{\alpha}_p))} = \tilde{\pi}^e,$$

since $p_1^n = \tilde{p}_1^n(\bar{\alpha}_p) = \tilde{p}_1^e(\bar{\alpha}_p)$ and $W'(v_2, \tilde{\Delta}) > 0$. By Lemma 2.A.4, it must be $\alpha > \hat{\alpha}$ when $\pi^n < \tilde{\pi}^e$. Thus, $\bar{\alpha}_p > \hat{\alpha}$.

Further, we have

$$\tilde{\pi}^n = \frac{-D(\tilde{p}_1^n(\bar{\alpha}_p), \tilde{p}_1^e(\bar{\alpha}_p))^2}{D'(\tilde{p}_1^n(\bar{\alpha}_p), \tilde{p}_1^e(\bar{\alpha}_p))} < \frac{-[1 + \alpha W'(v_2, \tilde{\Delta})]D(p_1^e(\bar{\alpha}_p), p_1^e(\bar{\alpha}_p))^2}{D'(p_1^e(\bar{\alpha}_p), p_1^e(\bar{\alpha}_p))} = \pi^e,$$

since $p_1^e(\bar{\alpha}_p) = \tilde{p}_1^n(\bar{\alpha}_p) = \tilde{p}_1^e(\bar{\alpha}_p)$ and $W'(v_2, \tilde{\Delta}) > 0$. By Lemma 2.A.4, it must be $\alpha > \bar{\alpha}$ when $\tilde{\pi}^n < \pi^e$. Thus, $\bar{\alpha}_p > \bar{\alpha}$. Hence, $\max\{\hat{\alpha}, \bar{\alpha}\} < \bar{\alpha}_p$. □

We denote the utility a consumer receives from the base good with v_1. The surplus of a classical consumer in the benchmark economy ($\alpha = 0$) is given by $U_c = v_1 - p_1 + W(v_2) - p_2 = v_1 - p_1^b$ since $p_2 = W(v_2)$ in any benchmark (and symmetric non-exploiting) equilibrium. Hence, not consuming the add-on does not decrease the surplus of a classical consumer. A classical consumer benefits, compared to the benchmark, from the presence of behavioral consumers when $p_1^* < p_1^b$. Otherwise, when $p_1^* > p_1^b$, classical consumers are harmed.

(a) Since $p_1^n = p_1^b$ in any symmetric non-exploiting equilibrium, the surplus of a classical consumer is the same as in the benchmark. Hence, they are unaffected by the presence of behavioral consumers. Further, the market is unchanged since prices are identical to the benchmark.

(b) (i) By Lemma 2.A.3, there exists symmetric exploiting equilibria with $p_1^* = p_1^e(\alpha)$ for $\alpha > \min\{\bar{\alpha}, \hat{\alpha}\}$. By Lemma 2.3.1, we have $p_1^e(\alpha) > p_1^b$ for $\alpha \in (\min\{\bar{\alpha}, \hat{\alpha}\}, \bar{\alpha}_p)$, which reduces a classical consumer's surplus compared to the benchmark. Thus, classical consumers are harmed by the presence of behavioral consumers. Lemma 2.B.1 proofs that $\alpha \in (\min\{\hat{\alpha}, \bar{\alpha}\}, \bar{\alpha}_p)$ exists. By Lemma 2.3.1, we have $p_1^e(\alpha) < p_1^b$ for all $\alpha > \bar{\alpha}_p$, which increases a classical consumer's surplus compared to the benchmark. Thus, classical consumers benefit by the presence of behavioral consumers.

(*ii*) By Lemma 2.A.3, asymmetric equilibria exist only if $\hat{\alpha} < \alpha < \bar{\alpha}$. Therefore, by Lemma 2.B.1, we have $\alpha < \bar{\alpha}_p$ in any asymmetric equilibrium, which implies, by Lemma 2.3.1, $\tilde{p}_1^n(\alpha) > p_1^b$ and $\tilde{p}_1^e(\alpha) > p_1^b$. Hence, regardless from which firm classical consumers buy the base good, their surplus is lower compared to the benchmark. Thus, classical consumers are harmed by the presence of behavioral consumers in any asymmetric equilibrium.

(c) By Lemma 2.A.3, there exists symmetric exploiting equilibria and asymmetric equilibria. We have $p_1^* \in \{p_1^e(\alpha), \tilde{p}_1^n(\alpha), \tilde{p}_1^e(\alpha)\} > p_1^b$ for all α by Lemma 2.3.1. Hence, a classical consumer's surplus is lower compared to the benchmark in any symmetric exploiting equilibrium or asymmetric equilibrium. Thus, classical consumers are harmed by the presence of behavioral consumers.

□

2.B.3 Proof of Proposition 2.3.2

(a) The surplus of a behavioral consumer, when the behavioral effect does not increase the add-on utility, is given by $U_b = v_1 - p_1 + W(v_2) - p_2$, where v_1 denotes the gross utility received from the base good. The surplus of a behavioral consumer, when the behavioral effect increases the add-on utility, is given by $\tilde{U}_b = v_1 - p_1 + W(v_2, \tilde{\Delta}) - p_2$.

(i) The condition that behavioral consumers are worse off by exploitation is independent of whether U_b or $\tilde{U}_b$ applies:

$$\begin{aligned} & U_b^{NE} > U_b^E \\ \Leftrightarrow \quad & v_1 - p_1^n + W(v_2) - p_2^{NE} > v_1 - p_1^e(\alpha) + W(v_2) - p_2^E \\ \Leftrightarrow \quad & p_1^e(\alpha) > p_1^n + W(v_2) - W(v_2, \tilde{\Delta}(p_1^e(\alpha))) \end{aligned} \tag{2.5}$$

$$
\begin{aligned}
& \tilde{U}_b^{NE} > \tilde{U}_b^{E} \\
\Leftrightarrow \quad & v_1 - p_1^n + W(v_2, \tilde{\Delta}(p_1^n)) - p_2^{NE} \\
& > v_1 - p_1^e(\alpha) + W(v_2, \tilde{\Delta} p_1^e(\alpha))) - p_2^E \\
\Leftrightarrow \quad & p_1^e(\alpha) > p_1^n + W(v_2) - W(v_2, \tilde{\Delta}(p_1^n)) \qquad (2.6)
\end{aligned}
$$

Observe that the term $W(v_2, \tilde{\Delta})$ in (2.5) and (2.6) is different because in the former, the reference price is $p_1^e(\alpha)$ and in the latter p_1^n. However, we show that the condition is satisfied for any reference price. The condition $p_1^e(\alpha) > p_1^n + W(v_2) - W(v_2, \tilde{\Delta})$ holds for all $\alpha \in [0, 1]$. It is immediate to see that the condition is satisfied when $\alpha \leq \bar{\alpha}_p$, which implies $p_1^e(\alpha) \geq p_1^n$ by Lemma 2.3.1, and since $W(v_2) < W(v_2, \tilde{\Delta})$. For asymmetric equilibria, we just need to substitute $p_1^e(\alpha)$ with $\tilde{p}_1^n(\alpha)$ or $\tilde{p}_1^e(\alpha)$, respectively. The condition is always satisfied since asymmetric equilibria only exist for $\alpha < \bar{\alpha}_p$ and $\tilde{p}_1^n(\alpha) \geq p_1^n$ and $\tilde{p}_1^e(\alpha) \geq p_1^n$ when $\alpha \leq \bar{\alpha}_p$.

When $\alpha > \bar{\alpha}_p$, which implies $p_1^e(\alpha) < p_1^n$, we need an intermediate step. Consider the following inequality and observe

$$
\begin{aligned}
\frac{[1 + \alpha W'(v_2, \tilde{\Delta})] D(p_1^e(\alpha), p_1^e(\alpha))}{-D'(p_1^e(\alpha), p_1^e(\alpha))} &> \frac{D(p_1^n, p_1^n)}{-D'(p_1^n, p_1^n)} \qquad (2.7) \\
\alpha > \frac{\frac{D(p_1^n, p_1^n)}{-D'(p_1^n, p_1^n)} \frac{-D'(p_1^e(\alpha), p_1^e(\alpha))}{D(p_1^e(\alpha), p_1^e(\alpha))} - 1}{W'(v_2, \tilde{\Delta})} &= \frac{\frac{\epsilon_{D(e)}}{\epsilon_{D(n)}} - 1}{W'(v_2, \tilde{\Delta})} \\
\alpha \geq 0 &> \frac{\frac{\epsilon_{D(e)}}{\epsilon_{D(n)}} - 1}{W'(v_2, \tilde{\Delta})},
\end{aligned}
$$

where $D(s) = D(p_1^s, p_1^s)$ for $s = n, e$. The last inequality follows from the fact that $\frac{\partial \epsilon_D}{\partial p_1} > 0$ when $\epsilon_D > \epsilon_W$ by the proof of Lemma 2.A.2. Thus, we have $\epsilon_{D(n)} > \epsilon_{D(e)}$ when $p_1^e(\alpha) < p_1^n$, which implies $\frac{\epsilon_{D(e)}}{\epsilon_{D(n)}} - 1 < 0$.

Now, we use the property of inequality (2.7) to show that $p_1^e(\alpha) < p_1^n + W(v_2) - W(v_2, \tilde{\Delta})$ never holds for $\alpha \in [0, 1]$.

$$
p_1^e(\alpha) < p_1^n + W(v_2) - W(v_2, \tilde{\Delta})
$$

$$\begin{aligned}
&\frac{[1+\alpha W'(v_2,\tilde{\Delta})]D(p_1^e(\alpha),p_1^e(\alpha))}{-D'(p_1^e(\alpha),p_1^e(\alpha))} - \alpha W(v_2,\tilde{\Delta}) < \\
&\frac{D(p_1^n,p_1^n)}{-D'(p_1^n,p_1^n)} - W(v_2,\tilde{\Delta}) \\
&\alpha > \underbrace{\frac{\frac{[1+\alpha W'(v_2,\tilde{\Delta})]D(p_1^e(\alpha),p_1^e(\alpha))}{-D'(p_1^e(\alpha),p_1^e(\alpha))} - \frac{D(p_1^n,p_1^n)}{-D'(p_1^n,p_1^n)}}{W(v_2,\tilde{\Delta})}}_{>0} + 1 > 1,
\end{aligned}$$

which is a contradiction for any $\alpha \in [0,1]$. Hence, it must be $p_1^e(\alpha) > p_1^n + W(v_2) - W(v_2,\tilde{\Delta})$ for any $\alpha \in [0,1]$, which implies that behavioral consumers are always better off in a non-exploiting equilibrium (or benchmark economy) than in an exploiting equilibrium.

(ii) Total consumer surplus is larger in a symmetric non-exploiting equilibrium when

$$\begin{aligned}
& CS^{NE} > CS^E(\alpha) - \alpha D(\cdot)[W(v_2,\tilde{\Delta}) - W(v_2)] \\
\Leftrightarrow\quad & \alpha D(\cdot)[W(v_2,\tilde{\Delta}) - W(v_2)] > CS^E(\alpha) - CS^{NE}
\end{aligned}$$

By Lemma 2.3.1 we have $p_1^e(\alpha) \geq p_1^n$ when $\alpha \leq \bar{\alpha}_p$, which implies $D(p_1^e(\alpha), p_1^e(\alpha)) \leq D(p_1^n, p_1^n)$. Thus, it must be $CS^E(\alpha) \leq CS^{NE}$ which implies $\alpha D(\cdot)[W(v_2,\tilde{\Delta}) - W(v_2)] > 0 \geq CS^E(\alpha) - CS^{NE}$ since $W(v_2,\tilde{\Delta}) > W(v_2)$.

(b) When $\epsilon_D < \epsilon_W$, then by Lemma 2.3.1 we have $p_1^e(\alpha) > p_1^n$. Following the argument of part (a), behavioral consumers are always better off in a non-exploiting equilibrium when $p_1^e(\alpha) > p_1^n$, which is the case when firms exploit.

Further, $p_1^e(\alpha) > p_1^n$ implies again $CS^E(\alpha) < CS^{NE}$ and thus, total consumer surplus is always lower under exploitation □

2.C Further Results

2.C.1 Monopoly and Perfect Competition

Suppose that base goods are perfectly differentiated, then each firm is a monopolist in its respective base good market. Further, suppose that $D(p_1)$ is strictly decreasing, twice continuously differentiable, $\lim_{p_1 \to \infty} D(p_1) = 0$ and satisfies $D(p_1)D''(p_1) < 2D'(p_1)^2$. Observe that the monopolist's maximization problem is similar to equation (2.1) without $p_{1,-j}$, and yields $\pi^n = \pi(p_1^n, W(v_2))$ when choosing the non-exploiting strategy and $\pi^e = \pi(p_1^e(\alpha), W(v_2, \tilde{\Delta}))$ when choosing the exploiting strategy. The profits and prices behave similarly to the symmetric outcomes with two firms. Therefore, we can directly apply Lemma 2.A.1 and Lemma 2.A.2, which implies that π^n is constant in α and π^e strictly increasing in α. Define the profit threshold $\hat{\alpha}$ such that $\pi^n = \pi^e$.

Lemma 2.C.1.

(i) If $\alpha < \hat{\alpha}$, then the monopolist does not exploit and sets $p_1^* = p_1^n$ and $p_2^* = W(v_2)$.

(ii) If $\alpha > \hat{\alpha}$, then the monopolist exploits and sets $p_1^* = p_1^e(\alpha)$ and $p_2^* = W(v_2, \tilde{\Delta})$.

Proof. The proof is analogous to the proof of Lemma 2.A.3. □

The remainder of the analysis is similar to the baseline model with two firms. The critical price threshold $\bar{\alpha}_p$ is unchanged. Therefore, Lemma 2.3.1 without asymmetric prices follows immediately. Further, analogous to Lemma 2.B.1, we have $\hat{\alpha} < \bar{\alpha}_p$. Thus, Proposition 2.C.1 (a) and (b) below follow and is analogous to Proposition 2.3.1 and 2.3.2.

Proposition 2.C.1 (Monopoly and perfect competition)**.**

(a) Under a monopolist, the presence of behavioral consumers harms classical consumers in any exploiting equilibrium except if $\alpha > \bar{\alpha}_p$ and

$\epsilon_D > \epsilon_W$. *Classical consumers are unaffected in any non-exploiting equilibrium.*

(b) *Behavioral consumers are worse off when a monopolist exploits them. For $\epsilon_D > \epsilon_W$, total consumer surplus is strictly lower when $\alpha D(\cdot)[W(v_2, \tilde{\Delta}) - W(v_2)] > CS^E(\alpha) - CS^{NE}$. The condition is always satisfied for $\alpha \leq \bar{\alpha}_p$. For $\epsilon_D < \epsilon_W$, total consumer surplus is strictly lower when a monopolist exploits.*

(c) *Classical consumers are never harmed by the presence of behavioral consumers under perfect price competition. Classical consumers benefit in any symmetric exploiting equilibrium and are unaffected in any symmetric non-exploiting equilibrium.*

Proof. (a) The proof is analogous to the proof to Proposition 2.3.1.

(b) The proof is analogous to the proof to Proposition 2.3.2.

(c) Firms must earn zero profits under perfect competition implying $p_{1,j}D(\cdot) = -p_2 Q(\cdot)$. Further, they must offer the lowest price given the zero profit constraint. Otherwise, firms would face zero demand. Thus, it must be $p_1^* = \min\{-W(v_2), -\alpha W(v_2, \tilde{\Delta})\}$. Hence, it is optimal to exploit behavioral consumers only if $\alpha W(v_2, \tilde{\Delta}) > W(v_2)$. The unique symmetric exploiting equilibrium exists if and only if $\alpha > \frac{W(v_2)}{W(v_2,\tilde{\Delta})}$. Otherwise, when $\alpha < \frac{W(v_2)}{W(v_2,\tilde{\Delta})}$, the unique symmetric non-exploiting equilibrium exists. In the benchmark economy with $\alpha = 0$, firms choose $p_2 = W(v_2)$ and $p_1^b = -W(v_2)$. Thus, in any symmetric non-exploiting equilibrium, firms set $p_1^n = p_1^b = -W(v_2)$ and classical consumers are unaffected by the presence of behavioral consumers. In any exploiting equilibrium, it must be $p_1^e = -\alpha W(v_2, \tilde{\Delta}) < -W(v_2) = p_1^b$. Hence, classical consumers have to pay strictly less in any exploiting equilibrium than in the benchmark and thus, benefit. Lastly, there exist no profitable deviations for firms. Changing p_2 leads to less add-on revenues and thus, a higher p_1 and zero base-good demand. Increasing

p_1 leads to zero demand and thus zero profits. Decreasing p_1 would lead to negative profits.

□

2.C.2 Proof of Proposition 2.4.1

(a) The surplus of a classical and behavioral consumer are characterized in the proof of Proposition 2.3.1 and 2.3.2, and given by U_c, U_b and $\tilde{U}_b$, respectively.

(i) Observe that $\frac{\partial p_1^e(\alpha)}{\partial \alpha} < 0$ when $\epsilon_D > \epsilon_W$ by Lemma 2.A.1. By definition, the ex-post equilibrium is identical to the ex-ante equilibrium when the policy is ineffective and firms still set $p_1^e(\alpha)$. Denote with α' the share of behavioral consumers ex-ante and with $\alpha'' < \alpha'$ the share ex-post. Since $\alpha'' < \alpha'$ and $\frac{\partial p_1^e(\alpha)}{\partial \alpha} < 0$, we have $p_1^e(\alpha'') > p_1^e(\alpha')$. Since $\frac{\partial W(v_2, \tilde{\Delta})}{\partial p_1} > 0$, the add-on price $p_2^* = W(v_2, \tilde{\Delta})$ also increases ex-post. Observe that the add-on surplus is unaffected for classical consumers and worse for behavioral consumers ex-post. The base-good surplus for any type is strictly lower ex-post since v_1 is unchanged and p_1 is strictly larger. Hence, behavioral and classical consumers are worse off by an ineffective policy. Educated consumers enjoy an increased add-on surplus from not buying anymore, $p_2 - W(v_2) = W(v_2, \tilde{\Delta}) - W(v_2)$. Hence, they benefit if $W(v_2, \tilde{\Delta}) - W(v_2) > p_1^e(\alpha'') - p_1^e(\alpha')$. Otherwise, they are worse off.

Note that if $\tilde{U}_b = v_1 - p_1 + W(v_2, \tilde{\Delta}) - p_2$ applies, the effect of an ineffective policy on behavioral and educated consumers is similar to that on classical consumers as all three types obtain zero add-on surplus before and after the policy.

(ii) We first prove the result for classical consumers. By definition, an effective policy leads to a non-exploiting equilibrium ex-post with $p_1^* = p_1^n = p_1^b$ and $p_2^* = W(v_2)$. The add-on surplus remains at zero since $p_2^* = W(v_2)$. Hence, a classical consumer benefits from an effective policy when $p_1^e(\alpha') > p_1^b$, which, by Proposition 2.3.1, is the case when the ex-ante equilibrium was harmful. Similarly, when the ex-ante equilibrium was beneficial, which implies $p_1^e(\alpha') < p_1^b$, then a classical consumer is worse off ex-post.

The condition that behavioral consumers benefit from an effective policy is independent of whether U_b or $\tilde{U}_b$ applies and is identical to the condition in the proof of Proposition 2.3.2. It must be $p_1^e(\alpha') > p_1^n + W(v_2) - W(v_2, \tilde{\Delta})$ for any $\alpha \in [0, 1]$, which implies that the surplus of behavioral consumers in the benchmark economy (or non-exploiting equilibrium) is always larger than in the exploiting equilibrium. Therefore, any effective policy benefits behavioral consumers.

Similarly, for educated consumers given our welfare specification. They benefit when

$$\begin{aligned} U_c^{NE} &> U_b^E(\alpha') \\ v_1 - p_1^n + W(v_2) - p_2 &> v_1 - p_1^e(\alpha') + W(v_2) - p_2 \\ p_1^e(\alpha') &> p_1^n + W(v_2) - W(v_2, \tilde{\Delta}), \end{aligned}$$

which is the same condition as for behavioral consumers.

Note that if $\tilde{U}_b$ applies, then educated consumers benefit only if $p^e(\alpha') > p_1^n$ because the add-on surplus is zero ex-ante and ex-post. Education reduces the perceived utility of the add-on from $W(v_2, \tilde{\Delta})$ to $W(v_2)$. In this case, the condition that educated consumers benefit from an effective policy is identical to the one of classical consumers.

(b) When $\epsilon_D > \epsilon_W$, then by Lemma 2.A.1, $p_1^e(\alpha)$ is increasing in α. Thus, any decrease in α reduces base-good and add-on prices. Therefore, any policy must be beneficial for consumers.

$\square$

2.C.3 Proof of Proposition 2.4.2

(i) A binding price floor implies $p_1^* \in \{p_1^e(\alpha), \tilde{p}_1^n(\alpha), \tilde{p}_1^e(\alpha)\} < \underline{p}_1$. Since $\frac{\partial W(v_2, \tilde{\Delta})}{\partial p_1} > 0$ and $p_2^* = W(v_2, \tilde{\Delta})$ for any exploiting firm, the add-on price increases.

(ii) Since $D'(\cdot) < 0$ and $|D'(p_{1,j}, p_{1,-j})| > \left|\frac{\partial D_j(p_{1,j}, p_{1,-j})}{\partial p_{1,-j}}\right|$, $p_1^* < \underline{p}_1$ implies $D(p_{1,j}^*, p_{1,-j}^*) > D(\underline{p}_1, \max\{p_{1,-j}^*, \underline{p}_1\})$.

The surplus of a classical and behavioral consumer are characterized in the proof of Proposition 2.3.1 and 2.3.2, and given by U_c, U_b and $\tilde{U}_b$, respectively. We can observe immediately that all consumers remaining in the market are worse since they have to pay a higher p_1. Classical consumers who do not buy anymore are worse off since it must be $v_1 - p_1^* \geq 0 > v_1 - \underline{p}_1$. Behavioral consumers who do not buy anymore benefit if $U_b = v_1 - p_1^* + W(v_2) - W(v_2, \tilde{\Delta}) < 0$. Otherwise, they are harmed.

□

2.C.4 Sequential Buying

Observe that the equilibrium entails mixing of p_2 with searching consumers. Depending on the chosen strategy, setting the monopolistic add-on price $p_2 \in \{W(v_2), W(v_2, \tilde{\Delta})\}$ with probability 1 is not optimal. A firm can profitably deviate by setting a slightly lower add-on price and capture the add-on demand of all searching consumers. The expected profit of a non-exploiting firm is given by

$$\begin{aligned}\mathbb{E}\pi_j^n(p_{1,j}, p_{1,-j}, p_2^n) =\\ p_{1,j}D_j(\cdot) + (1-\rho)D_j(\cdot)p_2^n + \rho[1 - F^n(p_2^n)][D_j(\cdot) + D_{-j}(\cdot)]p_2^n.\end{aligned}$$

The term $1 - F^n(p_2^n)$ denotes the probability to set a lower add-on price than the competitor. The expected profit of an exploiting firm is given by

$$\begin{aligned}\mathbb{E}\pi_j^e(p_{1,j}, p_{1,-j}, p_2^e) =\\ p_{1,j}D_j(\cdot) + \alpha(1-\rho)D_j(\cdot)p_2^e + \alpha\rho[1 - F^e(p_2^e)][D_j(\cdot) + D_{-j}(\cdot)]p_2^e.\end{aligned}$$

Firms can always obtain positive add-on profits by selling the add-on to loyal consumers at $p_2 \in \{W(v_2), W(v_2, \tilde{\Delta})\}$, and earn $(1-\rho)D_j(\cdot)W(v_2)$ or $\alpha(1-\rho)D_j(\cdot)W(v_2, \tilde{\Delta})$, respectively, in the aftermarket. Therefore, firms must be indifferent between mixing and just selling to loyal consumers at the monopolistic price. We show in the proof of Lemma 2.6.1 below that $F^n(W(v_2)) = 1$ and $F^e(W(v_2, \tilde{\Delta})) = 1$. This allows us to rewrite the expected profits accordingly

$$\mathbb{E}\pi_j^n(p_{1,j}, p_{1,-j}, W(v_2)) = [p_{1,j} + (1-\rho)W(v_2)]D_j(\cdot),$$
$$\mathbb{E}\pi_j^e(p_{1,j}, p_{1,-j}, W(v_2, \tilde{\Delta})) = [p_{1,j} + \alpha(1-\rho)W(v_2, \tilde{\Delta})]D_j(\cdot).$$

Observe that the maximization problems are similar to the baseline model and identical when $\rho = 0$. Thus, we can proceed like in the baseline model and derive the base-good prices and profits in the three different outcomes. The results of Lemma 2.A.1 and Lemma 2.A.2 are similar, we only need to adjust properly for the term $(1-\rho)$. Further, the equilibrium structure is identical to Lemma 2.A.3, with the only difference that firms mix over p_2 instead of setting an add-on price with probability 1, which we will prove below. The result of Lemma 2.3.1 is unchanged and we still have $\max\{\hat{\alpha}, \bar{\alpha}\} < \bar{\alpha}_p$. Therefore, Propositions 2.3.1 and 2.3.2 follow immediately. The derivations and proofs are available on request.

Proof of Lemma 2.6.1

(i) A non-exploiting firm must be indifferent between mixing over p_2 and setting $p_2 = W(v_2)$. Thus, we can derive the equilibrium price distribution $F^n(p_2^n)$

$$\mathbb{E}\pi_j^n(p_{1,j}, p_{1,-j}, p_2^n) = \mathbb{E}\pi_j^n(p_{1,j}, p_{1,-j}, W(v_2))$$
$$(1-\rho)D_j(\cdot)p_2^n + \rho[1 - F^n(p_2^n)][D_j(\cdot) + D_{-j}(\cdot)]p_2^n = (1-\rho)D_j(\cdot)W(v_2)$$
$$F^n(p_2^n) = 1 - \frac{(1-\rho)D_j(\cdot)[W(v_2) - p_2^n]}{\rho[D_j(\cdot) + D_{-j}(\cdot)]p_2^n}.$$

The upper bound is given by $W(v_2)$

$$F^n(W(v_2)) = 1 - \frac{(1-\rho)D_j(\cdot)[W(v_2) - W(v_2)]}{\rho[D_j(\cdot) + D_{-j}(\cdot)]W(v_2)} = 1.$$

Set $F^n(p_2^n) = 0$ to obtain the lower bound $\underline{p}_2^n$

$$\begin{aligned} F^n(\underline{p}_2^n) &= 0 \\ (1-\rho)D_j(\cdot)[W(v_2) - \underline{p}_2^n] &= \rho[D_j(\cdot) + D_{-j}(\cdot)]\underline{p}_2^n \\ \underline{p}_2^n &= \frac{(1-\rho)D_j(\cdot)W(v_2)}{D_j(\cdot) + \rho D_{-j}(\cdot)}. \end{aligned}$$

We can easily verify that $\mathbb{E}\pi_j^n(p_{1,j}, p_{1,-j}, \underline{p}_2^n) = \mathbb{E}\pi_j^n(p_{1,j}, p_{1,-j}, W(v_2))$, which implies that firms obtain the same expected profit for all prices on the equilibrium support. The price distribution $F^n(p_2^n)$ is continuous and atomless since $D(\cdot)$ is continuous, $W(v_2)$ is constant, and $\frac{\partial F^n(p_2^n)}{\partial p_2^n} > 0$. For a detailed proof see Baye and Morgan (2001).

(ii) The proof is analogous to part (i). We simply have to replace p_2^n with p_2^e and $W(v_2)$ with $W(v_2, \tilde{\Delta})$. Note that an exploiting firm must set an add-on price $p_2^e > W(v_2)$. Therefore, the lower bound is given by $\max\{\underline{p}_2^e, W(v_2)\}$. It is easily verifiable that $\mathbb{E}\pi_j^e(p_{1,j}, p_{1,-j}, W(v_2)) = \mathbb{E}\pi_j^e(p_{1,j}, p_{1,-j}, W(v_2, \tilde{\Delta}))$ when $\underline{p}_2^e < W(v_2)$.

□

2.C.5 Unit Demand

We use a Hotelling model to analyze the unit demand case with classical and behavioral consumers, which are uniformly distributed on the interval $[0, 1]$. Consumers buy at most one unit of the base good with valuation v_1 at price p_1. We suppose that v_1 is sufficiently large. Two firms are located at each extreme, $l \in \{0, 1\}$. They sell identical main products and add-ons, and produce at similar marginal costs c and zero, respectively. Without loss of generality, assume that firm j is located at $l = 0$ and firm $-j$ at $l = 1$. Buying a good imposes transportation costs t on the consumer. The rest of the setup is identical to the baseline model in Section 2.2, but we use an explicit WTP function $W(v_2, \tilde{\Delta}) = v_2\beta_i(1 + p_1 - p_2)$ with $\beta_i \in \{0, 1\}$.

Aftermarket

In the last stage, after buying the base good, consumers can buy an add-on with valuation v_2 at price p_2. A classical consumer ($\beta = 0$) buys the add-on when $v_2 \geq p_2$ and a behavioral consumer ($\beta = 1$) buys when $\frac{v_2(1+p_1)}{1+v_2} \geq p_2$. Similar to the baseline model, firms extract the entire rent and choose $p_2^* \in \{v_2, \frac{v_2(1+p_1)}{1+v_2}\}$ in equilibrium. Therefore, the add-on demand is given by $Q_j(p_{2,j}, D_j(p_{1,j}, p_{1,-j})) = \{D_j(\cdot), \alpha D_j(\cdot)\}$.

Firm's Problem

The base-good demand of either firm is determined by the indifferent consumer $\bar{x}$, who is located at $\bar{x} = \frac{1}{2} + \frac{p_{1,-j}-p_{1,j}}{2t}$. The demand and profit functions of firm j are given by

$$D_j(p_{1,j}, p_{1,-j}) = \bar{x} = \frac{1}{2} + \frac{p_{1,-j} - p_{1,j}}{2t},$$
$$\pi_j(p_{1,j}, p_{1,-j}, p_{2,j}) = \big[p_{1,j} - c\big] \left[\frac{1}{2} + \frac{p_{1,-j} - p_{1,j}}{2t}\right] + Q_j(p_{2,j}, D_j(.))p_{2,j}.$$

The base-good prices and firm profits in the symmetric non-exploiting and symmetric exploiting outcome are given by

$$p_1^n = t + c - v_2, \qquad \pi^n = \pi(p_1^n, p_1^n, v_2) = \frac{t}{2}$$
$$p_1^e = t + \frac{c - \frac{\alpha v_2}{1+v_2}}{1 + \frac{\alpha v_2}{1+v_2}}, \quad \pi^e = \pi\left(p_1^e, p_1^e, \frac{v_2(1+p_1^e)}{1+v_2}\right) = \frac{t}{2}\left(1 + \frac{\alpha v_2}{1+v_2}\right)$$

We can observe immediately that $\pi^e > \pi^n$ for all $\alpha > 0$.[41] The reason for this is the covered market assumption, which is often used in Hotelling models. However, possible asymmetric strategies enable the existence of symmetric non-exploiting equilibria.

[41] It can be shown that the introduction of behavioral consumers does not affect the optimal location of a firm.

The prices and profits in asymmetric outcomes are

$$\tilde{p}_1^n = t + \frac{2(c - v_2)}{3} + \frac{c - \frac{\alpha v_2}{1+v_2}}{3(1 + \frac{\alpha v_2}{1+v_2})}$$

$$\tilde{p}_1^e = t + \frac{c - v_2}{3} + \frac{2(c - \frac{\alpha v_2}{1+v_2})}{3(1 + \frac{\alpha v_2}{1+v_2})}$$

$$\tilde{\pi}^n = \pi\left(\tilde{p}_1^n, \tilde{p}_1^e, v_2\right) = \frac{1}{2t}\left[t + \frac{v_2(1 + v_2 + \alpha(v_2 - 1 - c))}{3(1 + v_2(1+\alpha))}\right]^2$$

$$\tilde{\pi}^e = \pi\left(\tilde{p}_1^e, \tilde{p}_1^n, \frac{v_2(1+\tilde{p}_1^e)}{1+v_2}\right) = \frac{\left[t(1 + v_2(1+\alpha)) - \frac{1}{3}v_2(1 + v_2 + \alpha(v_2 - 1 - c))\right]^2}{2t(1+v_2)(1 + v_2(1+\alpha))}.$$

Note that demands under asymmetric strategies can be negative. We focus on interior solutions and assume that $D(\tilde{p}_1^n, \tilde{p}_1^e) > 0$ and $D(\tilde{p}_1^e, \tilde{p}_1^n) > 0$.

2.C.6 Equilibrium

The equilibria characterization is similar to Lemma 2.A.3.[42]

Lemma 2.C.2.

(i) If $\alpha < \min\{\bar{\alpha}, \hat{\alpha}\}$, then both firms do not exploit and set $p_1^* = p_1^n$ and $p_2^* = v_2$.

(ii) If $\alpha > \max\{\bar{\alpha}, \hat{\alpha}\}$, then both firms exploit and set $p_1^* = p_1^e$ and $p_2^* = \frac{v_2(1+p_1^e)}{1+v_2}$.

(iii) If $\bar{\alpha} < \alpha < \hat{\alpha}$, then either both firms do not exploit symmetrically or both firms exploit symmetrically.

(iv) If $\hat{\alpha} < \alpha < \bar{\alpha}$, then firm j does not exploit and sets $p_1^* = \tilde{p}_1^n$ and $p_2^* = v_2$, and firm $-j$ exploits and sets $p_1^* = \tilde{p}_1^e$ and $p_2^* = \frac{v_2(1+p_1^e)}{1+v_2}$.

Proof. The proof is analogeous to the proof of Lemma 2.A.3. Note that $\frac{\partial \tilde{\pi}^n}{\partial \alpha} < 0$. Thus, the threshold $\bar{\alpha}$ exists. Further, we have $\pi^n > \tilde{\pi}^e$ and $\tilde{\pi}^n > \pi^e$ when $\alpha = 0$. Since $\frac{\partial \pi^n}{\partial \alpha} = 0$, $\frac{\partial \pi^e}{\partial \alpha} > 0$, $\frac{\partial \tilde{\pi}^e}{\partial \alpha} > 0$, and $\frac{\partial \tilde{\pi}^n}{\partial \alpha} < 0$, the thresholds $\hat{\alpha}$ and $\bar{\alpha}$ must be unique. □

[42]If $D(\tilde{p}_1^n, \tilde{p}_1^e) = 0$ or $D(\tilde{p}_1^e, \tilde{p}_1^n) = 0$, only symmetric equilibria exists.

The critical price threshold is given by $\bar{\alpha}_p = \frac{1+v_2}{1+c-v_2}$. This leads to the following result similar to Lemma2.3.1.

Lemma 2.C.3.

(i) Suppose $1 + c > v_2$. If $\alpha \in (\min\{\bar{\alpha}, \hat{\alpha}\}, \bar{\alpha}_p)$, then the base good is more expensive in any symmetric exploiting or asymmetric equilibrium than in the benchmark. If $\alpha > \bar{\alpha}_p$, then the base good is cheaper in any symmetric exploiting equilibrium.

(ii) Suppose $1 + c < v_2$. The base good is always more expensive in any symmetric exploiting or asymmetric equilibrium than in the benchmark.

Proof. (i)

$$\begin{aligned} & \alpha < \frac{1+v_2}{1+c-v_2} = \bar{\alpha}_p \\ \Leftrightarrow \quad & c + \alpha v_2(1+c-v_2) < c + v_2(1+v_2) \\ \Leftrightarrow \quad & (c-v_2)\left(1+\frac{\alpha v_2}{1+v_2}\right) < c - \frac{\alpha v_2}{1+v_2} \\ \Leftrightarrow \quad & t + c - v_2 < t + \frac{c - \frac{\alpha v_2}{1+v_2}}{1+\frac{\alpha v_2}{1+v_2}} \\ \Leftrightarrow \quad & p_1^b < p_1^e \end{aligned}$$

$$\begin{aligned} & \alpha < \frac{1+v_2}{1+c-v_2} \\ \Leftrightarrow \quad & \frac{1}{3}(c-v_2) < \frac{c-\frac{\alpha v_2}{1+v_2}}{3(1+\frac{\alpha v_2}{1+v_2})} \\ \Leftrightarrow \quad & t + c - v_2 < t + \frac{2(c-v_2)}{3} + \frac{c-\frac{\alpha v_2}{1+v_2}}{3(1+\frac{\alpha v_2}{1+v_2})} \\ \Leftrightarrow \quad & p_1^b < \tilde{p}_1^n \end{aligned}$$

$$\begin{aligned} & \alpha < \frac{1+v_2}{1+c-v_2} \\ \Leftrightarrow \quad & 2(c-v_2) < \frac{2(c-\frac{\alpha v_2}{1+v_2})}{3(1+\frac{\alpha v_2}{1+v_2})} \\ \Leftrightarrow \quad & t + c - v_2 < t + \frac{c-v_2}{3} + \frac{2\left(c-\frac{\alpha v_2}{1+v_2}\right)}{3(1+\frac{\alpha v_2}{1+v_2})} \end{aligned}$$

$$\Leftrightarrow \quad p_1^b < \tilde{p}_1^e$$

(ii)

$$\begin{aligned} p_1^b &< p_1^e \\ t + c - v_2 &< t + \frac{c - \frac{\alpha v_2}{1+v_2}}{1 + \frac{\alpha v_2}{1+v_2}} \\ \alpha(1 + c - v_2) &< 1 + v_2 \\ \alpha > 0 &> \frac{1 + v_2}{1 + c - v_2} \end{aligned}$$

Since $1 + c - v_2 < 0$, the direction of inequality reverses when dividing. The proof for $p_1^b < \tilde{p}_1^n$ and $p_1^b < \tilde{p}_1^e$ when $1 + c - v_2 < 0$ is analogeous.

□

Similar to Lemma 2.B.1, the price threshold is always larger than the profit thresholds when $1 + c - v_2 > 0$. When $1 + c - v_2 < 0$, then $\bar{\alpha}_p < 0$, which corresponds to the case of $\epsilon_D < \epsilon_W$.

Lemma 2.C.4. *Suppose* $1 + c > v_2$. *Then* $\max\{\hat{\alpha}, \tilde{\alpha}\} < \bar{\alpha}_p$.

Proof. Suppose $\alpha = \bar{\alpha}_p = \frac{1+v_2}{1+c-v_2}$. Then

$$\begin{aligned} \tilde{\pi}^e &> \pi^n \\ \frac{t(1+c)}{2(1+c-z)} &> \frac{t}{2} \\ 0 &> -\frac{tv_2}{2} \end{aligned}$$

By Lemma 2.C.2, it must be $\alpha > \hat{\alpha}$ when $\pi^n < \tilde{\pi}^e$. Thus, $\bar{\alpha}_p > \hat{\alpha}$.

Further, when $\alpha = \bar{\alpha}_p = \frac{1+v_2}{1+c-v_2}$, then

$$\begin{aligned} \pi^e &> \tilde{\pi}^n \\ \frac{t(1+c)}{2(1+c-z)} &> \frac{t}{2} \\ 0 &> -\frac{tv_2}{2} \end{aligned}$$

Note that $\pi^n = \tilde{\pi}^n$ and $\pi^e = \tilde{\pi}^e$ when $\alpha = \bar{\alpha}_p = \frac{1+v_2}{1+c-v_2}$. By Lemma 2.C.2, it must be $\alpha > \bar{\alpha}$ when $\tilde{\pi}^n < \pi^e$. Thus, $\bar{\alpha}_p > \bar{\alpha}$. Hence, $\max\{\hat{\alpha}, \tilde{\alpha}\} < \bar{\alpha}_p$. $\square$

Given the results of Lemma 2.C.2, Lemma 2.C.3 and Lemma 2.C.4, Proposition 2.C.2 follows immediately, which analogeous to the main finding in the baseline model stated by Proposition 2.3.1.

Proposition 2.C.2 (Unit demand)**.**

(a) Behavioral consumers do not affect the market in any symmetric non-exploiting equilibrium.

(b) Suppose $1 + c > v_2$. Then the presence of behavioral consumers: (i) harms classical consumers in any symmetric exploiting equilibrium if $\alpha < \bar{\alpha}_p$ and benefits otherwise, (ii) harms classical consumers in any asymmetric equilibrium.

(c) Suppose $1 + c < v_2$. Then the presence of behavioral consumers harms classical consumers in any symmetric exploiting or asymmetric equilibrium.

Proof. The proof is analogous to the proof of Proposition 2.3.1, where $1 + c > v_2$ corresponds to the case of $\epsilon_D > \epsilon_W$ and $1 + c < v_2$ corresponds to $\epsilon_D < \epsilon_W$. $\square$

2.C.7 Cheaper Base Good than Add-on

In the baseline model, we focused on the case that the behavioral mechanism affects the add-on WTP positively by restricting Δ to be positive. Let us now consider the opposite when the add-on is more expensive than the base good. Then, the behavioral mechanism decreases the add-on WTP. We suppose the same setup as in the baseline model but allow $\tilde{\Delta} = \beta_i \Delta(p_2, p_1)$ to be negative and focus on the case of cheap base goods and expensive add-ons such that $\Delta < 0$ in any equilibrium. Further, for simplicity, we consider monopolistic base good markets. Crucially, behavioral consumers now have a lower WTP for the add-on, which has several implications. The add-on demand is given by

$$Q(p_2, D(p_1)) = \begin{cases} D(p_1) & \text{if } p_2 \leq W(v_2, \tilde{\Delta}), \\ (1-\alpha)D(p_1) & \text{if } W(v_2, \tilde{\Delta}) < p_2 \leq W(v_2), \\ 0 & \text{if } p_2 > W(v_2). \end{cases}$$

Contrary to the baseline model, all consumers buy the add-on if it is priced at the WTP of behavioral consumers, while only the fraction $(1-\alpha)$ accepts the add-on offer when $p_2 = W(v_2)$. The profit function, adjusted for the monopoly case, is still given by Equation (2.1) and the firm chooses between the non-exploiting strategy ($p_2^* = W(v_2)$) and the exploiting strategy ($p_2^* = W(v_2, \tilde{\Delta})$). Due to the negative behavioral effect, exploiting implies now lowering the add-on price below the WTP of classical consumers and selling the add-on to all. The non-exploiting profit $\pi^n = \pi(p_1^n(\alpha), W(v_2))$ is strictly decreasing in α, while the exploiting profit $\pi^e = \pi(p_1^e, W(v_2, \tilde{\Delta}))$ is independent of the share of behavioral consumers. Thus, we can define the profit threshold $\alpha = \hat{\alpha} \Leftrightarrow \pi^n = \pi^e$. Similarly to Lemma 2.A.3, for a share below the threshold, the monopolist does not exploit behavioral consumers and sets $p_2^* = W(v_2)$. There are only a few behavioral consumers that do not buy the add-on. When α is sufficiently large, the monopolist selects the exploiting strategy as the missed revenue in the aftermarket would be too high otherwise.

Interestingly, none of the results with $\Delta < 0$ depend on the semi-elasticities ϵ_D and ϵ_W. The optimal non-exploiting price $p_1^n(\alpha)$ is strictly increasing in α. The optimal base good price, when all consumers purchase the add-on (p_1^e), is independent of the share of behavioral consumers, like in the baseline model. Further, the outcome of the benchmark economy with $\alpha = 0$ is now different from both, the exploiting and non-exploiting equilibrium. The base good is always the cheapest in the benchmark, $p_1^b < \min\{p_1^n(\alpha), p_1^e\}$. The simple reason for this is that firms in after-sales markets redistribute add-on earnings to lower the base good price to attract more consumers.[43]

[43]For this reason, $p_1^n(\alpha)$ is increasing in α as the add-on earnings decline with more behavioral

In the benchmark economy, all consumers purchase the add-on at the price $W(v_2) > W(v_2, \tilde{\Delta})$, which clearly yields higher add-on profits than in any (non-)exploiting equilibrium. This implies that even a few behavioral consumers already affect the economy. Because not everyone buys the add-on in the non-exploiting equilibrium, the base good becomes more expensive consequently. Hence, by not accepting the additional offer, behavioral consumers indirectly increase the base good price. In the exploiting equilibrium, all consumers purchase the add-on but at a lower price than in the benchmark economy. Therefore, in contrast to Proposition 2.3.1, the presence of behavioral consumers always affects classical consumers: When behavioral consumers with $\Delta < 0$ are present, in any equilibrium, the base good is more expensive than in the benchmark case.

Proposition 2.C.3 (Cheaper base good than add-on). *Suppose* $\Delta < 0$.

(i) The presence of behavioral consumers harms a classical consumer in any non-exploiting equilibrium for all $\alpha > 0$.

(ii) If $p_1^e - p_1^b > W(v_2) - W(v_2, \tilde{\Delta})$, *then the presence of behavioral consumers harms a classical consumer in any exploiting equilibrium. Otherwise, a classical consumer benefits.*

Importantly, classical consumers are harmed when the monopolist *does not exploit* behavioral consumers. In any non-exploiting equilibrium, classical consumers pay the same for the add-on as in the benchmark economy but strictly more for the base good when $\alpha > 0$. Thus, they are clearly worse off. The impact in an exploiting equilibrium is ambiguous. Compared to the benchmark, classical consumers have to pay more for the base good but less for the add-on. Which effect dominates determines whether classical consumers benefit or are harmed by the presence of behavioral consumers.

consumers in the population.

2.C.8 Proofs Cheaper Base Good than Add-on

The monopolist's maximization problems given a chosen strategy are

$$\begin{aligned}\max_{p_1} \pi^n(p_1, W(v_2)) &= \max_{p_1}[p_1 + (1-\alpha)W(v_2)]D(p_1),\\ \max_{p_1} \pi^e(p_1, W(v_2, \tilde{\Delta})) &= \max_{p_1}[p_1 + W(v_2, \tilde{\Delta})]D(p_1).\end{aligned}$$

Maximizing each expression with respect to p_1 yields the prices and profits given the monopolist exploits or not

$$\begin{aligned}p_1^n(\alpha) &= \frac{-D(p_1^n(\alpha))}{D'(p_1^n(\alpha))} - (1-\alpha)W(v_2),\\ \pi^n(p_1^n(\alpha), W(v_2)) &= \frac{-D(p_1^n(\alpha))^2}{D'(p_1^n(\alpha))},\\ p_1^e &= \frac{-[1 + W'(v_2, \tilde{\Delta})]D(p_1^e)}{D'(p_1^e)} - W(v_2, \tilde{\Delta}),\\ \pi^e(p_1^e, W(v_2, \tilde{\Delta})) &= \frac{-[1 + W'(v_2, \tilde{\Delta})]D(p_1^e)^2}{D'(p_1^e)}.\end{aligned}$$

In contrary to the baseline model, the non-exploiting base-good price and profit depend on α, while the exploiting base-good price and profit are independent of α.

Lemma 2.C.5.

(a) $p_1^n(\alpha)$ is strictly increasing in α and π^n is strictly decreasing in α.

(b) p_1^e and π^e are constant in α.

Proof. (a)

$$\begin{aligned}\frac{\partial p_1^n(\alpha)}{\partial \alpha} &= \frac{W(v_2)}{2 - \frac{D(p_1^n(\alpha))D''(p_1^n(\alpha))}{D'(p_1^n(\alpha))^2}} > 0\\ \frac{\partial \pi^n}{\partial \alpha} &= -D(p_1^n(\alpha))\left[2 - \frac{D(p_1^n(\alpha))D''(p_1^n(\alpha))}{D'(p_1^n(\alpha))^2}\right]\frac{\partial p_1^n(\alpha)}{\partial \alpha}\\ &= -D(p_1^n(\alpha))W(v_2) < 0.\end{aligned}$$

(b) Taking $\frac{\partial p_1^e}{\partial \alpha}$ and rearranging yields

$$\frac{\partial p_1^e}{\partial \alpha}\left[(1+W'(v_2,\tilde{\Delta}))\left(2-\frac{D(p_1^e)D''(p_1^e)}{D'(p_1^e)^2}\right)+\frac{W''(v_2,\Delta)D(p_1^e)}{D'(p_1^e)}\right]=0,$$
$$\frac{\partial \pi^e}{\partial \alpha}=-(1+W'(v_2,\tilde{\Delta}))D(p_1^e)\left[2-\frac{D(p_1^e)D''(p_1^e)}{D'(p_1^e)^2}\right]\underbrace{\frac{\partial p_1^e}{\partial \alpha}}_{=0}=0.$$

□

Since π^n is strictly decreasing and π^e is constant in α, we can define the profit threshold $\hat{\alpha}$ and characterize the equilibria.

Lemma 2.C.6.

(i) If $\alpha < \hat{\alpha}$, then the monopolist does not exploit and sets $p_1^* = p_1^n$ and $p_2^* = W(v_2)$.

(ii) If $\alpha > \hat{\alpha}$, then the monopolist exploits and sets $p_1^* = p_1^e(\alpha)$ and $p_2^* = W(v_2,\tilde{\Delta})$.

Proof. The proof is analogous to the proof to Lemma 2.A.3. □

The base good is always the cheapest in the benchmark, $p_1^b = \frac{-D(p_1^b)}{D'(p_1^b)} - W(v_2) < \min\{p_1^n(\alpha), p_1^e\}$.

Lemma 2.C.7. *$p_1^b < \min\{p_1^n(\alpha), p_1^e\}$ for all $\alpha > 0$.*

Proof. First, observe that $p_1^b = p_1^n(\alpha) \Leftrightarrow 0 = \alpha W(v_2)$ is feasible only when $\alpha = 0$. By Lemma 2.C.5, $p_1^n(\alpha)$ is strictly increasing in α. Hence, since p_1^b is independent of α, it must follow that $p_1^b < p_1^n(\alpha)$ when $\alpha > 0$.
We prove $p_1^b < p_1^e$ in several steps. First, observe that $p_1^b \neq p_1^e$ for all $\alpha \in \mathbb{R}$ because

$$\begin{aligned} & p_1^b = p_1^e \\ \Leftrightarrow\quad & -W(v_2) = \frac{-W'(v_2,\tilde{\Delta})D(p_1^e)}{D'(p_1^e)} - W(v_2,\tilde{\Delta}) \\ \Leftrightarrow\quad & \frac{W'(v_2,\tilde{\Delta})D(p_1^e)}{D'(p_1^e)} = W(v_2) - W(v_2,\tilde{\Delta}), \end{aligned}$$

which is a contradiction since $\frac{W'(v_2,\tilde{\Delta})D(p_1^b)}{D'(p_1^b)} < 0$ and $W(v_2) - W(v_2, \tilde{\Delta}) > 0$. Hence, since p_1^b and p_1^e are both constant in α, it must be either $p_1^b < p_1^e \; \forall\alpha$ or $p_1^b > p_1^e \; \forall\alpha$.

Next, observe that $p_1^e = p_1^n(\alpha)$ when

$$\alpha = \bar{\alpha}_p = 1 - \underbrace{\frac{W'(v_2, \tilde{\Delta})D(p_1^e)}{D'(p_1^e)W(v_2)}}_{<0} - \underbrace{\frac{W(v_2, \tilde{\Delta})}{W(v_2)}}_{<1} > 0.$$

Since $p_1^b < p_1^n(\alpha)$ for all $\alpha > 0$ and $p_1^e = p_1^n(\alpha)$ when $\alpha = \bar{\alpha}_p > 0$, it follows that $p_1^b < p_1^e$ for $\alpha \geq \bar{\alpha}_p$. But since p_1^b and p_1^e are both constant in α, it must be $p_1^b < p_1^e$ for any α. □

Now can we prove the statements in Proposition 2.C.3.

Proof of Proposition 2.C.3

The proof follows closely the proof of Proposition 2.3.1.

(i) By Lemma 2.C.6 and Lemma 2.C.7, we have $p_1^b < p_1^n(\alpha)$ and $p_2 = W(v_2)$ in any non-exploiting equilibrium. Hence, classical consumers pay the same as in the benchmark economy for the add-on, but strictly more for the base good, which reduces a classical consumer's surplus compared to the benchmark. Thus, classical consumers are harmed by the presence of behavioral consumers.

(ii) By Lemma 2.C.6 and Lemma 2.C.7, we have $p_1^b < p_1^e$ and $p_2 = W(v_2, \tilde{\Delta}) < W(v_2)$ in any exploiting equilibrium. Hence, compared to the benchmark, classical consumers pay strictly less ($W(v_2) - W(v_2, \tilde{\Delta}) > 0$) for the add-on and strictly more for the base good ($p_1^b - p_1^e < 0$). If $p_1^e - p_1^b > W(v_2) - W(v_2, \tilde{\Delta})$, the negative effect dominates, which reduces a classical consumer's surplus compared to the benchmark. Thus, classical consumers are harmed by the presence of behavioral consumers. Otherwise, if $p_1^e - p_1^b < W(v_2) - W(v_2, \tilde{\Delta})$, the positive effect dominates, which increases a classical consumer's surplus compared to the benchmark. Thus, classical consumers benefit by the presence of behavioral consumers.

Chapter 3

Complements, Merger Incentives and Context Effects

3.1 Introduction

We are interested in how context effects in purchasing decisions of complementary products affect vertical integration and market formation. Microsoft's recent $69 billion acquisition of the game publisher Activision Blizzard in October 2023 marks the largest merger in the gaming industry and the history of Microsoft.[1] This case gained a lot of attention and faced many difficulties, especially from regulators and competition authorities in the US, UK, and Europe.[2] While we observe a growing trend of mergers in the gaming market, there are less active industries with complementary products like traveling and accommodation. One reason could be that consumer preferences are unfavorable for mergers.

We consider an economy where consumers purchase complementary products and have non-standard preferences due to potential context effects, which distort their choice. An unusual price may attract consumers' attention, distracting them from the prices of other complements. For instance, a great deal on a gaming console catches consumers' attention, resulting in

[1]See CBO's report "Research and Development in the Pharmaceutical Industry", available at Statista.
[2]See "Microsoft's $69 Billion Activision Blizzard Acquisition Finally Approved", available at Forbes.

them being less price-sensitive for additional games. Alternatively, consumers may relate a cheap complement to the price of the expensive one. An expensive flight makes hotels look relatively cheap, letting consumers accept a higher price.

An extensive literature analyzes and documents decision-makers with context-dependent preferences, most notably salience effects.[3] In several papers, Bordalo et al. (2012, 2013); Bordalo, Gennaioli, and Shleifer (2016); Bordalo et al. (2020) provide a theoretical framework for salience supported by a large amount of experimental and empirical evidence (Dertwinkel-Kalt, Köhler, Lange, and Wenzel, 2017; Dertwinkel-Kalt and Köster, 2020; Dertwinkel-Kalt et al., 2020; Dessaint and Matray, 2017). Despite the popularity of context effects, markets with consumers holding context-dependent preferences are rarely studied. Yet, it is important to understand how markets with those consumers are shaped, especially for policy considerations: *"[...] from a theoretical perspective, a very important question to address is how salient thinking shapes strategic considerations of economic actors and the resulting equilibrium outcome"* (Herweg, Müller, and Weinschenk, 2018, p. 107). Firms may take advantage of behavioral consumers by designing specific contexts, for instance, making certain attributes more salient, implementing decoy effects, or using drip pricing.

This paper aims to shed light on the optimal pricing and merger incentives with firms selling complementary products. In particular, when competition is asymmetric across markets and consumers have context-dependent preferences that underweight the price of one complement. In the model, consumers must buy one unit of each product A and B to receive utility. In market A, two firms compete in a standard Hotelling model, while a monopolist offers product B. We first derive a benchmark economy without behavioral effects and consider vertical integration between the monopolist and one of the competitive firms. As usual, with vertical mergers, prices in the competitive market drop because double marginalization is reduced. The merged firm then recoups rents with a price increase for the monopolistic

[3] See Bordalo et al. (2022) for an overview. Relative thinking (Bushong et al., 2021; Somerville, 2022) and focusing (Kőszegi and Szeidl, 2013) also consider decision-makers with context-dependent preferences.

good. The merger is incentive-compatible when formation costs are not too large.

We then introduce context effects into the model. Consumers underweight the price of the non-salient product in their purchasing decisions. We analyze the case when the monopolistic good is salient, and thus, consumers underrate the price of the competitive good. This has crucial impacts on merger incentives. A price change in the competitive market affects demand less than in the benchmark, implying less intense competition. This effect harms the efficiency gains of vertical integration as the rival does not need to react as strongly as in the benchmark when the merged firm lowers its price. In fact, when context effects are sufficiently strong, the merged firm may even set a larger price than the rival. Because the vertically integrated firm cannot fully utilize its advantages from removing double marginalization, merger incentives are lower than in the benchmark case. They may actually be negative, implying that vertical integration does not occur even if formation costs are zero.

Literature Review. Several papers consider salience effects in market situations, though none of them looks at complementary goods or at merger incentives. Bordalo et al. (2013) show how consumers choose products depending on the salience of prices and qualities. Bordalo et al. (2016) then analyze how firms compete for consumer attention with salient attributes. Apffelstaedt and Mechtenberg (2021) study the optimal product line of a retailer when consumers are subject to context effects. Helfrich and Herweg (2017) and Dertwinkel-Kalt and Köster (2022) analyze a competitive retail market with a monopolistic upstream manufacturer. Both papers focus on how salience affects quality choice and online sales. Inderst and Obradovits (2023) consider hidden prices, while Inderst and Obradovits (2020) look at loss leading. Both study how salience impacts product quality in the respective setup. Yin, Jiang, and Zhou (2023) examine how context-dependent preferences affect a firm's optimal bundling strategy of two products. In contrast to our model, the monopolist is also active in the second market. Narasimhan and Turut (2013) show that in the presence of consumers with context-dependent preferences, competitors rather imitate rivals than horizontally differentiate when improving products. Chen and Turut (2013) reveal that context-dependent consumers influence whether existing technology should be improved or disruptive innovations should be explored.

Some papers about complementary products study vertical integration, but without context effects. Economides and Salop (1992) analyze the effects of vertical and horizontal mergers on prices, but do not look at merger incentives. Depending on the competition intensity, a total merger may lower or increase the price of the composite good. Economides (2005) finds negative incentives for vertical integration when product competition is intense because then, prices drop too much. In our model, incentives can be negative because competition becomes less intense with context effects, which harms the efficiency gains of vertical integration.[4] Heeb (2003) studies vertical integration in complementary markets in a similar setup to ours, in which a monopolist produces an essential good, and a duopoly offers complementary products. Depending on the valuation of the complementary good, different pricing regimes arise. In his model, vertical integration is always incentive-compatible. In a similar model, Akgün, Caffarra, Etro, and Stillman (2020) analyze the effects of merger on consumer surplus.

The contribution of this paper is twofold. It adds further understanding of how consumers with context-dependent preferences affect markets and provides an explanation of why vertical integration in some related markets may not occur.

3.2 Model

Consumers with taste x are uniformly distributed on $[0, 1]$ and buy at most one unit of a composite product A_iB. Firms $i \in \{1, 2\}$ with symmetric marginal costs $c_A > 0$ are located at the extreme points $l_i \in \{0, 1\}$ and offer the component A_i at price p_{Ai}. Component B is offered by a monopolist M

[4] Economides (2005) studies a model with four firms and several variants of integration, but not the particular case of our model.

with $c_B > 0$ at price p_B. Consumers buying the product A_iB obtain utility

$$v - p_{Ai} - p_B - t|x - l_i|, \tag{3.1}$$

where $v > c_A + c_B$ and $t > 0$ captures the transportation costs. We define $\bar{v} = v - c_A - c_B$ to ease notation.

Assumption 1. $t < \frac{\bar{v}}{2}$.

Assumption 1 guarantees that all consumers are served and the market is covered in equilibrium. In a covered market, the demand for firm i depends on the indifferent consumer buying A_1B or A_2B, $\hat{x} = \frac{t+p_{A2}-p_{A1}}{2t}$. The demands are $D_{A_1B} = \hat{x}$ and $D_{A_2B} = 1-\hat{x}$. The demand for the monopolist is fixed at 1. Further, suppose that firms 1 and 2 play symmetric strategies in equilibrium, $p^*_{A1} = p^*_{A2} = p^*_A$, implying equal demand, $\hat{x} = \frac{1}{2}$ in equilibrium. We solve the game for Nash equilibria in pure strategies.

3.3 Benchmark

We first analyze the classical setup without any context effects as a benchmark. We characterize the equilibria with three firms and with vertical integration in a covered market. Then, we analyze the incentives to merge. All proofs are in the Appendix.

3.3.1 Three firms

Any equilibrium that covers the market must satisfy $p^*_A + p^*_B = v - \frac{t}{2}$. This fulfills the participation constraint of the indifferent consumer $\hat{x}$ with equality. For any lower total price, the monopolist could increase p_B without losing demand, obtaining strictly larger profits. Any larger price leads to an uncovered market, which is not an equilibrium given Assumption 1.[5]

The minimum price a firm sets in equilibrium is $\underline{p}_j = c_j + \frac{t}{2}$, where $j \in \{A, B\}$. Either firm must obtain a mark-up of at least $\frac{t}{2}$. Otherwise, it

[5] We show this in the proof of Lemma 3.3.1.

can profitably deviate by increasing its price. This implies a maximum price of $\bar{p}_j = v - \frac{t}{2} - \underline{p}_{-j} = v - c_{-j} - t$. Observe that $\underline{p}_j < \bar{p}_j$ for $t < \frac{2\bar{v}}{3}$. In market A, however, competition is so intense that p_A^* is bounded from above by $c_A + t$.[6] Thus, in equilibrium, we have $p_A^* \in [c_A + \frac{t}{2}, c_A + t]$. We capture the markup a firm i sets on the competitive good by $\frac{t}{y-1}$, where $y \in [2, 3]$.[7] Lemma 3.3.1 characterizes the equilibrium structure.

Lemma 3.3.1. *In any pure-strategy equilibrium, firms set prices*

$$p_A^* = c_A + \frac{t}{y-1} \quad \textit{and} \quad p_B^* = v - c_A - \frac{t(1+y)}{2(y-1)},$$

where $y \in [2, 3]$. *Equilibrium profits are* $\pi_M^* = \bar{v} - \frac{t(1+y)}{2(y-1)}$ *and* $\pi_i^* = \frac{t}{2(y-1)}$.

The equilibrium intervals for the price of the monopolistic good and profits are:

$$p_B^* \in [v - c_A - \frac{3t}{2}, v - c_A - t], \quad \pi_M^* \in [\bar{v} - \frac{3t}{2}, \bar{v} - t] \quad \text{and} \quad \pi_i^* \in [\frac{t}{4}, \frac{t}{2}].$$

3.3.2 Vertical Intregration

Suppose that firm 1 and M merge. The new firm $\hat{M}$ produces both complements, A_1 and B, sells them individually, and chooses p_{A1} and p_B to maximize profits, while firm 2 produces only component A_2.[8] Assumption 1 still guarantees that the market is covered in equilibrium. Crucially, we cannot assume $p_{A1}^* = p_{A2}^*$ anymore. Thus, any equilibrium must satisfy $\frac{p_{A1}^* + p_{A2}^*}{2} + p_B^* = v - \frac{t}{2}$ or, equivalently, $p_B^* + p_{A1}^* = v - t\hat{x}$. We suppose that negative prices are not feasible. Further, when marginal costs are sufficiently large ($c_A \geq t$), the merged firm has the power to exclude firm 2 from the market by setting p_{A1} sufficiently low, $p_{A1} \leq p_{A2} - t$.[9] In this case, firm 2 is indifferent between any choice of p_{A2} because all

[6] If $p_A^* > c_A + t$, firm i can profitably deviate by lowering its price and capture demand from rival $-i$.

[7] While a more direct notation would be possible, this specification is more convenient once we analyze vertical integration.

[8] We do not consider bundling in this chapter. But note, when $\hat{M}$ cannot relocate, then bundling is not profitable given Assumption 1.

[9] Absent Assumption 1, the merged firm could also exclude by increasing p_B when the market is uncovered.

lead to zero demand (and profits), which implies the existence of multiple equilibria. Therefore, we restrict the equilibrium set by using the standard refinement that excludes equilibria in weakly dominated strategies. Then, in any admissible equilibrium, both firms are active as shown in Lemma 3.3.2 below.

Similar to the case with three firms, prices must be sufficiently high in equilibrium

$$\underline{p}_{A2} = c_A + \frac{t + p_{A1} - c_A}{3}, \quad \underline{p}_{A1} = c_A + \frac{t + p_{A2} - c_A}{3} - \frac{2}{3}(p_B^* - c_B),$$
$$\underline{p}_B = c_B + \frac{t + c_A - p_{A1}^*}{2}.$$

Because the merged firm can cross-subsidize with the revenue of the monopolistic product, the minimum price on the competitive product is lower. The upper bound for the price of the competitive product is given by $\bar{p}_{Ai} = c_A + \frac{t+p_{A-i}-c_A}{2}$. Thus, in equilibrium, firm 2 sets a markup of $\frac{t+p_{A1}-c_A}{y}$ with $y \in [2,3]$. Further, we find, given firm 2 is active, that the best response of the merged firm is to price the competitive good above marginal costs and impose half the markup that firm 2 does.

Lemma 3.3.2.

(i) In any admissible, pure-strategy equilibrium, both firms are active and set prices $p_{A1}^* = c_A + \frac{t}{2y-1}$, $p_{A2}^* = c_A + \frac{2t}{2y-1}$, *and* $p_B^* = v - c_A - \frac{t(1+y)}{2y-1}$, *where* $y \in [2,3]$. *The equilibrium profits are* $\pi_{\hat{M}}^* = \bar{v} - \frac{t(2y^2-1)}{(2y-1)^2}$ *and* $\pi_2^* = \frac{2t(y-1)}{(2y-1)^2}$.

(ii) When $p_{A2} \geq c_A + 2t$ *and* $t \leq c_A$, *there exists an inadmissible, pure strategy equilibrium, in which firm 2 is inactive and firm* $\hat{M}$ *sets prices* $p_{A1} \leq c_A - t$, $p_B = v - p_{A1} - t$, *and obtains* $\pi_{\hat{M}} = \bar{v} - t$.

The demand for A_1B in an admissible equilibrium is $\hat{x} = \frac{y}{2y-1}$ and for A_2B we have $1 - \hat{x} = \frac{y-1}{2y-1}$. This leads to $p_{A1}^* \in \left[c_A + \frac{t}{5}, c_A + \frac{t}{3}\right]$, $p_{A2}^* \in \left[c_A + \frac{2t}{5}, c_A + \frac{2t}{3}\right]$, $p_B^* \in \left[v - c_A - t, v - c_A - \frac{4t}{5}\right]$, $\pi_{\hat{M}}^* \in \left[\bar{v} - \frac{17t}{25}, \bar{v} - \frac{7t}{9}\right]$, and $\pi_2^* \in \left[\frac{4t}{25}, \frac{2t}{9}\right]$. In general, the merged firm prefers to have the rival in the market because it sells component B to consumers located far away. A price cut to cover the market alone affects profits stronger than the increased demand for A_1. In case (ii), though, firm 2 sets a price so large, that it is better for $\hat{M}$ to serve the market alone. This leads to zero demand and profits for firm 2 even if the merged firm sets $p^*_{A1} = c_A + \frac{t}{2y-1}$.[10] Because $\pi_2 \geq 0$ for $p_{A2} < c_A + 2t$, it follows that $p_{A2} \geq c_A + 2t$ is a weakly dominated strategy.

Further, we have $\pi^*_{\hat{M}} = \bar{v} - \frac{t(2y^2-1)}{(2y-1)^2} > \bar{v} - t$ for $p_{A2} < c_A + 2t$. Hence, in this case, the merged firm prefers the rival to be active. In other words, when $p_{A2} < c_A + 2t$, then excluding by setting $p_{A1} + p_B = v - t$ is strictly dominated by choosing $p^*_{A1} = c_A + \frac{t}{2y-1}$ and $p^*_B = v - c_A - \frac{t(1+y)}{2y-1}$ and the inadmissible equilibrium does not exist. Therefore, the only admissible equilibria in pure strategies, given our refinement, is when both firms are active and set prices as characterized by Lemma 3.3.2 (i). Note that when $t > c_A$, then the merged firm does not have the power to exclude firm 2 from the market. Hence, both firms are active in this case.

We observe that total industry profits are lower after the merger. As we show in the next section, this is not necessarily the case when consumers underweight the price of the competitive good. While the competitive products become cheaper, the monopolistic complement is more expensive after vertical integration. Overall, the composite product is cheaper at the merged firm and more expensive at firm 2 compared to the case with three firms.

Merger Incentives

Consider now the incentive to merge and suppose that integration incurs some fixed formation costs $F > 0$. The incentive constraint is given by $\pi_{\hat{M}} - F > \pi_1 + \pi_M$. Further, we suppose that y remains constant and is not affected by the merger.

Lemma 3.3.3. *Suppose vertical integration incurs fixed costs $F > 0$ and does not affect y. Then, merger occurs iff $F < \frac{t(3y-2)}{2(y-1)(2y-1)^2}$.*

[10] We obtain $\hat{x} = 1$ for $p^*_{A1} = c_A + \frac{t}{2y-1}$ and $p_{A2} \geq c_A + 2t$.

The benefit of integrating is $\Delta\pi = \pi_{\hat{M}} - \pi_1 - \pi_M \in [\frac{7t}{100}, \frac{2t}{9}]$, which is strictly decreasing in y and strictly increasing in t. Thus, merger incentives are stronger when competition is less intense, resembling the findings of Economides (2005).

3.4 Context Effects on the Competitive Good

We now suppose that consumers are subject to context effects. This distorts their decision-making as they weigh attributes differently depending on the context. For example, product B could be salient and attract consumers' attention.[11] This distracts consumers from the price of component A, causing them to spend in total more than planned (Bordalo et al., 2022). Formally, they weight p_{Ai} by $\delta \in (0, 1)$. Similar to Bordalo et al. (2016), behavioral consumers perceive the utility as

$$v - \delta p_{Ai} - p_B - t|x - l_i| \tag{3.2}$$

when observing the offer. We impose a rather strong assumption that salience is exogenously given and cannot be manipulated by firms. This allows us to focus on analyzing how context effects affect optimal pricing and merger incentives, given that one product is salient.[12] Further, like in Bordalo et al. (2013, 2016), we suppose that the weight $\delta \in (0, 1)$ is the degree of salient thinking and is independent of the chosen prices, and thus, a constant parameter.

Assumption 1 still guarantees a covered market in equilibrium. The indifferent consumer is now $\hat{x} = \frac{t+\delta(p_{A2}-p_{A1})}{2t}$. First, consider the three firm case before vertical integration occurs. Supposing again symmetry in equilibrium, $p^*_{A2} = p^*_{A1}$, we obtain $\hat{x} = \frac{1}{2}$. Any equilibrium covering the market must satisfy $\delta p^*_A + p^*_B = v - \frac{t}{2}$, which implies $p^*_B = v - \frac{t}{2} - \delta p^*_A$.

[11] According to Bordalo et al. (2022), three reasons may cause salience in consumer choice: contrast, surprise, and prominence. Prices of components A and B may differ greatly, creating contrast effects. One price could be surprisingly lower or higher than expected, or one component could be much more advertised to get attention.

[12] An immediate continuation of this chapter is to analyze the opposite case, when the competitive good is salient, and consumers weight p_B with $\delta \in (0, 1)$.

Further, similar to the benchmark case, it must be $p_A^* \in [c_A + \frac{t}{2\delta}, c_A + \frac{t}{\delta}]$ in equilibrium. To simplify notation, denote $\tilde{v} = v - \delta c_A - c_B$, where $\tilde{v} > \bar{v}$ since $\delta \in (0,1)$.

Lemma 3.4.1. *In any pure-strategy equilibrium, firms set prices*

$$p_A^* = c_A + \frac{t}{\delta(y-1)} \quad \textit{and} \quad p_B^* = v - \delta c_A - \frac{t(1+y)}{2(y-1)},$$

where $y \in [2,3]$. *Equilibrium profits are* $\pi_M^* = \tilde{v} - \frac{t(1+y)}{2(y-1)}$ *and* $\pi_i^* = \frac{t}{2\delta(y-1)}$.

This leads to $p_B^* \in [v - \delta c_A - \frac{3t}{2}, v - \delta c_A - t]$, $\pi_M^* \in [\tilde{v} - \frac{3t}{2}, \tilde{v} - t]$ and $\pi_i^* \in [\frac{t}{4\delta}, \frac{t}{2\delta}]$. Depending on how we define consumer welfare, according to Equation (3.1) or (3.2), some consumers may obtain a negative surplus. Suppose that δ is just a temporary effect at the moment of purchase and the effective utility a consumer experiences is defined as in the benchmark case with $\delta = 1$. Then, consumers around $\hat{x}$ will receive a negative surplus but do not realize this until after consumption. We can interpret these as consumers who got tricked by an appealing or surprising offer, and then regret the purchase or are disappointed after consumption.

3.4.1 Vertical Intregration

Suppose again that firm 1 and M merge. Any equilibrium must satisfy $\frac{\delta(p_{A1}^* + p_{A2}^*)}{2} + p_B^* = v - \frac{t}{2}$ or, equivalently, $p_B^* + \delta p_{A1}^* = v - t\hat{x}$ and the necessary minimum prices are

$$\underline{p}_{A2} = c_A + \frac{t + \delta(p_{A1} - c_A)}{3\delta}, \quad \underline{p}_{A2} = c_A + \frac{t + \delta(p_{A1} - c_A)}{3\delta} - \frac{2}{3}(p_B^* - c_B),$$
$$\underline{p}_B = c_B + \frac{t + c_A - p_{A1}^*}{2}.p$$

The maximum prizes for product A_i are

$$\bar{p}_{A1} = \bar{p}_{A2} = c_A + \frac{t + \delta(p_{A-i} - c_A)}{2\delta}$$

Lemma 3.4.2.

(i) In any pure-strategy equilibrium, both firms are active and set prices

$$p^*_{A1} = c_A + \frac{t(1 + y(1 - \delta))}{\delta(2y - 1)},\ p^*_{A2} = c_A + \frac{t(3 - \delta)}{\delta(2y - 1)},\ \text{and}$$
$$p^*_B = v - \delta c_A - \frac{t(3 - \delta)(1 + y)}{2(2y - 1)},$$

where $y \in [2, 3]$*. Equilibrium profits are*

$$\pi^*_{\hat{M}} = \tilde{v} - \frac{t(2\delta(2y + 3y^2 - 1) + \delta^2(1 - 2y - y^2) - (1 + y)^2)}{2\delta(2y - 1)^2},$$
$$\pi^*_2 = \frac{t(3 - \delta)^2(y - 1)}{2\delta(2y - 1)^2}$$

(ii) No equilibrium exists that excludes firm 2 *from the market.*

In contrast to the benchmark, no equilibrium exists, in which firm 2 is inactive. Since consumers underweight the price for component A_i, the merged firm must set an even lower price, $p_{A1} \leq \delta c_A - t$, to exclude firm 2 from the market. But, because consumers are less price-sensitive, the merged firm wants to take advantage of this by increasing p_{A1}. It turns out that setting $p_{A1} \leq \delta c_A - t$ cannot constitute an equilibrium as firm $\hat{M}$ has a profitable deviation to set p_B as low and p_{A1} as high as possible when firm 2 is inactive.

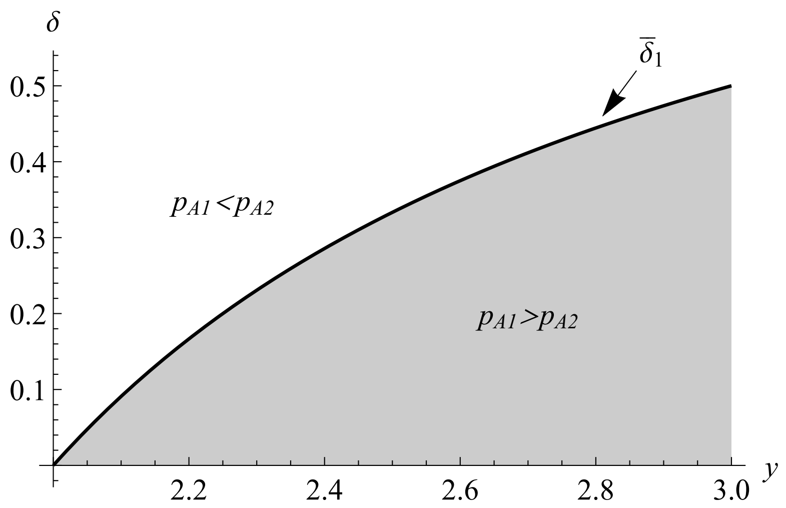
In equilibrium, we have the following intervals for prices and profits:

$$p^*_{A1} \in \left[c_A + \frac{t(4 - 3\delta)}{5\delta}, c_A + \frac{t(3 - 2\delta)}{3\delta}\right], p^*_{A2} \in \left[c_A + \frac{t(3 - 2\delta)}{5\delta}, c_A + \frac{t(3 - 2\delta)}{3\delta}\right]$$
$$p^*_B \in \left[v - \delta c_A - \frac{t(3 - \delta)}{2}, v - \delta c_A - \frac{2t(3 - \delta)}{5}\right], \pi^*_2 \in \left[\frac{t(3 - \delta)^2}{25\delta}, \frac{t(3 - \delta)^2}{18\delta}\right].$$

For $y = 2$, we have $\pi^*_{\hat{M}} = \tilde{v} - \frac{t(30\delta - 7\delta^2 - 9)}{18\delta}$, and for $y = 3$, it is $\pi^*_{\hat{M}} = \tilde{v} - \frac{t(32\delta - 7\delta^2 - 8)}{25\delta}$.[13] The demand for product A_1B is $\hat{x} = \frac{1 - \delta + y(1 + \delta)}{2(2y - 1)} \in [\frac{2 + \delta}{5}, \frac{3 + \delta}{6}]$, which is strictly lower than in the benchmark case with $\delta = 1$. The market share for the merged firm may even be lower than that of firm 2. This is

[13]Depending on δ, profits for the merged firm could be larger for $y = 2$ than for $y = 3$.

Figure 3.1: Competitive Prices



Prices for component A_i with behavioral consumers weighting p_{Ai} by $\delta \in (0, 1)$.

the case when context effects are sufficiently strong and y sufficiently high, $\delta < \frac{y-2}{y-1} = \bar{\delta}_1$, where $\bar{\delta}_1 \in [0, \frac{1}{2}]$, or equivalently when $y > \frac{2-\delta}{1-\delta}$. If $\delta < \bar{\delta}_1$, consumers underweight the price strongly and firm 2 sets a rather low price, then the merged firm sets a larger price for the competitive good, $p^*_{A1} > p^*_{A2}$, as depicted in Figure 3.1. In fact, in this case, product A_1 is more expensive, and component B is cheaper after vertical integration.[14] This implies that the merged firm wants to generate its profits rather from selling component A_1 than B. In the benchmark case, the merged firm collected the rents through the monopolistic product B, by lowering p^*_{A1} and increasing p_B after vertical integration. With behavioral consumers, this happens only if context effects are not too strong (and y is not too large). Note that $p^*_{A1} > p^*_{A2}$ is impossible when $y = 2$. That is when firms price the component A_i at the upper bound of the equilibrium interval.

This pricing pattern can be explained by a direct and an indirect channel. First, the merged firm wants to take advantage of behavioral consumers underweighting the price by increasing p_{A1}. Second, context effects soften

[14]The difference in p_B is $\frac{t(y-2-\delta(y-1))(1+y)}{2(y-1)(2y-1)}$, which is positive for $\delta \in (0, \bar{\delta}_1)$, implying a lower p_B after the merger.

the competition in the market of product A. This is immediately observable from the indifferent consumer $\hat{x} = \frac{t+\delta(p_{A2}-p_{A1})}{2t}$. A price change affects the demand less than in the benchmark when $\delta = 1$.[15] This allows firms to set a higher price p_{Ai} in equilibrium. But also, the merged firm can steal fewer consumers from firm 2 by setting a low p_{A1}. Traditionally, vertical integration reduces double marginalization and enables the merged firm to exert pressure on its rival by lowering p_{A1}, which, in turn, lowers p_{A2} due to strategic complementarity. Then, the rent is captured by increasing the price for the monopolistic good B. Salience, however, weakens this effect. Firm 2 does not need to react as strongly to a price change as in the benchmark. Hence, the efficiency gains from vertical integration are less powerful. All told, the merged firm wants to increase p_{A1} because consumers are less sensitive to it, which also makes competition less intense, reducing the gains from merging.

Interestingly, when $\delta < \bar{\delta}_1$, profits of firm 2 increase because of the merger, and total industry profits rise as well. Otherwise, both decline for $\delta > \bar{\delta}_1$, like in the benchmark. Product A_i is more expensive after the merger for $\delta > \bar{\delta}_1$ and cheaper when $\delta < \bar{\delta}_1$. The opposite is true for the monopolistic product B, which becomes more expensive after the merger for $\delta < \bar{\delta}_1$. The total price for the composite good A_1B, $p^*_{A1} + p^*_B$, is lower for $\delta > \bar{\delta}_1$ and higher otherwise. The composite good A_2B becomes cheaper when $\delta \in (\bar{\delta}_1, \frac{2}{y+1})$ and is more expensive otherwise.

Merger Incentives

The incentive constraint to merge is still given by $\pi_{\hat{M}} - F > \pi_1 + \pi_M$. Define $\bar{\delta}_2 = \frac{1+y^2-y}{y(2+y)-1}$, and observe that $\bar{\delta}_2 \in [\frac{3}{7}, \frac{1}{2}]$ and $\bar{\delta}_1 \leq \bar{\delta}_2$ for any $y \in [2, 3]$.

Lemma 3.4.3. *Suppose vertical integration incurs fixed costs $F > 0$ and does not affect y. Then, vertical integration occurs iff*

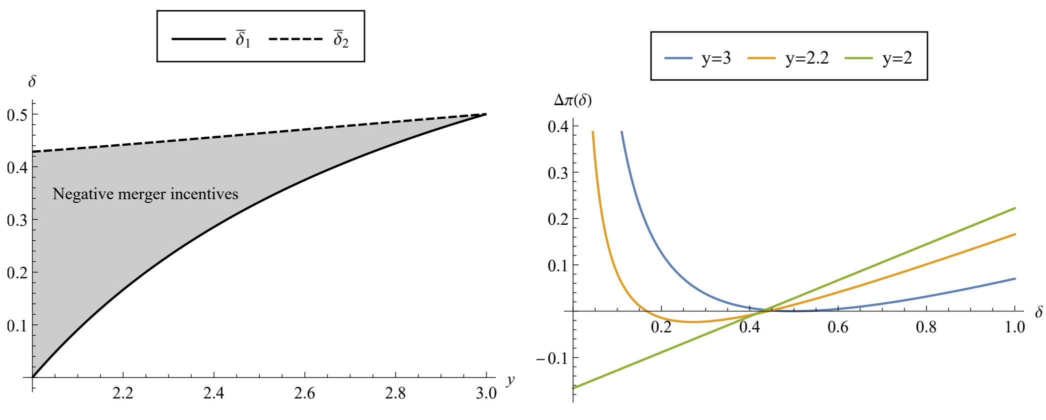
$$F < \frac{t(2 + \delta(y-1) - y)[y - 1 - y^2 + \delta(y(2+y) - 1)]}{2\delta(y-1)(2y-1)^2}. \tag{3.3}$$

When $\bar{\delta}_1 \leq \delta \leq \bar{\delta}_2$, then vertical integration is never profitable.

[15] In the extreme case of $\delta \to 0$, prices have no effect on demand.

When $\bar{\delta}_1 \leq \delta \leq \bar{\delta}_2$, the right-hand side of the inequality (3.3) in Lemma 3.4.3, the benefit of integration ($\Delta\pi$), is non-positive. Hence, even if formation costs are zero, vertical integration is not profitable. For instance, when $y = 2$, then $\Delta\pi = \frac{t(7\delta-3)}{18}$, $\bar{\delta}_1 = 0$ and $\bar{\delta}_2 = \frac{3}{7}$, implying $\Delta\pi \leq 0$ for $0 < \delta \leq \frac{3}{7}$. The findings of Lemma 3.4.3 are illustrated in Figure 3.2. The gray region in the left panel indicates when $\Delta\pi$ is negative, which is more likely when δ and y are low. Recall that a low y implies rather high prices in the competitive market A. Note that since $\delta \geq \bar{\delta}_1$, we have $p^*_{A1} \leq p^*_{A2}$ in this region. Merger incentives are negative when context effects are so strong that the efficiency gains from the merger are heavily reduced, but not strong enough that the merged firm can compensate for this by exploiting consumers with an expensive product A_1. The right panel comprises merger incentives for several fixed values of y. Except for the boundary case $y = 2$, we can observe that merger incentives are U-shaped in δ, implying that vertical integration becomes more likely if context effects are very strong. When $\delta > \bar{\delta}_2$ or $\delta < \bar{\delta}_1$, then the right-hand side of the inequality (3.3) is strictly positive. This is the white region in the left panel and when the functions are positive in the right panel of Figure 3.2. In this case, vertical integration occurs when formation costs are not too large.

Figure 3.2: Merger Incentives



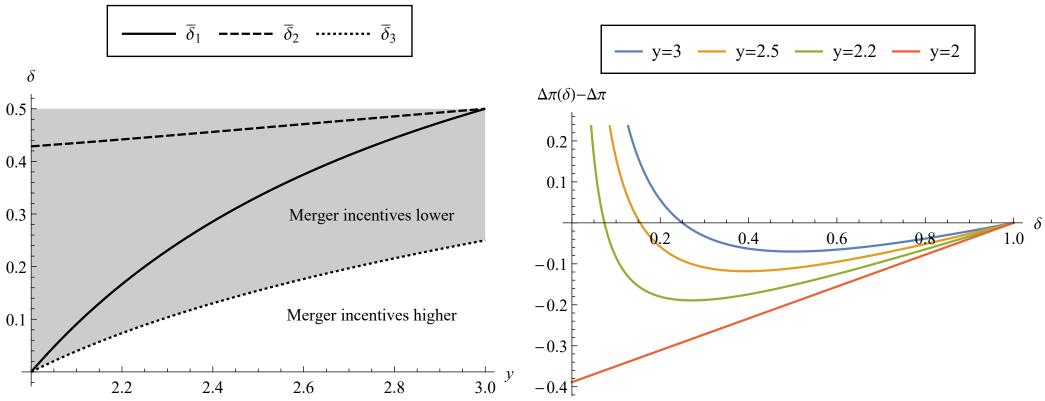
Merger incentives with behavioral consumers weighting p_{Ai} by $\delta \in (0, 1)$.

As described before, context effects soften competition and, thus, the efficiency gains of vertical integration. In other words, the merged firm cannot fully utilize the advantages of reducing double marginalization, which lowers the incentive to merge. For that reason, the incentives in the right panel of Figure 3.2 decrease when we start from the benchmark case ($\delta = 1$) and then reduce δ.

This competition effect makes it more difficult for vertical integration to be profitable compared to the benchmark. Thus, context effects reduce the likelihood that mergers occur, unless they are very strong. To analyze how context effects change the merger incentives compared to the benchmark, define $\bar{\delta}_3 = \frac{3y+y^3-2-3y^2}{1-3y+y^2+y^3}$, where $\bar{\delta}_3 \in [0, \frac{1}{4}]$ and $\bar{\delta}_3 \leq \bar{\delta}_1$.

Proposition 3.4.1. *Context effects on the competitive good decrease merger incentives if $\delta > \bar{\delta}_3$. Otherwise, merger incentives are weakly higher.*

Figure 3.3: Change in Merger Incentives



Difference of merger incentives between the benchmark case and with behavioral consumers.

For $\delta > \bar{\delta}_3$, then merger incentives are lower with salience. The reason emerges from the previous discussion. Context effects make competition less intense, which reduces the gains from vertical integration. Yet, for $\delta < \bar{\delta}_1$, the merged firm sets a larger price for product A than the competitor, taking advantage of behavioral consumers. Then, if $\delta < \bar{\delta}_3$, exploiting consumers by increasing p_{A1} dominates the loss of efficiency gains, and incentives are increased. However, this requires that consumers almost ignore the price of component A. It remains an empirical question of whether such markets with these extreme consumer preferences exist. The result of Proposition 3.4.1 is represented in Figure 3.3. The left panel depicts again the relationship between context effects (δ) and y. In the gray area, above $\bar{\delta}_3$, merger incentives are strictly lower than in the benchmark case. Below, in the white region, merger incentives are strictly larger. The right panel shows the difference in merger incentives for fixed values of y. We observe that lower incentives are more likely the lower y. In the extreme case of $y = 2$, merger incentives are strictly lower for any $\delta \in (0, 1)$.

3.5 Conclusion

This chapter studies markets with complementary products and asymmetric competition when consumers underweight the prize of one component due to context effects. In this case, the purchasing decision is distorted, and consumers accept a higher price for a complement than if context effects are absent. When the competitive good's price is underweighted, then competition becomes less intense. This reduces the efficiency gains, and thus the profitability, of vertical integration. Hence, the presence of such consumers affects the market formation as merger incentives are lower due to this competition effect and become less likely. Only if the effects are very strong, then we find an increase in merger incentives.

The model comes with some strict assumptions and limitations. An immediate next step is to analyze how context effects on the monopolistic good affect market formation. First derivations show that merger incentives increase and vertical integration becomes more likely. With context effects on the monopolistic product, competition intensity is not affected. Thus, the efficiency gains of vertical integration are not diminished. On the contrary, because consumers underweight the price of the monopolistic goods, the merged firm can recoup its advantages disproportionately high, making a merger more profitable.

Follow-up work should focus on microfoundation for context effects to understand where they are coming from and how firms may manipulate them, which would endogenize which complement is salient. This would add another strategic element to the game and may deliver further results. For completeness, the model should be extended to uncovered markets to gain an understanding of whether this is important for the results on market formation. Further, analyzing asymmetric firms in the competitive market may be worthwhile. For policy recommendations, it is important to understand which firm the monopolist prefers to acquire and what are the different market implications. Lastly, a welfare analysis could provide interesting insights. Traditionally, vertical integrations are efficient and less of a concern for competition authorities. This may change for markets involving consumers with context-dependent preferences.

3.A Proofs

3.A.1 Proof of Lemma 3.3.1

First, since $p^*_{A1} = p^*_{A2} = p^*_A$, we have $\hat{x} = \frac{1}{2}$. The indifferent consumer purchases if $v - p^*_A - p^*_B - t\hat{x} \geq 0$, which reduces to $p^*_A + p^*_B \leq v - \frac{t}{2}$. Thus, any equilibrium must satisfy $p^*_A + p^*_B = v - \frac{t}{2}$. Suppose not. When $p^*_A + p^*_B < v - \frac{t}{2}$, the monopolist can increase p^*_B without losing demand, which is strictly profitable. When $p^*_A + p^*_B > v - \frac{t}{2}$, the market is uncovered. Then firm i is a local monopolist and maximizes $\pi_i = (p_{Ai} - c_A)D_{A_iB}$, where $D_{A_iB} = \frac{v - p_{Ai} - p_B}{t}$. Firm M maximizes $\pi_M = (p_B - c_B)(D_{A_1B} + D_{A_2B})$. The equilibrium candidate is

$$p_{A1} = p_{A2} = \frac{v + 2c_A - c_B}{3}, \qquad p_B = \frac{v - c_A + 2c_B}{3}$$
$$D_{A_1B} = D_{A_2B} = \frac{\bar{v}}{3t}, \qquad \pi_1 = \pi_2 = \frac{\bar{v}^2}{9t}, \qquad \pi_M = \frac{2\bar{v}^2}{9t}.$$

In an uncovered market, it must be $D_{A_1B} + D_{A_2B} < 1$, which is true for $t > \frac{2\bar{v}}{3} > \frac{\bar{v}}{2}$. Given Assumption 1, this condition is not satisfied, and an equilibrium with uncovered markets does not exist. Thus, it must be $p^*_A +$

$p_B^* = v - \frac{t}{2}$ and a covered market in equilibrium.

Second, we prove that $p_A^* \in \left[c_A + \frac{t}{2}, c_A + t\right]$. Define $z_A > 0$ as the markup that firm i earns in equilibrium. Then, $p_A^* = c_A + z_A$ and $\pi_i^* = \frac{z_A}{2}$. Consider the deviation $p_A^d = c_A + z_A + \epsilon$, where $\epsilon > 0$, which yields $\pi_i^d = (z_A+\epsilon)(\frac{1}{2} - \frac{\epsilon}{t}) = \frac{z_A}{2} + \epsilon[\frac{1}{2} - \frac{z_A+\epsilon}{t}]$. This is profitable if $z_A < \frac{t}{2} - \epsilon$. Hence, it must be $z_A \geq \frac{t}{2}$ in equilibrium. Consider now the deviation $p_A^d = c_A + z_A - \epsilon$, which yields $\pi_i^d = (z_A - \epsilon)(\frac{1}{2} + \frac{\epsilon}{2t}) = \frac{z_A}{2} + \epsilon[\frac{z_A-\epsilon}{2t} - \frac{1}{2}]$. This is profitable if $z_A > t + \epsilon$. Hence, in equilibrium, it must be $z_A \leq t$. Combining the statements leads to $z_A \in [\frac{t}{2}, t]$ in equilibrium, implying $p_A^* \in \left[c_A + \frac{t}{2}, c_A + t\right]$.

The equilibrium condition $p_A^* + p_B^* = v - \frac{t}{2}$ implies
$p_B^* \in \left[v - c_A - \frac{3t}{2}, v - c_A - t\right]$. Define $z_B > 0$ as the markup that firm M earns in equilibrium. Then, $p_B^* = c_B + z_B$ and $\pi_M^* = z_B$. Consider the deviation $p_B^d = c_B + z_B + \epsilon$, which yields $\pi_M^d = (z_B + \epsilon)(1 - \frac{2\epsilon}{t}) = z_B + \epsilon - \frac{2z_B\epsilon + 2\epsilon^2}{t}$. This is profitable if $z_B < \frac{t}{2} - \epsilon$. Hence, in equilibrium it must be $z_B \geq \frac{t}{2}$ and $p_B^* \geq c_B + \frac{t}{2}$. Observe that $v - c_A - t > v - c_A - \frac{3t}{2} \geq c_B + \frac{t}{2}$ for $t \leq \frac{\bar{v}}{2}$. Hence, no profitable upward deviation exists for firm M. Further, no profitable downward deviation for firm M exists, since demand is unaffected. Thus, $p_B^* - \epsilon$ would lead to strictly lower profits.

Conjecture: $p_A^* = c_A + \frac{t}{y-1}$ with $y \in [2, 3]$ is the best response of firm i. Then it must be $p_B^* = v - c_A - \frac{t}{y-1} - \frac{t}{2} = v - c_A - \frac{t(1+y)}{2(y-1)}$. Since $\frac{t(1+y)}{2(y-1)} \in [t, \frac{3t}{2}]$ for $y \in [2, 3]$, no profitable deviation for firm M exists. Further, since $\frac{t}{y-1} \in [\frac{t}{2}, t]$ for $y \in [2, 3]$, no profitable deviation for firm i exists. The equilibrium profits are given by $\pi_i^* = \frac{1}{2}(p_A^* - c_A) = \frac{t}{2(y-1)}$ and $\pi_i^* = p_B^* - c_B = \bar{v} - \frac{t(1+y)}{2(y-1)}$.

□

3.A.2 Proof of Lemma 3.3.2

We prove part (i) in several steps. First, we show that given Assumption 1, no equilibrium exists that leads to an uncovered market. Second, we derive the interval of equilibrium prices for firm 2 and form a conjecture for the best response. Lastly, we derive the best response for the merged firm and verify the conjecture.

Suppose both firms are active. Similar to the proof of Lemma 3.3.1, it must be $p_B^* = v - p_{A1} - t\hat{x} = v - \frac{t+p_{A1}^*+p_{A2}^*}{2}$ in an equilibrium with a covered market. When $p_B^* < v - \frac{t+p_{A1}^*+p_{A2}^*}{2}$, firm $\hat{M}$ can increase p_B^* without losing demand. When $p_B^* > v - \frac{t+p_{A1}^*+p_{A2}^*}{2}$, the market is uncovered. In this case, firm 2 is a local monopolist, who maximizes $\pi_2 = (p_{A2}-c_A)D_{A_2B}$. Taking the first-order condition, we obtain $p_{A2} = \frac{v+c_A-p_B}{2}$. The merged firm maximizes

$$\pi_{\hat{M}} = (p_B - c_B)(D_{A_1B} + D_{A_2B}) + (p_{A1} - c_A)D_{A_1B}.$$

The first-order conditions with respect to p_{A1} and p_B yield $p_{A1} = \frac{v+c_A+c_B}{2} - p_B$ and $4p_B = 2v + c_A + 2c_B - 2p_{A1} - p_{A2}$. Solving all three first-order conditions simultaneously, we obtain $p_{A1} = \frac{v+5c_A-c_B}{6}$, $p_B = \frac{v-c_A+2c_B}{3}$ and $p_{A2} = \frac{v+2c_A-c_B}{3}$. This implies $D_{A_1B} = \frac{\bar{v}}{2t}$ and $D_{A_2B} = \frac{\bar{v}}{3t}$. Hence, aggregated demand is $D_{A_1B}+D_{A_2B} = \frac{5\bar{v}}{6t}$, which must be strictly lower than 1. Otherwise, the market is covered. Observe that $\frac{5\bar{v}}{6t} < 1$ only if $t > \frac{5\bar{v}}{6}$. A contradiction. Hence, given Assumption 1, the market is covered in equilibrium.

Similar to the proof of Lemma 3.3.1, it must be $p_{A2}^* \in \left[\frac{2c_A+t+p_{A1}}{3}, \frac{c_A+t+p_{A1}}{2}\right]$: Define $z_{A2} > 0$ as the markup that firm 2 earns in equilibrium. Then, $p_{A2}^* = c_A+z_{A2}$ and $\pi_i^* = z_{A2}(1-\hat{x})$. Consider the deviation $p_{A2}^d = c_A+z_{A2}+\epsilon$, which yields $\pi_{A2}^d = (z_{A2} + \epsilon)(1 - \hat{x} - \frac{\epsilon}{t}) = z_{A2}(1 - \hat{x}) + \epsilon(1 - \hat{x} - \frac{z_{A2}+\epsilon}{t})$. This is profitable if $z_{A2} < t(1 - \hat{x}) - \epsilon$. Hence, using $\hat{x} = \frac{1}{2} + \frac{p_{A2}^*-p_{A1}}{2t}$ and $p_{A2}^* = c_A + z_{A2}$, it must be $z_{A2} \geq \frac{t+p_{A1}-c_A}{3}$ in equilibrium. Consider now the deviation $p_{A2}^d = c_A + z_{A2} - \epsilon$, which yields $\pi_{A2}^d = (z_{A2} - \epsilon)(1 - \hat{x} + \frac{\epsilon}{2t}) = z_{A2}(1-\hat{x})-\epsilon(1-\hat{x})+\frac{z_{A2}\epsilon-\epsilon^2}{2t}$. This is profitable if $z_{A2} > 2t(1-\hat{x})+\epsilon$. Hence, in equilibrium, it must be $z_{A2} \leq \frac{t+p_{A1}-c_A}{2}$. Combining the statements leads to $z_{A2} \in [\frac{t+p_{A1}-c_A}{3}, \frac{t+p_{A1}-c_A}{2}]$, implying $p_{A2}^* \in \left[\frac{2c_A+t+p_{A1}}{3}, \frac{c_A+t+p_{A1}}{2}\right]$. We have $\frac{2c_A+t+p_{A1}}{3} < \frac{c_A+t+p_{A1}}{2}$, when $p_{A1} > c_A - t$. This must be the case when firm 2 is active ($\hat{x} < 1$). Suppose $p_{A1} = c_A - t$, then $\hat{x} = 1 + \frac{p_{A2}-c_A}{2t} \geq 1$ since it must be $p_{A2} \geq c_A$. Otherwise, firm 2 would obtain negative profits.

Conjecture for p_{A2}^*: From $p_{A2}^* \in \left[c_A + \frac{t+p_{A1}-c_A}{3}, c_A + \frac{t+p_{A1}-c_A}{2}\right]$, we can deduce that the markup firm 2 obtains in equilibrium, is given by $(t+p_{A1}-c_A)$ divided by some constant, which we capture with y. Hence, consider $p_{A2}^* = c_A + \frac{t+p_{A1}-c_A}{y}$ with $y \in [2, 3]$ as the best response of firm 2. Further, from the covered market condition, we obtain $p_B = v - p_{A1} - t\hat{x}$. Plugging this into the profit function of the merged firm and using $\hat{x} = \frac{t+p_{A2}-p_{A1}}{2t}$ yields:

$$\pi_{\hat{M}} = v - \frac{p_{A1}+p_{A2}+t}{2} - c_B + (p_{A1}-c_A)\left(\frac{1}{2}+\frac{p_{A2}-p_{A1}}{2t}\right)$$

Taking the first-order condition, we obtain:

$$p^*_{A1} = \frac{c_A + p_{A2}}{2} = c_A + \frac{t}{2y-1},$$

which implies $p^*_{A2} = c_A + \frac{2t}{2y-1}$. The demand for A_1B in equilibrium is then $\hat{x} = \frac{y}{2y-1}$ and for A_2B we have $1-\hat{x} = \frac{y-1}{2y-1}$. Thus, we obtain $p^*_B = v - c_A - \frac{t(1+y)}{2y-1}$. The equilibrium profits are $\pi^*_{\hat{M}} = \bar{v} - \frac{t(2y^2-1)}{(2y-1)^2}$ and $\pi^*_2 = \frac{2t(y-1)}{(2y-1)^2}$.

We verify now that p^*_{A2} is indeed the best response to p^*_{A1} and p^*_B. First, suppose the best response of firm 2, given p^*_{A1} and p^*_B, leads to an uncovered market. Then, $\pi_2 = (p_{A2}-c_A)(\frac{c_A-p_{A2}}{t}+\frac{1+y}{2y-1})$. The first-order condition yields $p_{A2} = c_A + \frac{t(1+y)}{2(2y-1)}$. Note that this solution is only feasible when $p_{A2} > p^*_{A2}$. Otherwise, the market is covered.

$$c_A + \frac{t(1+y)}{2(2y-1)} > c_A + \frac{2t}{2y-1}$$
$$y > 3$$

A contradiction since $y \in [2,3]$. Thus, the best response of firm 2, given p^*_{A1} and p^*_B, must lead to a cover market.

Then, we have $\pi^{cov}_2 = (p_{A2}-c_A)(\frac{1}{2}+\frac{1}{2(2y-1)}+\frac{c_A-p_{A2}}{2t})$. The first-order condition yields $p_{A2} = c_A + \frac{yt}{2y-1}$. Note that this solution is only feasible when $p_{A2} \leq p^*_{A2}$. Otherwise, the market becomes uncovered.

$$c_A + \frac{yt}{2y-1} \leq c_A + \frac{2t}{2y-1}$$
$$y \leq 2$$

Observe that for $y = 2$, we have $p_2 = c_A + \frac{2t}{3} = p^*_{A2}$. Any $y < 2$ is not feasible. For any $y \in (2,3]$, firm 2 sets the largest price p_{A2} that just covers the market. We can observe this from

$$\frac{\partial \pi_2^{cov}}{\partial p_{A2}} = \frac{1}{2} + \frac{1}{2(2y-1)} + \frac{c_A - p_{A2}}{t} = \frac{y}{2y-1} + \frac{c_A - p_{A2}}{t},$$

which is positive when $p_{A2} - c_A < \frac{yt}{2y-1}$. In this case, it is optimal to set the largest price that just covers the market. Plugging in p_{A2}^*, which is this largest price, we obtain $\frac{\partial \pi_2^{cov}}{\partial p_{A2}} > 0$ if $y > 2$. Hence, for $y \in [2,3]$, $p_{A2}^* = c_A + \frac{2t}{2y-1}$ is indeed the best response for firm 2.

Further, given Assumption 1, no profitable deviation for the merged firm exists. Suppose $p_B^* + \epsilon$. Then

$$\begin{aligned}\hat{\pi}_M^d &= \left(\bar{v} - \frac{t(1+y)}{2y-1} + \epsilon\right)\left(1 - \frac{2\epsilon}{t}\right) + \frac{t}{2y-1}\left(\frac{y}{2y-1} - \frac{\epsilon}{t}\right) \\ &= \bar{v} - \frac{t(2y^2-1)}{(2y-1)^2} + \epsilon\left[\frac{4y}{2y-1} - \frac{2(\bar{v}+\epsilon)}{t}\right],\end{aligned}$$

which is profitable if

$$\begin{aligned}\frac{4y}{2y-1} &> \frac{2(\bar{v}+\epsilon)}{t} \\ t &> \frac{(2y-1)(\bar{v}+\epsilon)}{2y}.\end{aligned}$$

A contradiction since $t > \frac{3\bar{v}}{4}$ for $y = 2$ and $\frac{\partial\left[\frac{(2y-1)(\bar{v}+\epsilon)}{2y}\right]}{\partial y} > 0$. Suppose $p_{A1}^* + \epsilon$. Then, similarly to above, $\hat{\pi}_M^d = \bar{v} - \frac{t(2y^2-1)}{(2y-1)^2} + \epsilon[\frac{2y}{2y-1} - \frac{\bar{v}+\epsilon}{t}]$, which is profitable if $t > \frac{(2y-1)(\bar{v}+\epsilon)}{2y}$. Similar to before, a contradiction. Lastly, suppose $p_{A1}^* - \epsilon$. Then

$$\begin{aligned}\hat{\pi}_M^d &= \bar{v} - \frac{t(1+y)}{2y-1} + \left(\frac{t}{2y-1} - \epsilon\right)\left(\frac{y}{2y-1} + \frac{\epsilon}{2t}\right) \\ &= \bar{v} - \frac{t(2y^2-1)}{(2y-1)^2} - \epsilon\left[\frac{1}{2} + \frac{\epsilon}{2t}\right] < \pi_{\hat{M}}^*.\end{aligned}$$

(ii) Depending on whether the market is covered, firm $\hat{M}$ can set either

p_{A1} sufficiently low or p_B sufficiently large to exclude firm 2. First, suppose firm 2 is not active, and the market is not covered. Then, $\hat{M}$ maximizes $\pi_{\hat{M}} = (p_{A1} + p_B - c_A - c_B)(\frac{v-p_{A1}-p_B}{t})$. The sum of profit-maximizing prices is given by $p_{A1}+p_B = \frac{v+c_A+c_B}{2}$, where $p_B \geq v-p_{A2}$ is needed to exclude firm 2. The demand is $D_{A_1B} = \frac{v-c_A-c_B}{2t} = \frac{\bar{v}}{2t}$. Note that an uncovered market, in this case, requires $D_{A_1B} < 1$. A contradiction by Assumption 1, since $t < \frac{\bar{v}}{2}$ implies $D_{A_1B} > 1$. Thus, given Assumption 1, there exists no equilibrium, in which firm 2 is excluded and the market is uncovered.

Now, suppose firm 2 is not active, and the market is covered. When $t \leq c_A$, firm $\hat{M}$ can set $p_{A1} \leq c_A - t$ and firm 2 has zero demand (or would obtain negative profits). Otherwise, if $p_{A1} > c_A-t$, firm 2 can set $p_{A2} = c_A+\epsilon$ and obtain strictly positive demand and profits. To cover the market, firm $\hat{M}$ then sets $p_B = v - p_{A1} - t$. The sum of profit-maximizing prices is $p_{A1}+p_B = v-t$ and the resulting profits are $\pi_{\hat{M}} = p_{A1}+p_B-c_A-c_B = \bar{v}-t$. Firm 2 obtains $\pi_2 = 0$ for any p_{A2} given that $p_{A1} \leq c_A-t$. Thus, no profitable deviation for firm 2 exists. For firm $\hat{M}$, there exists a profitable deviation if p_{A2} is sufficiently high. For convenience, we denote $p_{A2} = c_A + z_{A2}$. Consider the solution from part (i): $p_{A1} = \frac{c_A+p_{A2}}{2} = c_A+\frac{z_{A2}}{2}$, implying $\hat{x} = \frac{1}{2}+\frac{z_{A2}}{4t}$, and $p_B = v-\frac{p_{A1}+p_{A2}+t}{2} = v-c_A-\frac{3z_{A2}}{4}-\frac{t}{2}$. Then, $\pi^d_{\hat{M}} = \bar{v}-\frac{t}{2}-\frac{3z_{A2}}{4}+\frac{z_{A2}}{2}(\frac{1}{2}+\frac{z_{A2}}{4t})$. First, observe that $\hat{x} \geq 1$ for $z_{A2} \geq 2t$. Hence, for any $p_{A2} \geq c_A+2t$, demand for firm $\hat{M}$ is fixed at 1. In that case, for $z_{A2} \geq 2t$, we have $\pi^d_{\hat{M}} = \bar{v}-\frac{t}{2}-\frac{z_{A2}}{4} \leq \bar{v} - t$. Hence, the deviation is not profitable when $p_{A2} \geq c_A + 2t$. Note that $\pi^d_{\hat{M}} = \bar{v} - t$ for $z_{A2} = 2t$. Suppose that $z_{A2} < 2t$. Then $\frac{\partial \pi^d_{\hat{M}}}{\partial z_{A2}} = \frac{z_{A2}}{4t} - \frac{1}{2} < 0$. Hence, $\pi^d_{\hat{M}} > \bar{v} - t$ for $z_{A2} < 2t$, implying a profitable deviation. Thus, excluding firm 2 is an equilibrium only if $p_{A2} \geq c_A + 2t$. Note that the deviation $v - t - \epsilon$ is not profitable since demand does not increase and $v - t + \epsilon$ to uncover the market is not profitable for $t < \frac{\bar{v}}{2}$ as shown above.

When $t > c_A$, then $p_{A1} \leq c_A - t$ is not feasible since prices must be non-negative. Observe that when $p_{A1} = 0$, firm 2 can set $p_{A2} = c_A + \epsilon$ and obtain positive demand $1 - \hat{x} = \frac{1}{2} - \frac{c_A}{2t} - \frac{\epsilon}{2t}$, leading to $\pi_2 = \epsilon(\frac{1}{2} - \frac{c_A}{2t} - \frac{\epsilon}{2t})$, which is strictly positive when $t > c_A + \epsilon$. This holds for any $p_{A1} \geq 0$. Thus, firm 2 is active when $t > c_A$. A contradiction. Hence, an equilibrium that excludes firm 2 from the market exists only if $p_{A2} \geq c_A + 2t$ and $t \leq c_A$. Otherwise, either firm $\hat{M}$ or firm 2 has a profitable deviation.

Observe that setting $p_{A2} \geq c_A + 2t$ is weakly dominated by $p_{A2} < c_A + 2t$. With the former strategy, firm 2 always obtains $\pi_2 = 0$; with the latter, depending on the action of firm $\hat{M}$, profits are weakly positive, $\pi_2 \geq 0$. Thus, by excluding equilibria in weakly dominated strategies, the equilibrium given by $p_{A1} + p_B = v - t$ and $p_{A2} \geq c_A + 2t$ is inadmissible. □

3.A.3 Proof of Lemma 3.3.3

$$\begin{aligned}
\pi_{\hat{M}} - F &> \pi_1 + \pi_M \\
\bar{v} - \frac{t(2y^2-1)}{(2y-1)^2} - F &> \bar{v} - \frac{t(1+y)}{2(y-1)} + \frac{t}{2(y-1)} \\
\frac{t[y(2y-1)^2 - 2(y-1)(2y^2-1)]}{2(y-1)(2y-1)^2} &> F \\
\frac{t(3y-2)}{2(y-1)(2y-1)^2} &> F
\end{aligned}$$

□

3.A.4 Proof of Lemma 3.4.1

First, since $p^*_{A1} = p^*_{A2} = p^*_A$, we have $\hat{x} = \frac{1}{2}$. The indifferent consumer purchases if $v - \delta p^*_A - p^*_B - t\hat{x} \geq 0$, which reduces to $\delta p^*_A + p^*_B \leq v - \frac{t}{2}$. Thus, any equilibrium must satisfy $\delta p^*_A + p^*_B = v - \frac{t}{2}$. Suppose not. When $\delta p^*_A + p^*_B < v - \frac{t}{2}$, the monopolist can increase p^*_B without losing demand, which is strictly profitable. When $\delta p^*_A + p^*_B > v - \frac{t}{2}$, the market is uncovered. Then firm i is a local monopolist and maximizes $\pi_i = (p_{Ai} - c_A)D_{A_iB}$, where $D_{A_iB} = \frac{v - \delta p_{Ai} - p_B}{t}$. Firm M maximizes $\pi_M = (p_B - c_B)(D_{A_1B} + D_{A_2B})$. The equilibrium candidate is

$$p_{A1} = p_{A2} = \frac{v + 2\delta c_A - c_B}{3\delta}, \qquad p_B = \frac{v - \delta c_A + 2c_B}{3}$$
$$D_{A_1B} = D_{A_2B} = \frac{\tilde{v}}{3t}, \qquad \pi_1 = \pi_2 = \frac{\tilde{v}^2}{9\delta t}, \qquad \pi_M = \frac{2\tilde{v}^2}{9t}.$$

In an uncovered market, it must be $D_{A_1B} + D_{A_2B} < 1$, which holds for $t > \frac{2\tilde{v}}{3} > \frac{\tilde{v}}{2} > \frac{\bar{v}}{2}$ since $\tilde{v} > \bar{v}$ for $\delta \in (0,1)$. A contradiction, given Assumption 1, an equilibrium with uncovered markets does not exist. Hence, it must be $\delta p_A^* + p_B^* = v - \frac{t}{2}$ in equilibrium with a covered market.

Second, we prove that $p_A^* \in \left[c_A + \frac{t}{2\delta}, c_A + \frac{t}{\delta}\right]$. Define $z_A > 0$ as the markup that firm i earns in equilibrium. Then, $p_A^* = c_A + z_A$ and $\pi_i^* = \frac{z_A}{2}$. Consider the deviation $p_A^d = c_A + z_A + \epsilon$, where $\epsilon > 0$, which yields $\pi_i^d = (z_A + \epsilon)(\frac{1}{2} - \frac{\delta\epsilon}{t}) = \frac{z_A}{2} + \epsilon[\frac{1}{2} - \frac{\delta z_A + \delta\epsilon}{t}]$. This is profitable if $z_A < \frac{t}{2\delta} - \epsilon$. Hence, it must be $z_A \geq \frac{t}{2\delta}$ in equilibrium. Consider now the deviation $p_A^d = c_A + z_A - \epsilon$, which yields $\pi_i^d = (z_A - \epsilon)(\frac{1}{2} + \frac{\delta\epsilon}{2t}) = \frac{z_A}{2} + \epsilon[\frac{\delta z_A - \delta\epsilon}{2t} - \frac{1}{2}]$. This is profitable if $z_A > \frac{t}{\delta} + \epsilon$. Hence, in equilibrium, it must be $z_A \leq \frac{t}{\delta}$. Combining the statements leads to $z_A \in [\frac{t}{2\delta}, \frac{t}{\delta}]$, implying $p_A^* \in \left[c_A + \frac{t}{2\delta}, c_A + \frac{t}{\delta}\right]$.

The equilibrium condition $\delta p_A^* + p_B^* = v - \frac{t}{2}$ implies $p_B^* \in \left[v - \delta c_A - \frac{3t}{2}, v - c_A - t\right]$. Similar to the proof of Lemma 3.3.1, there exists no profitable deviation since δ does not affect the deviation $p_B^d = c_B + z_B + \epsilon$, which leads to $\pi_M^d = (z_B + \epsilon)(1 - \frac{2\epsilon}{t})$ like in the benchmark case. Since $v - \delta c_A - \frac{3t}{2} > v - c_A - \frac{3t}{2}$ it must be $v - \delta c_A - \frac{3t}{2} > c_B$ for $t \leq \frac{\bar{v}}{2}$.

Conjecture: $p_A^* = c_A + \frac{t}{\delta(y-1)}$ with $y \in [2,3]$ is the best response of firm i. Then it must be $p_B^* = v - \delta c_A - \frac{t(1+y)}{2(y-1)}$. Since $\frac{t(1+y)}{2(y-1)} \in [t, \frac{3t}{2}]$ for $y \in [2,3]$, no profitable deviation for firm M exists. Further, since $\frac{t}{\delta(y-1)} \in [\frac{t}{2\delta}, \frac{t}{\delta}]$ for $y \in [2,3]$, no profitable deviation for firm i exists. The equilibrium profits are given by $\pi_i^* = \frac{1}{2}(p_A^* - c_A) = \frac{t}{2\delta(y-1)}$ and $\pi_i^* = p_B^* - c_B = \tilde{v} - \frac{t(1+y)}{2(y-1)}$.

□

3.A.5 Proof of Lemma 3.4.2

The proof follows a similar structure to the proof of Lemma 3.3.2. Suppose both firms are active. Then, it must be $p_B^* = v - \delta p_{A1} - t\hat{x} = v - \frac{t + \delta(p_{A1}^* + p_{A2}^*)}{2}$ in an equilibrium with a covered market. When $p_B^* < v - \frac{t+\delta(p_{A1}^* + p_{A2}^*)}{2}$, firm $\hat{M}$ can increase p_B^* without losing demand. When $p_B^* > v - \frac{t+\delta(p_{A1}^* + p_{A2}^*)}{2}$, the market is uncovered. In this case, firm 2 is a local monopolist, who maximizes $\pi_2 = (p_{A2} - c_A)D_{A_2B}$. Taking the first-order condition, we obtain

$p_{A2} = \frac{v+\delta c_A - p_B}{2\delta}$. The merged firm maximizes

$$\pi_{\hat{M}} = (p_B - c_B)(D_{A_1B} + D_{A_2B}) + (p_{A1} - c_A)D_{A_1B}.$$

The first-order conditions with respect to p_{A1} and p_B yield

$$p_{A1} = \frac{v + \delta(c_A + c_B) - (1+\delta)p_B}{2\delta} \text{ and } 4p_B = 2v + c_A + 2c_B - (1+\delta)p_{A1} - \delta p_{A2}.$$

Solving all three first-order conditions simultaneously, we obtain

$$\begin{aligned} p_{A1} &= \frac{(4-3\delta)(v-c_B) + (6\delta + \delta^2 - 2)c_A}{2(5\delta - 1 - \delta^2)}, \\ p_B &= \frac{(4\delta - 2)(v - \delta c_A) + 2\delta(3-\delta)c_B}{2(5\delta - 1 - \delta^2)}, \\ p_{A2} &= \frac{(3-\delta)(v-c_B) + (7\delta - 2 - \delta^2)c_A}{2(5\delta - 1 - \delta^2)}. \end{aligned}$$

This implies $D_{A_1B} = \frac{\delta(2+\delta)(v-\delta c_A - c_B)}{2t(5\delta-1-\delta^2)}$ and $D_{A_2B} = \frac{\delta(3-\delta)(v-\delta c_A - c_B)}{2t(5\delta-1-\delta^2)}$. Hence, aggregated demand is $D_{A_1B} + D_{A_2B} = \frac{5\delta(v-\delta c_A - c_B)}{2t(5\delta-1-\delta^2)}$, which must be strictly lower than 1. Otherwise, the market is covered. Recall that the condition for an uncovered market in Lemma 3.3.2 was $t > \frac{5\bar{v}}{6}$. We have

$$\frac{5\bar{v}}{6} < \frac{5\delta\tilde{v}}{2(5\delta - 1 - \delta^2)} < t.$$

The first inequality holds since $\bar{v} < \tilde{v}$ and

$$\begin{aligned} \frac{5}{6} &< \frac{5\delta}{2(5\delta - 1 - \delta^2)} \\ \frac{5\delta - 1 - \delta^2}{3} &< \delta \\ 0 &< (1-\delta)^2. \end{aligned}$$

Hence, $D_{A_1B} + D_{A_2B} > 1$ given Assumption 1, a contradiction. The market must be covered in equilibrium.

In equilibrium, it must be $p^*_{A2} \in \left[\frac{2\delta c_A + t + \delta p_{A1}}{3\delta}, \frac{\delta c_A + t + \delta p_{A1}}{2\delta}\right]$: Define $z_{A2} > 0$

as the markup that firm 2 earns in equilibrium. Then, $p^*_{A2} = c_A + z_{A2}$ and $\pi^*_i = z_{A2}(1-\hat{x})$. Consider the deviation $p^d_{A2} = c_A + z_{A2} + \epsilon$, which yields $\pi^d_{A2} = (z_{A2}+\epsilon)(1-\hat{x}-\frac{\delta\epsilon}{t}) = z_{A2}(1-\hat{x})+\epsilon(1-\hat{x}-\frac{\delta z_{A2}+\delta\epsilon}{t})$. This is profitable if $z_{A2} < \frac{t(1-\hat{x})}{\delta} - \epsilon$. Hence, using $\hat{x} = \frac{1}{2} + \frac{\delta(p^*_{A2}-p_{A1})}{2t}$ and $p^*_{A2} = c_A + z_{A2}$, it must be $z_{A2} \geq \frac{t+\delta(p_{A1}-c_A)}{3\delta}$ in equilibrium. Consider now the deviation $p^d_{A2} = c_A + z_{A2} - \epsilon$, which yields $\pi^d_{A2} = (z_{A2}-\epsilon)(1-\hat{x}+\frac{\delta\epsilon}{2t}) = z_{A2}(1-\hat{x}) - \epsilon(1-\hat{x}) + \frac{\delta z_{A2}\epsilon - \delta\epsilon^2}{2t}$. This is profitable if $z_{A2} > \frac{2t}{\delta}(1-\hat{x}) + \epsilon$. Hence, in equilibrium, it must be $z_{A2} \leq \frac{t+\delta(p_{A1}-c_A)}{2\delta}$. Combining the statements leads to $p^*_{A2} \in \left[c_A + \frac{t+\delta(p_{A1}-c_A)}{3\delta}, c_A + \frac{t+\delta(p_{A1}-c_A)}{2\delta}\right]$.

Similar to the proof of Lemma 3.3.2, we conjecture $p^*_{A2} = c_A + \frac{t+\delta(p_{A1}-c_A)}{\delta y}$ with $y \in [2,3]$ as the best response of firm 2. Further, from the covered market condition, we obtain $p_B = v - \delta p_{A1} - t\hat{x}$. Plugging this into the profit function of the merged firm and using $\hat{x} = \frac{t+\delta(p_{A2}-p_{A1})}{2t}$ yields:

$$\pi_{\hat{M}} = v - \frac{\delta p_{A1} + \delta p_{A2} + t}{2} - c_B + (p_{A1} - c_A)\left(\frac{1}{2} + \frac{\delta p_{A2} - \delta p_{A1}}{2t}\right)$$

Taking the first-order condition, we obtain:

$$p^*_{A1} = \frac{c_A + p_{A2}}{2} + \frac{t(1-\delta)}{2\delta} = c_A + \frac{t(1+y(1-\delta))}{\delta(2y-1)},$$

which implies $p^*_{A2} = c_A + \frac{t(3-\delta)}{\delta(2y-1)}$. The demand for A_1B in equilibrium is then $\hat{x} = \frac{1-\delta+y(1+\delta)}{2(2y-1)}$ and for A_2B we have $1-\hat{x} = \frac{(3-\delta)(y-1)}{2(2y-1)}$. Thus, we obtain $p^*_B = v - \delta c_A - \frac{t(3-\delta)(1+y)}{2(2y-1)}$. The equilibrium profits are

$$\pi^*_{\hat{M}} = \tilde{v} - \frac{t(2\delta(2y+3y^2-1) + \delta^2(1-2y-y^2) - (1+y)^2)}{2\delta(2y-1)^2},$$
$$\pi^*_2 = \frac{t(3-\delta)^2(y-1)}{2\delta(2y-1)^2}.$$

We verify now that p^*_{A2} is indeed the best response to p^*_{A1} and p^*_B. First, suppose the best response of firm 2, given p^*_{A1} and p^*_B, leads to an uncovered market. Then, $\pi_2 = (p_{A2} - c_A)(\frac{\delta(c_A - p_{A2})}{t} + \frac{(3-\delta)(1+y)}{2(2y-1)})$. The first-order condition yields $p_{A2} = c_A + \frac{t(3-\delta)(1+y)}{4\delta(2y-1)}$. Note that this solution is only feasible when $p_{A2} > p^*_{A2}$. Otherwise, the market is covered.

$$c_A + \frac{t(3-\delta)(1+y)}{4\delta(2y-1)} > c_A + \frac{t(3-\delta)}{\delta(2y-1)}$$
$$y > 3$$

A contradiction since $y \in [2,3]$. Thus, the best response of firm 2, given p^*_{A1} and p^*_B, must lead to a cover market.

Then, we have $\pi_2^{cov} = (p_{A2} - c_A)(\frac{1}{2} + \frac{1+y(1-\delta)}{2(2y-1)} + \frac{\delta(c_A - p_{A2})}{2t})$. The first-order condition yields $p_{A2} = c_A + \frac{ty(3-\delta)}{2\delta(2y-1)}$. Note that this solution is only feasible when $p_{A2} \leq p^*_{A2}$. Otherwise, the market becomes uncovered.

$$c_A + \frac{ty(3-\delta)}{2\delta(2y-1)} \leq c_A + \frac{t(3-\delta)}{\delta(2y-1)}$$
$$y \leq 2$$

Observe that $p_2 = c_A + \frac{t(3-\delta)}{3\delta} = p^*_{A2}$ for $y = 2$. For any $y \in (2,3]$, firm 2 sets the largest price p_{A2} that just covers the market. We can observe this from

$$\frac{\partial \pi_2^{cov}}{\partial p_{A2}} = \frac{y(3-\delta)}{2(2y-1)} + \frac{\delta(c_A - p_{A2})}{t},$$

which is positive when $\delta(p_{A2} - c_A) < \frac{ty(3-\delta)}{2(2y-1)}$. In this case, it is optimal to set the largest price that just covers the market. Plugging in p^*_{A2}, which is this largest price, we obtain $\frac{\partial \pi_2^{cov}}{\partial p_{A2}} > 0$ if $y > 2$. Hence, for $y \in [2,3]$, $p^*_{A2} = c_A + \frac{t(3-\delta)}{\delta(2y-1)}$ is indeed the best response for firm 2.

Further, given Assumption 1, no profitable deviation for the merged firm exists. Suppose $p^*_B + \epsilon$. Then

$$\begin{aligned}
\hat{\pi}^d_M &= \left(\tilde{v} - \frac{t(3-\delta)(1+y)}{2(2y-1)} + \epsilon\right)\left(1 - \frac{2\epsilon}{t}\right) \\
&\quad + \frac{t(1+y(1-\delta))}{\delta(2y-1)}\left(\frac{1-\delta+y(1+\delta)}{2(2y-1)} - \frac{\epsilon}{t}\right) \\
&= \pi^*_{\hat{M}} + \epsilon\left[\frac{y(6\delta-\delta^2-1)-(1-\delta)^2}{\delta(2y-1)} - \frac{2(\tilde{v}+\epsilon)}{t}\right],
\end{aligned}$$

which is profitable if

$$\begin{aligned}\frac{y(6\delta-\delta^2-1)-(1-\delta)^2}{\delta(2y-1)} &> \frac{2(\tilde{v}+\epsilon)}{t}\\ t &> \frac{2\delta(2y-1)(\tilde{v}+\epsilon)}{y(6\delta-\delta^2-1)-(1-\delta)^2} > \frac{\bar{v}}{2}.\end{aligned}$$

The last inequality holds since $\tilde{v} > \bar{v}$, $\epsilon > 0$ and

$$\begin{aligned}\frac{2\delta(2y-1)}{y(6\delta-\delta^2-1)-(1-\delta)^2} &> \frac{1}{2}\\ 1+y+\delta(y(2+\delta)-6+\delta) &> 0,\end{aligned}$$

where the left-hand side is strictly increasing in y and clearly holds for $y = 2$. Thus, a contradiction given Assumption 1 and the deviation is not profitable. Suppose now $p^*_{A1} + \epsilon$. Then, similarly to above,

$$\begin{aligned}\hat{\pi}^d_M &= \left(\tilde{v} - \frac{t(3-\delta)(1+y)}{2(2y-1)}\right)\left(1-\frac{\delta\epsilon}{t}\right)\\ &+ \left(\frac{t(1+y(1-\delta))}{\delta(2y-1)}+\epsilon\right)\left(\frac{1-\delta+y(1+\delta)}{2(2y-1)}-\frac{\delta\epsilon}{t}\right)\\ &= \pi^*_{\hat{M}} + \epsilon\left[\frac{y(6\delta-\delta^2-1)-(1-\delta)^2}{2(2y-1)} - \frac{\delta(\tilde{v}+\epsilon)}{t}\right],\end{aligned}$$

which is profitable if

$$\begin{aligned}\frac{y(6\delta-\delta^2-1)-(1-\delta)^2}{2(2y-1)} &> \frac{\delta(\tilde{v}+\epsilon)}{t}\\ t &> \frac{2\delta(2y-1)(\tilde{v}+\epsilon)}{y(6\delta-\delta^2-1)-(1-\delta)^2} > \frac{\bar{v}}{2}.\end{aligned}$$

Similar to before, a contradiction. Lastly, suppose $p^*_{A1} - \epsilon$. Then

$$\begin{aligned}\hat{\pi}^d_M &= \tilde{v} - \frac{t(3-\delta)(1+y)}{2(2y-1)} + \left(\frac{t(1+y(1-\delta))}{\delta(2y-1)} - \epsilon\right)\left(\frac{1-\delta+y(1+\delta)}{2(2y-1)} + \frac{\delta\epsilon}{2t}\right)\\ &= \pi^*_{\hat{M}} + \epsilon\left[\frac{1+y(1-\delta)}{2(2y-1)} - \frac{1-\delta+y(1+\delta)}{2(2y-1)} + \frac{\delta\epsilon}{2t}\right] = \pi^*_{\hat{M}} - \delta\epsilon\left[\frac{1}{2}+\frac{\epsilon}{2t}\right] < \pi^*_{\hat{M}}.\end{aligned}$$

(ii) First, suppose firm 2 is not active, and the market is not covered. Then, $\hat{M}$ maximizes $\pi_{\hat{M}} = (p_{A1} + p_B - c_A - c_B)(\frac{v-\delta p_{A1}-p_B}{t})$. It must be $p_B \geq v - \delta p_{A2}$. Otherwise, firm 2 could gain positive market share (and profits) by setting $p_{A2} = c_A + \epsilon$. Since consumers underweight p_{A1} by δ, firm $\hat{M}$ rather wants to increase p_{A1} than p_B and set p_B as low as possible. Thus, we can suppose $p_B = v - \delta p_{A2}$, which just excludes firm 2. Then, $\pi_{\hat{M}} = (p_{A1} + v - \delta p_{A2} - c_A - c_B)(\frac{\delta(p_{A2}-p_{A1})}{t})$. Maximizing yields

$$p_{A1} = \frac{(1+\delta)p_{A2} - \bar{v}}{2}, \quad D_{A_1B} = \frac{\delta(\bar{v} + (1-\delta)p_{A2})}{2t}, \quad \pi_{\hat{M}} = \frac{\delta(\bar{v} + (1-\delta)p_{A2})^2}{4t}.$$

Observe that it must be $\bar{v} + (1-\delta)p_{A2} < \frac{2t}{\delta}$. Otherwise, the market is not uncovered. This implies $\pi_{\hat{M}} = \frac{\delta}{4t}(\bar{v}+(1-\delta)p_{A2})^2 < \frac{\delta}{4t}\left(\frac{2t}{\delta}\right)^2 = \frac{t}{\delta}$. Consider the deviation $p^d_B = 0$, and p^d_{A1} maximizes $\pi^d_{\hat{M}} = (p^d_{A1} - c_A - c_B)(\max\{\frac{v-\delta p^d_{A1}}{t}, 1\})$. Suppose $\frac{v-\delta p^d_{A1}}{t} < 1$. Then maximizing yields $p^d_{A1} = \frac{v+\delta c_A+\delta c_B}{2\delta}$, implying $D_{A_1B} = \frac{v-\delta c_A-\delta c_B}{2t} > \frac{\bar{v}}{2t} > 1$. A contradiction given Assumption 1. Thus, it must be $p^d_{A1} = \frac{v-t}{\delta}$, $D_{A_1B} = 1$ and $\pi^d_{\hat{M}} = \frac{v-t}{\delta} - c_A - c_B$, which is strictly profitable given Assumption 1

$$\begin{aligned} \frac{v-t}{\delta} - c_A - c_B &> \frac{t}{\delta} \\ v - \delta c_A - \delta c_B &> 2t \\ \frac{v - \delta c_A - \delta c_B}{2} &> \frac{\bar{v}}{2} > t, \end{aligned}$$

implying $\frac{v-t}{\delta} - c_A - c_B > \frac{t}{\delta} > \frac{\delta(\bar{v}+(1-\delta)p_{A2})^2}{4t}$. Thus, given Assumption 1, there exists no equilibrium, in which firm 2 is excluded and the market is uncovered.

Now, suppose firm 2 is not active, the market is covered, and $t \leq \delta c_A$. For any $p_{A1} \leq c_A - \frac{t}{\delta}$, firm 2 has zero demand. To cover the market, firm $\hat{M}$ then sets $p_B = v - \delta p_{A1} - t$. Then, $p_{A1} + p_B = v - t + (1-\delta)p_{A1}$ and $\pi_{\hat{M}} = \bar{v} - t + (1-\delta)p_{A1}$. Observe that profits are strictly increasing in p_{A1}. Hence, the constraint on p_{A1} must hold with equality to maximize profits,

$p_{A1} = c_A - \frac{t}{\delta}$. This implies $\pi_{\hat{M}} = v - \delta c_A - c_B - \frac{t}{\delta}$.

Consider again the deviation $p^d_B = 0$, and $p^d_{A1} = \frac{v-t}{\delta}$, leading to $\pi^d_{\hat{M}} = \frac{v-t}{\delta} - c_A - c_B$, which is strictly profitable,

$$\begin{aligned}\frac{v-t}{\delta} - c_A - c_B &> v - \delta c_A - c_B - \frac{t}{\delta}\\ (1-\delta)v - (1-\delta)\delta c_A &> 0\\ (1-\delta)(v - \delta c_A) &> 0\end{aligned}$$

since $v > c_A$. Suppose now $t > \delta c_A$. Then, $p_{A1} \leq c_A - \frac{t}{\delta}$ is not feasible since prices must be non-negative. Similar to the proof of Lemma 3.3.2, firm 2 can set $p_{A2} = c_A + \epsilon$ and obtain positive demand and profit. Thus, given Assumption 1, there exists no equilibrium, in which firm 2 is excluded and the market is covered. □

3.A.6 Proof of Lemma 3.4.3

The benefit of merging is

$$\begin{aligned}\Delta\pi &= \pi^*_{\hat{M}} - \pi^*_M - \pi^*_i\\ &= \tilde{v} - \frac{t(2\delta(2y+3y^2-1) + \delta^2(1-2y-y^2) - (1+y)^2)}{2\delta(2y-1)^2} - \tilde{v} + \frac{t(\delta(1+y)-1)}{2\delta(y-1)}\\ &= \frac{t(2+\delta(y-1)-y)[y-1-y^2+\delta(y(2+y)-1)]}{2\delta(y-1)(2y-1)^2}.\end{aligned}$$

Thus, it follows that merger incentives are satisfied iff

$$F < \frac{t(2+\delta(y-1)-y)[y-1-y^2+\delta(y(2+y)-1)]}{2\delta(y-1)(2y-1)^2}.$$

Observe that when $\bar{\delta}_1 \leq \delta \leq \bar{\delta}_2$, then $2 + \delta(y-1) - y \leq 0$ and $y - 1 - y^2 + \delta(y(2+y)-1) \geq 0$. Hence it follows $\Delta\pi \leq 0 < F$ for $\delta \in [\bar{\delta}_1, \bar{\delta}_2]$. Otherwise, both terms are either positive or negative, and thus, merger incentives are positive.

□

3.A.7 Proof of Proposition 3.4.1

The merger incentives with $\delta = 1$ from Lemma 3.3.3 are $\frac{t(3y-2)}{2(y-1)(2y-1)^2}$. The merger incentives with $\delta \in (0,1)$ from Lemma 3.4.3 are $\frac{t(2+\delta(y-1)-y)[y-1-y^2+\delta(y(2+y)-1)]}{2\delta(y-1)(2y-1)^2}$. Hence, for $\delta \in (0,1)$, merger incentives are lower when

$$\begin{aligned}\frac{t(2+\delta(y-1)-y)[y-1-y^2+\delta(y(2+y)-1)]}{2\delta(y-1)(2y-1)^2} &< \frac{t(3y-2)}{2(y-1)(2y-1)^2}\\ t(2+\delta(y-1)-y)[y-1-y^2+\delta(y(2+y)-1)] - t\delta(3y-2) &< 0\\ t(\delta-1)(2+\delta-3(1+\delta)y+(3+\delta)y^2-(1-\delta)y^3) &< 0\end{aligned} \tag{3.4}$$

This inequality holds when

$$\begin{aligned}2+\delta-3(1+\delta)y+(3+\delta)y^2-(1-\delta)y^3 &> 0\\ \delta &> \frac{3y+y^3-2-3y^2}{1-3y+y^2+y^3} = \bar{\delta}_3.\end{aligned}$$

Otherwise, when $\delta \leq \bar{\delta}_3$, then the left-hand side of inequality (3.4) is non-negative and positive for $\delta < \bar{\delta}_3$. Thus, merger incentives are weakly higher with $\delta \in (0, \bar{\delta}_3]$. □

Bibliography

AHMETOGLU, G., A. FURNHAM, AND P. FAGAN (2014): "Pricing practices: A critical review of their effects on consumer perceptions and behaviour," *Journal of Retailing and Consumer Services*, 21, 696–707.

AKERLOF, G. A. (1982): "Labor contracts as partial gift exchange," *The quarterly journal of economics*, 97, 543–569.

AKGÜN, U., C. CAFFARRA, F. ETRO, AND R. STILLMAN (2020): "On the welfare impact of mergers of complements: Raising rivals' costs versus elimination of double marginalization," *Economics Letters*, 195, 109429.

ALEVY, J. E., C. E. LANDRY, AND J. A. LIST (2015): "Field experiments on the anchoring of economic valuations," *Economic Inquiry*, 53, 1522–1538.

AMIR, R., H. LIU, D. MACHOWSKA, AND J. RESENDE (2019): "Spillovers, subsidies, and second-best socially optimal R&D," *Journal of Public Economic Theory*, 21, 1200–1220.

APFFELSTAEDT, A. AND L. MECHTENBERG (2021): "Competition for context-sensitive consumers," *Management Science*, 67, 2828–2844.

ARIELY, D., G. LOEWENSTEIN, AND D. PRELEC (2003): "Coherent Arbitrariness: Stable Demand Curves Without Stable Preferences," *The Quarterly Journal of Economics*, 118, 73–106.

ARMSTRONG, M. (2015): "Search and ripoff externalities," *Review of Industrial Organization*, 47, 273–302.

ARMSTRONG, M. AND Y. CHEN (2009): "Inattentive consumers and product quality," *Journal of the European Economic Association*, 7, 411–422.

ARMSTRONG, M. AND J. VICKERS (2012): "Consumer protection and contingent charges," *Journal of Economic Literature*, 50, 477–93.

ASCHHOFF, B. AND T. SCHMIDT (2008): "Empirical evidence on the success of R&D cooperation – happy together?" *Review of Industrial Organization*, 33, 41–62.

AZAR, O. H. (2007): "Relative thinking theory," *The Journal of Socio-Economics*, 36, 1–14.

——— (2011): "Do people think about absolute or relative price differences when choosing between substitute goods?" *Journal of Economic Psychology*, 32, 450–457.

——— (2019): "Do fixed payments affect effort? Examining relative thinking in mixed compensation schemes," *Journal of Economic Psychology*, 70, 52–66.

BAGWELL, K. AND R. W. STAIGER (1994): "The sensitivity of strategic and corrective R&D policy in oligopolistic industries," *Journal of International Economics*, 36, 133–150.

BARDEY, D., B. JULLIEN, AND J.-M. LOZACHMEUR (2016): "Health insurance and diversity of treatment," *Journal of Health Economics*, 47, 50–63.

BAVLY, G., Y. HELLER, AND A. SCHREIBER (2022): "Social welfare in search games with asymmetric information," *Journal of Economic Theory*, 202, 105462.

BAYE, M. R. AND J. MORGAN (2001): "Information gatekeepers on the internet and the competitiveness of homogeneous product markets," *American Economic Review*, 91, 454–474.

Becker, W. and J. Dietz (2004): "R&D cooperation and innovation activities of firms – evidence for the German manufacturing industry," *Research Policy*, 33, 209–223.

Beggs, A. and K. Graddy (2009): "Anchoring Effects: Evidence from Art Auctions," *American Economic Review*, 99, 1027–39.

Bergman, O., T. Ellingsen, M. Johannesson, and C. Svensson (2010): "Anchoring and cognitive ability," *Economics Letters*, 107, 66–68.

Blake, T., S. Moshary, K. Sweeney, and S. Tadelis (2021): "Price salience and product choice," *Marketing Science*, 40, 619–636.

Boone, J. (2008a): "Competition: Theoretical parameterizations and empirical measures," *Journal of Institutional and Theoretical Economics (JITE)/Zeitschrift für die gesamte Staatswissenschaft*, 587–611.

——— (2008b): "A new way to measure competition," *The Economic Journal*, 118, 1245–1261.

Bordalo, P., N. Gennaioli, and A. Shleifer (2012): "Salience theory of choice under risk," *The Quarterly journal of economics*, 127, 1243–1285.

——— (2013): "Salience and consumer choice," *Journal of Political Economy*, 121, 803–843.

——— (2016): "Competition for attention," *The Review of Economic Studies*, 83, 481–513.

——— (2020): "Memory, attention, and choice," *The Quarterly journal of economics*, 135, 1399–1442.

——— (2022): "Salience," *Annual Review of Economics*, 14, 521–44.

Bourreau, M., B. Jullien, and Y. Lefouili (2021): "Mergers and demand-enhancing innovation," *TSE Working Papers No. 18-907.*

Bracha, A., U. Gneezy, and G. Loewenstein (2015): "Relative pay and labor supply," *Journal of labor economics*, 33, 297–315.

Brown, J., T. Hossain, and J. Morgan (2010): "Shrouded attributes and information suppression: Evidence from the field," *The Quarterly Journal of Economics*, 125, 859–876.

Brunner, P., I. Letina, and A. Schmutzler (2022): "Research joint ventures: The role of financial constraints," *arXiv preprint arXiv:2207.04856v1*.

Bryan, K. A. and J. Lemus (2017): "The direction of innovation," *Journal of Economic Theory*, 172, 247–272.

Bushong, B., M. Rabin, and J. Schwartzstein (2021): "A model of relative thinking," *The Review of Economic Studies*, 88, 162–191.

Caggese, A. (2019): "Financing constraints, radical versus incremental innovation, and aggregate productivity," *American Economic Journal: Macroeconomics*, 11, 275–309.

Caloghirou, Y., S. Ioannides, and N. S. Vonortas (2003): "Research joint ventures," *Journal of Economic Surveys*, 17, 541–570.

Cassiman, B. (2000): "Research joint ventures and optimal R&D policy with asymmetric information," *International Journal of Industrial Organization*, 18, 283–314.

Cassiman, B. and R. Veugelers (2002): "R&D cooperation and spillovers: some empirical evidence from Belgium," *American Economic Review*, 92, 1169–1184.

Chatterjee, P. (2010): "Consumer response to promotions in the presence of surcharge: implications for online retailing," *Journal of Customer Behaviour*, 9, 117–134.

Chen, Y. and Ö. Turut (2013): "Context-dependent preferences and innovation strategy," *Management Science*, 59, 2747–2765.

Chen, Z. and P. Rey (2012): "Loss leading as an exploitative practice," *American Economic Review*, 102, 3462–82.

Chetty, R., A. Looney, and K. Kroft (2009): "Salience and Taxation: Theory and Evidence," *American Economic Review*, 99, 1145–77.

Choi, J. P. (1993): "Cooperative R&D with product market competition," *International Journal of Industrial Organization*, 11, 553–571.

Competition Market Authority (2022): "Evidence Review of Online Choice Architecture and Consumer and Competition Harm," *UK Government Report*.

Czarnitzki, D. and H. Hottenrott (2011): "R&D investment and financing constraints of small and medium-sized firms," *Small Business Economics*, 36, 65–83.

d'Aspremont, C. and A. Jacquemin (1988): "Cooperative and noncooperative R&D in duopoly with spillovers," *American Economic Review*, 78, 1133–1137.

Denicolò, V. and M. Polo (2018): "Duplicative research, mergers and innovation," *Economics Letters*, 166, 56–59.

Department of Transportation (2022): "Enhancing Transparency of Airline Ancillary Service Fees," *Federal Register*, 2022-22214, 63718–63738.

Dertwinkel-Kalt, M., K. Köhler, M. R. Lange, and T. Wenzel (2017): "Demand shifts due to salience effects: Experimental evidence," *Journal of the European Economic Association*, 15, 626–653.

Dertwinkel-Kalt, M. and M. Köster (2020): "Salience and skewness preferences," *Journal of the European Economic Association*, 18, 2057–2107.

——— (2022): "Attention to online sales: The role of brand image concerns," *Journal of Economics & Management Strategy*, 31, 64–89.

Dertwinkel-Kalt, M., M. Köster, and M. Sutter (2020): "To buy or not to buy? Price salience in an online shopping field experiment," *European Economic Review*, 130, 103593.

Dessaint, O. and A. Matray (2017): "Do managers overreact to salient risks? Evidence from hurricane strikes," *Journal of Financial Economics*, 126, 97–121.

Diamond, P. A. (1971): "A model of price adjustment," *Journal of economic theory*, 3, 156–168.

Duso, T., L.-H. Röller, and J. Seldeslachts (2014): "Collusion through joint R&D: An empirical assessment," *Review of Economics and Statistics*, 96, 349–370.

Economides, N. (2005): "The incentive for vertical integration," *Available at SSRN 692041.*

Economides, N. and S. C. Salop (1992): "Competition and integration among complements, and network market structure," *The Journal of Industrial Economics*, 105–123.

Ellison, G. (2005): "A model of add-on pricing," *The Quarterly Journal of Economics*, 120, 585–637.

Ellison, G. and S. F. Ellison (2009): "Search, obfuscation, and price elasticities on the internet," *Econometrica*, 77, 427–452.

Erat, S. and S. R. Bhaskaran (2012): "Consumer mental accounts and implications to selling base products and add-ons," *Marketing Science*, 31, 801–818.

Falvey, R., J. Poyago-Theotoky, and K. T. Teerasuwannajak (2013): "Coordination costs and research joint ventures," *Economic Modelling*, 33, 965–976.

Farrell, J. and C. Shapiro (2000): "Scale economies and synergies in horizontal merger analysis," *Antitrust LJ*, 68, 685.

Fauli-Oller, R. and J. Sandonis (2003): "To merge or to license: implications for competition policy," *International Journal of Industrial Organization*, 21, 655–672.

Federal Trade Commission (2012): "The Economics of Drip Pricing," Tech. rep., U.S. Government.

——— (2022): "Unfair or Deceptive Fees Trade Regulation Rule," *Federal Register*, 87 FR 67413, 67413 – 67424.

Federico, G., G. Langus, and T. Valletti (2017): "A simple model of mergers and innovation," *Economics Letters*, 157, 136–140.

——— (2018): "Horizontal mergers and product innovation," *International Journal of Industrial Organization*, 59, 1–23.

Fehr, E., G. Kirchsteiger, and A. Riedl (1993): "Does fairness prevent market clearing? An experimental investigation," *The quarterly journal of economics*, 108, 437–459.

Feldman, N. E. and B. J. Ruffle (2015): "The impact of including, adding, and subtracting a tax on demand," *American Economic Journal: Economic Policy*, 7, 95–118.

Ferraro, P., K. D. Messer, P. Shukla, and C. Weigel (2021): "Behavioral Biases among Producers: Experimental Evidence of Anchoring in Procurement Auctions," .

Finkelstein, A. (2009): "E-ztax: Tax salience and tax rates," *The Quarterly Journal of Economics*, 124, 969–1010.

Frisch, D. (1993): "Reasons for framing effects," *Organizational behavior and human decision processes*, 54, 399–429.

Fudenberg, D., D. K. Levine, and Z. Maniadis (2012): "On the robustness of anchoring effects in WTP and WTA experiments," *American Economic Journal: Microeconomics*, 4, 131–45.

Fumagalli, C., M. Motta, and E. Tarantino (2022): "Shelving or developing? The acquisition of potential competitors under financial constraints," *CSEW Working Paper no. 637*.

Furnham, A. and H. C. Boo (2011): “A literature review of the anchoring effect,” *The journal of socio-economics*, 40, 35–42.

Gabaix, X. and D. Laibson (2006): “Shrouded attributes, consumer myopia, and information suppression in competitive markets,” *The Quarterly Journal of Economics*, 121, 505–540.

Gamp, T. (2015): “Search, differentiated products, and obfuscation,” *Beitraege zur Jahrestagung des Vereins fÃ¼r Socialpolitik 2015: Oekonomische Entwicklung - Theorie und Politik - Session: Monopoly and Oligopoly No. D24-V2.*

Gilbert, R. (2019): “Competition, mergers, and R&D diversity,” *Review of Industrial Organization*, 465–484.

Goyal, S. and J. L. Moraga-Gonzalez (2001): “R&d networks,” *Rand Journal of Economics*, 686–707.

Greenleaf, E. A., E. J. Johnson, V. G. Morwitz, and E. Shalev (2016): “The price does not include additional taxes, fees, and surcharges: A review of research on partitioned pricing,” *Journal of Consumer Psychology*, 26, 105–124.

Grossman, G. M. and C. Shapiro (1986): “Research joint ventures: an antitrust analysis,” *JL Econ. & Org.*, 2, 315.

Hall, B. H. and J. Lerner (2010): “The financing of R&D and innovation,” in *Handbook of the Economics of Innovation*, Elsevier, vol. 1, 609–639.

Hart, O. and J. Moore (2008): “Contracts as reference points,” *The Quarterly journal of economics*, 123, 1–48.

Heeb, R. (2003): “Randal heeb innovation and vertical integration in complementary markets,” *Journal of Economics & Management Strategy*, 12, 387–417.

HEIDHUES, P., J. JOHNEN, AND B. KŐSZEGI (2021): "Browsing versus studying: A pro-market case for regulation," *The Review of Economic Studies*, 88, 708–729.

HEIDHUES, P. AND B. KŐSZEGI (2017): "Naivete-based discrimination," *The Quarterly Journal of Economics*, 132, 1019–1054.

——— (2018): "Behavioral industrial organization," *Handbook of Behavioral Economics: Applications and Foundations 1*, 1, 517–612.

HELFRICH, M. AND F. HERWEG (2017): "Salience in retailing: Vertical restraints on internet sales," .

HERWEG, F., D. MÜLLER, AND P. WEINSCHENK (2018): "Salience in markets," in *Handbook of Behavioral Industrial Organization*, Edward Elgar Publishing, 75–113.

HERZ, H. AND D. TAUBINSKY (2017): "What Makes a Price Fair? An Experimental Study of Transaction Experience and Endogenous Fairness Views," *Journal of the European Economic Association*, 16, 316–352.

HOLTON, R. H. (1957): "Price discrimination at retail: the supermarket case," *The Journal of Industrial Economics*, 6, 13–32.

HOSSAIN, T. AND J. MORGAN (2006): "... plus shipping and handling: Revenue (non) equivalence in field experiments on ebay," *The BE Journal of Economic Analysis & Policy*, 6.

HOTELLING, H. (1929): "Stability in Competition," *The Economic Journal*, 39, 41–57.

HOWELL, S. T. (2017): "Financing innovation: Evidence from R&D grants," *American Economic Review*, 107, 1136–64.

HUCK, S. AND B. WALLACE (2015): "The impact of price frames on consumer decision making: Experimental evidence," *Experimental evidence*, 1–47.

Inderst, R. and M. Obradovits (2020): "Loss leading with salient thinkers," *The RAND Journal of Economics*, 51, 260–278.

——— (2023): "Excessive competition on headline prices," *International Economic Review*, 64, 783–808.

Ioannidis, K., T. Offerman, and R. Sloof (2020): "On the effect of anchoring on valuations when the anchor is transparently uninformative," *Journal of the Economic Science Association*, 6, 77–94.

Jacowitz, K. E. and D. Kahneman (1995): "Measures of anchoring in estimation tasks," *Personality and Social Psychology Bulletin*, 21, 1161–1166.

Jacquemin, A. (1988): "Cooperative agreements in R&D and European antitrust policy," *European Economic Review*, 32, 551–560.

Johnson, J. P. (2017): "Unplanned purchases and retail competition," *American Economic Review*, 107, 931–65.

Kahneman, D. and A. Tversky (1979): "Prospect theory: An analysis of decision under risk," *Econometrica*, 47, 363–391.

Kamien, M. I., E. Muller, and I. Zang (1992): "Research joint ventures and R&D cartels," *American Economic Review*, 1293–1306.

Kamien, M. I. and I. Zang (2000): "Meet me halfway: Research joint ventures and absorptive capacity," *International Journal of Industrial Organization*, 18, 995–1012.

Karle, H., F. Kerzenmacher, H. Schumacher, and F. Verboven (2022): "Search Costs and Diminishing Sensitivity," .

Katz, M. L. (1986): "An analysis of cooperative research and development," *RAND Journal of Economics*, 527–543.

Katz, M. L. and C. Shapiro (1985): "On the licensing of innovations," *RAND Journal of Economics*, 16, 504–520.

Kerr, W. R. and R. Nanda (2015): "Financing innovation," *Annual Review of Financial Economics*, 7, 445–462.

Kosfeld, M. and U. Schüwer (2016): "Add-on Pricing in Retail Financial Markets and the Fallacies of Consumer Education*," *Review of Finance*, 21, 1189–1216.

Kőszegi, B. and A. Szeidl (2013): "A model of focusing in economic choice," *The Quarterly journal of economics*, 128, 53–104.

Krieger, J., D. Li, and D. Papanikolaou (2022): "Missing novelty in drug development," *Review of Financial Studies*, 35, 636–679.

Leahy, D. and J. P. Neary (1997): "Public policy towards R&D in oligopolistic industries," *American Economic Review*, 642–662.

Letina, I. (2016): "The road not taken: competition and the R&D portfolio," *RAND Journal of Economics*, 47, 433–460.

Letina, I. and A. Schmutzler (2019): "Inducing variety: A theory of innovation contests," *International Economic Review*, 60, 1757–1780.

Letina, I., A. Schmutzler, and R. Seibel (2021): "Killer acquisitions and beyond: policy effects on innovation strategies," *University of Zurich, Department of Economics Working Paper no. 358.*

Link, A. N. (1998): "Case study of R&D efficiency in an ATP joint venture," *The Journal of Technology Transfer*, 23, 43–51.

Mancusi, M. L. and A. Vezzulli (2014): "R&D and credit rationing in SMEs," *Economic Inquiry*, 52, 1153–1172.

Maniadis, Z., F. Tufano, and J. A. List (2014): "One Swallow Doesn't Make a Summer: New Evidence on Anchoring Effects," *American Economic Review*, 104, 277–90.

Martin, S. (1996): "R&D joint ventures and tacit product market collusion," *European Journal of Political Economy*, 11, 733–741.

Michel, C. (2017): "Market regulation of voluntary add-on contracts," *International Journal of Industrial Organization*, 54, 239–268.

Miyagiwa, K. (2009): "Collusion and research joint ventures," *Journal of Industrial Economics*, 57, 768–784.

Mohnen, P., F. C. Palm, S. S. Van Der Loeff, and A. Tiwari (2008): "Financial constraints and other obstacles: Are they a threat to innovation activity?" *De Economist*, 156, 201–214.

Moon, P., K. Keasey, and D. Duxbury (1999): "Mental accounting and decision making: The relationship between relative and absolute savings," *Journal of Economic Behavior & Organization*, 38, 145–153.

Moraga-González, J. L., E. Motchenkova, and S. Nevrekar (2022): "Mergers and innovation portfolios," *The RAND Journal of Economics*, 53, 641–677.

Morwitz, V. G., E. A. Greenleaf, and E. J. Johnson (1998): "Divide and prosper: consumers' reactions to partitioned prices," *Journal of marketing research*, 35, 453–463.

Motta, M. and E. Tarantino (2021): "The effect of horizontal mergers, when firms compete in prices and investments," *International Journal of Industrial Organization*, 78, 102774.

Mowen, M. M. and J. C. Mowen (1986): "An empirical examination of the biasing effects of framing on business decisions," *Decision Sciences*, 17, 596–602.

Narasimhan, C. and Ö. Turut (2013): "Differentiate or imitate? The role of context-dependent preferences," *Marketing Science*, 32, 393–410.

Office of Fair Trading (2013): "Partitioned Pricing Research: The economics of partitioned pricing," *UK Government Report*.

Rabin, M. (1993): "Incorporating fairness into game theory and economics," *The American economic review*, 1281–1302.

Ranyard, R. and D. Abdel-Nabi (1993): "Mental accounting and the process of multiattribute choice," *Acta Psychologica*, 84, 161–177.

Rasch, A., M. Thöne, and T. Wenzel (2020): "Drip pricing and its regulation: Experimental evidence," *Journal of Economic Behavior & Organization*, 176, 353–370.

Robbert, T. and S. Roth (2014): "The flip side of drip pricing," *Journal of Product & Brand Management*, 23, 413–419.

Röller, L.-H., R. Siebert, and M. Tombak, Mikhel (2007): "Why firms form (or do not form) RJVs," *Economic Journal*, 117, 1122–1144.

Santana, S., S. K. Dallas, and V. G. Morwitz (2020): "Consumer reactions to drip pricing," *Marketing Science*, 39, 188–210.

Savignac, F. (2008): "Impact of financial constraints on innovation: What can be learned from a direct measure?" *Economics of Innovation and New Technology*, 17, 553–569.

Schmutzler, A. (2013): "Competition and investment: A unified approach," *International Journal of Industrial Organization*, 31, 477–487.

Sexton, S. (2015): "Automatic bill payment and salience effects: Evidence from electricity consumption," *Review of Economics and Statistics*, 97, 229–241.

Shapiro, C. (1985): "Patent licensing and R&D rivalry," *American Economic Review, Papers and Proceedings*, 75, 25–30.

——— (1994): "Aftermarkets and Consumer Welfare: Making Sense of Kodak Symposium: Post-Chicago Economics," *Antitrust Law Journal*, 63, 483.

Singh, N. and X. Vives (1984): "Price and quantity competition in a differentiated duopoly," *The Rand journal of economics*, 546–554.

Somerville, J. (2022): "Range-Dependent Attribute Weighting in Consumer Choice: An Experimental Test," *Econometrica*, 90, 799–830.

Sovinsky, M. (2022): "Do research joint ventures serve a collusive function?" *Journal of the European Economic Association*, 20, 430–475.

Spiegler, R. (2006): "Competition over agents with boundedly rational expectations," *Theoretical Economics*, 1, 207–231.

——— (2011): *Bounded rationality and industrial organization*, Oxford University Press.

——— (2016): "Choice complexity and market competition," *Annual Review of Economics*, 8, 1–25.

Suzumura, K. (1992): "Cooperative and noncooperative R&D in an oligopoly with spillovers," *American Economic Review*, 1307–1320.

Taubinsky, D. and A. Rees-Jones (2018): "Attention variation and welfare: theory and evidence from a tax salience experiment," *The Review of Economic Studies*, 85, 2462–2496.

Thaler, R. (1980): "Toward a positive theory of consumer choice," *Journal of economic behavior & organization*, 1, 39–60.

——— (1985): "Mental accounting and consumer choice," *Marketing science*, 4, 199–214.

The White House (2022): "The President's Initiative on Junk Fees and Related Pricing Practices," Tech. rep., U.S. Government.

——— (2023): "Remarks of President Joe Biden - State of the Union Address as Prepared for Delivery," Tech. rep., U.S. Government.

Tufano, F. (2010): "Are true preferences revealed in repeated markets? An experimental demonstration of context-dependent valuations," *Experimental Economics*, 13, 1–13.

Tversky, A. and D. Kahneman (1974): "Judgment under Uncertainty: Heuristics and Biases," *Science*, 185, 1124–1131.

——— (1981): "The Framing of Decisions and the Psychology of Choice," *Science*, 211, 453–458.

Veugelers, R. (1997): "Internal R&D expenditures and external technology sourcing," *Research Policy*, 26, 303–315.

Vilasuso, J. and M. R. Frascatore (2000): "Public policy and R&D when research joint ventures are costly," *Canadian Journal of Economics/Revue canadienne d'économique*, 33, 818–839.

Wenner, L. M. (2015): "Expected prices as reference points-theory and experiments," *European Economic Review*, 75, 60–79.

Xia, L. and K. B. Monroe (2004): "Price partitioning on the internet," *Journal of Interactive marketing*, 18, 63–73.

Yin, Q., B. Jiang, and S. X. Zhou (2023): "Effects of consumers' context-dependent preferences on product bundling," *Production and Operations Management*, 32, 1674–1691.

Yoon, S., N. M. Fong, and A. Dimoka (2019): "The robustness of anchoring effects on preferential judgments." *Judgment & Decision Making*, 14.

Statement of Authorship

Selbständigkeitserklärung

Ich erkläre hiermit, dass ich diese Arbeit selbständig verfasst und keine anderen als die angegebenen Quellen benutzt habe. Alle Stellen, die wörtlich oder sinngemäss aus Quellen entnommen wurden, habe ich als solche gekennzeichnet. Mir ist bekannt, dass andernfalls der Senat gemäss Artikel 36 Absatz 1 Buchstabe o des Gesetzes vom 5. September 1996 über die Universität zum Entzug des aufgrund dieser Arbeit verliehenen Titels berechtigt ist.

Bern, 22. Dezember 2023

Philipp Brunner