ECONtribute
MARKETS & PUBLIC POLICY

ECONtribute
Discussion Paper No. 153

Imperfect Competition in Derivatives Markets

Christina Brinkmann

March 2022

www.econtribute.de

UNIVERSITÄT BONN
Universität zu Köln

Funding by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany´s Excellence Strategy – EXC 2126/1– 390838866 is gratefully acknowledged.

Imperfect competition in derivatives markets

Christina Brinkmann*

This version: March 20, 2022

Abstract

Since the push towards central clearing in derivatives markets after the global financial crisis, an open question has been how the development has affected competition. This paper models imperfect competition between dealers in derivatives markets. Two risk-neutral dealers offer derivatives to risk-averse clients with a hedging need, and compete in price (fee) and quality (default probability). I find that with such two-dimensional competition, for given default probabilities, an equilibrium in prices exists that is preferred by both dealers. In this equilibrium the dealer with the lower default probability makes larger profits - a feature, that can produce market discipline to keep the own default probability low. If a central counterparty (CCP) is introduced as an innovation that removes the quality dimension of the competition, this market force pushing for higher qualities vanishes.

JEL classification: G12, G23, G28, L13, L15

Keywords: Derivatives, OTC Markets, Central Clearing, Imperfect Competition, Vertical Product Differentiation

---

*University of Bonn. E-mail: christina-brinkmann@uni-bonn.de. I thank Christian Kubitza for many helpful conversations particulary in the early stages of the project. I also thank Jo Braithwaite, Dominik Damast, Nektaria Glynia (discussant), Hendrik Hakenes, Tobias Herbst, Eugen Kovac, Stephan Lauermann, David Murphy, Martin Peitz and Farzad Saidi for helpful comments and discussions, as well as participants of the Finance seminar at the University of Bonn, the 14th RGS Doctoral Conference in Economics and the 8th CRC TR 224 Workshop for Young Researchers. I gratefully acknowledge financial support from ECONtribute, funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany´s Excellence Strategy – EXC 2126/1– 390838866. Funding by the German Research Foundation (DFG) through CRC TR 224 (Project C03) is gratefully acknowledged.

1 Introduction

Derivatives markets are large with a notional outstanding of $580 trillion at the end of 2020 (BIS global OTC derivatives statistics). They enable transfer of market risks between entities with different risk bearing capacities (e.g. bank-affiliated dealers, asset managers, non-financial corporations). The opacity and interconnectedness of derivatives markets prior to 2008 played a significant role in amplifying the instabilities of the financial system exposed during the global financial crisis (Gregory, 2014). Ever since, making derivatives markets more resilient has been a priority for policy makers and, in particular, there has been a push towards mandatory central clearing via a central counterparty (CCP). The topic has regained interest in the wake of the market turmoil in March 2020 and a subsequent proposal to centrally clear US Treasury securities and repos (Group of Thirty Working Group on Treasury Market Liquidity, 2021). A CCP interposes itself between two counterparties in derivatives markets and thereby insulates the contracting parties from counterparty risk, i.e. the risk that the counterparty defaults and does not honor its contractual obligations. On the one hand, CCPs can support financial stability through enforced margining (Biais et al., 2016), netting (Duffie and Zhu, 2011) as well as transparency for better regulatory oversight. On the other hand, CCPs change the structure of a highly concentrated market, since most market participants buy derivatives from few large dealers. A key question remains how central clearing affects the oligopolistic competition in derivatives markets.

The first step towards answering this question is to capture the nature of the competition in the absence of central clearing. To that end, I develop a model of imperfect competition in derivatives markets: two risk-neutral dealers sell insurance to clients who wish to hedge against a common macro risk and are heterogeneous in their degree of risk aversion. Dealers can differentiate their products in two dimensions: they choose their own default probability – interpreted as (inverse) insurance quality - and the price. I find that, if default probability is a dimension of the competition, in equilibrium the safer dealer has greater profits, which can produce market discipline for choosing low default probabilities. As explained further below, one simple way of interpreting a CCP then is to see it as an innovation that removes the quality dimension of the competition. In the present model framework this corresponds to pure price competition without market forces that push for higher quality choices.

Even before the push towards CCPs a key feature of derivatives markets has been its hub-and-spoke structure with (typically bank-affiliated) dealers at the core and clients in the periphery (Abad et al., 2016). The number of dealers is typically small (see e.g. Duffie and Zhu (2011)) and thus carefully modeling dealers' market power seems crucial. To that end I carry over key elements from models of vertical product differentiation from the IO literature with two firms (here dealers) and a continuum of consumers (here clients). In our context, the good sold is insurance against a macro risk that is common to all clients and clients differ in their risk aversion.

In the model clients evaluate a derivative in terms of price and default probability of the dealer, i.e. there is competition in two dimensions. For the same price, clients prefer a dealer with a lower default probability (higher quality). Between two dealers with the same default probability (i.e. offering the same quality), they prefer a lower price. One may argue that in practice clients cared much more about the price, e.g. would rather save 0.25 basis points on each swap and accept a 50 basis points higher probability of default of their counterparty. The marginal rates of substitution between default probability and price may be small, but the numerical example from above nonetheless suggests that the trade-off is operating in the background.

In the model dealers choose own default probabilities first and then set the price (or *fee*) for establishing the client-dealer relationship. A way to think about an institution choosing their own default probability is that they decide which measures to undertake to ensure their solvency such as setting aside capital, having balanced trading books, etc. Client's hedging need arises endogenously, they choose which dealer to transact with and subsequently the two parties trade the optimal contract.

My first result is to show that the market is segmented as one would expect: more risk-averse clients buy from the safer dealer at a higher price, while less risk-averse clients buy from the unsafer dealer at a lower price. The main result is existence of a Nash equilibrium in prices for given quality choices. I impose few parameter restrictions to obtain the result and discuss their economic meaning in the text. The equilibrium needs not be unique, but if there are multiple equilibria one is preferred by both sellers.

The second main result is that in any such Nash equilibrium in prices the quality-leader enjoys larger profits. This makes the position of quality-leader more attractive and I demonstrate that this can give rise to upward pressure on qualities. The intuition is simple: When products are differentiated not only in price, but also in quality, firms (i.e. dealers in this context) have an incentive to soften price competition by choosing distinct qualities. Since the high-quality firm has higher profits in equilibrium, this position is the more attractive one. The high-quality firm wants to keep the leadership position in terms of quality and to avoid being overtaken quality-wise by the other firm. This can produce upward pressure on the quality choices - a phenomenon I call *market discipline in qualities*.

An endogenously arising market force pushing for higher qualities is in disconnect with the standard result on vertical product differentiation (Shaked and Sutton, 1982), however, which, in the widely adopted version in (Tirole, 1988), suggests that maximal differentiation in qualities emerges. Therefore, in the appendix I revisit the standard model of vertical product differentiation and show how under less restrictive assumptions qualities are subject to push *and* pull factors and I clarify what it takes to yield endogenous market discipline even in the standard set-up.

The notion of market discipline may not have been the focus previously, because for consumer goods quality does not implicitly embed an aspect of stability of the system. In the insurance and derivatives context, however, the level of costly effort undertaken by individual participants to ensure low levels of own default probability is connected to financial stability. Hence, a market force providing an incentive to ensure a low level of counterparty risk apart from regulation is relevant for an assessment of the market microstructure.

In practice, the hub-and-spoke structure remains unaltered in a derivatives market with a CCP. A CCP steps in between every bilateral derivatives contract, becoming the seller to every buyer and the buyer to every seller. The two contracting parties do not directly face the risk of each other's default anymore, but now depend on the loss absorption capacities of the CCP. Market participants that meet certain criteria can become members of the CCP, enabling them to clear contracts as described and mandating that they participate in loss sharing mechanisms in case another member defaults. Typically, large dealers are clearing members of the CCP, while other market participants access central clearing as their clients (Financial Stability Board (2018)). For example, at the large London-based CCP LCH more than 80% of all client transactions in interest rate derivatives are with five clearing members only (see public disclosure item 19.1.3.2, https://www.lch.com/resources/ccp-disclosures). If two dealers are members of the CCP, from the perspective of the client both dealers offer the same probability of contract continuity. Primarily this is due to porting, that is, in case one clearing member defaults, the portfolios of the clients of the defaulted clearing member get ported to another solvent clearing member, called *backup* clearing member (see e.g. Braithwaite and Murphy (2020) for details).

One straightforward, but perhaps simplistic, way of introducing this into the model is to assume that the presence of a CCP removes any difference in quality between contracts. In that case pure price competition prevails, removing the upward pressure on the qualities from before and providing no incentive to keep qualities above the regulatory minimum.

One may wonder whether the phenomenon may simply be seen as a form of moral hazard. The seminal paper by Biais et al. (2016) embeds a moral hazard problem á la Holmstrom and Tirole (1997) into a derivatives set-up in order to study the role of margins set by a CCP. In their model the seller needs to exert costly effort to ensure solvency, but subsequently may not be the beneficiary of such effort if the derivative turns into a liability instead of an asset for him - something unknown upon trading. Margins then serve as a means to mitigate this moral hazard problem on the side of the seller of the derivative. I do not adopt the moral hazard and information structure from Biais et al. (2016). Instead, in my model whether there are incentives to ensure a low own default probability or not depends on the industrial organization of the market. This could enrich our thinking about the roots of moral hazard beyond the standard effort vs no effort decision. The analysis differs in two other respects. I focus on the competition between two sellers, while Biais et al. (2016) assume perfect competition among a continuum of sellers. Additionally, all buyers and sellers are members of the CCP in Biais et al. (2016), while the set-up in this paper allows for client clearing via a hub-and-spoke structure in which most market participants are clients.

I contribute to a growing literature on derivatives markets and central clearing. Seminal contributions have examined netting benefits (Duffie and Zhu, 2011), transparency (Acharya and Bisin, 2014) and the role of margins (Biais et al., 2012, 2016). Recently, the focus has shifted towards the question how loss sharing mechanisms of CCPs should optimally be designed (Cucic, 2022; Huang and Zhu, 2021; Kuong and Maurin, 2020; Wang et al., 2022), whether loss sharing rules have heterogeneous effects across different types of market participants (Kubitza et al., 2021) and incentives of a profit-maximizing CCP (Capponi and Cheng, 2018; Huang, 2019). Carapella and Monnet (2020) study the effect of central clearing on the entry decision of dealers in derivatives markets. The idea is that, if more dealers enter as a result of the regulation, there is more intense competition and a resulting lower level of spreads may alter incentives to invest in efficient technologies ex ante. A key difference to my model is that in Carapella and Monnet (2020) all agents are risk-neutral and the focus is on search frictions for dealers that intermediate derivatives, as in the search literature on OTC markets pioneered by Duffie et al. (2005).

Another related literature is that on vertical product differentiation, initiated by Gabszewicz and Thisse (1979, 1980). Shaked and Sutton (1982) establish under full market coverage and in the absence of production costs that competition in price and quality produces an incentive to choose distinct qualities in order to soften price competition. Compared to the treatment in Tirole (1988), who otherwise closely follows them, Shaked and Sutton (1982) have a slightly different utility function and an additional entry-stage upfront. Tirole (1988) in fact mentions the generalizations I analyse in the appendix and hypothesizes that the principle of differentiation is more robust nonetheless. This article proves this hypothesis and refines it, as the richer set-up allows for push *and* pull factors impacting the quality decisions. Unlike the work by Moorthy (1988, 1991), who lifts the same assumptions and numerically computes and compares outcomes using quadratic costs, I use a general convex cost function and derive the push factor directly from profit-maximizing incentives.

The rest of the paper is organized as follows. Section 2 presents the model. Section 3 shows market segmentation, section 4 equilibrium existence and section 5 upward pressure on qualities. Section 6 concludes. All proofs are in the appendix.

2 Model

2.1 Set-up

There is a continuum of protection buyers (pb), also called clients, with a hedging need and there are two protection sellers (ps), also called dealers, with the capability to sell derivatives.

Timing. There are four points in time, $t \in \{0, 1, 2, 3\}$. In the first two periods, the protection sellers engage in competition in price and quality. Specifically, in $t = 0$ protection sellers simultaneously choose "quality" in the form of their own default probabilities $b_i, i \in \{1, 2\}$. In $t = 1$ they observe each other's default probability and choose fees $\gamma_i, i \in \{1, 2\}$ for establishing client-dealer relationships.[1] Upon observing the protection sellers' choices (b_1, γ_1) and (b_2, γ_2), protection buyers decide from whom to buy in $t = 2$. Lastly, protection buyers' endowment risk materializes and payments are exchanged in $t = 3$.

Protection buyers. Protection buyers face an uncertain endowment risk $\tilde{x} \in \{\overline{\theta}, \underline{\theta}\}$ with $\underline{\theta} < 0 < \overline{\theta}$. $\tilde{x}$ materializes in $t = 3$ taking the value $\underline{\theta}$ with probability p and $\overline{\theta}$ with probability $(1 - p)$. Suppose $E[\tilde{x}] = 0$.[2] The endowment risk is the *same* across all protection buyers. Protection buyers are risk-averse with strictly increasing and strictly concave utility function $u : \mathbb{R} \to \mathbb{R}$, $u \in \mathcal{C}^2$. Specifically, they have the following utility function that exhibits constant absolute risk aversion (CARA)

$$u_a(x) = \frac{1}{a}\left(1 - \exp(-ax)\right) \qquad a \gneq 0. \tag{1}$$

The limit case $a = 0$ yields $u_0(x) = x$, i.e. risk neutrality for all payments. An increase in a corresponds to being more risk-averse. Protection buyers are characterized by their degree of absolute risk aversion a and a is assumed to be uniformly distributed over the interval $[\underline{a}, \overline{a}]$, $\underline{a} > 0$. (1) parameterizes the degree of absolute risk aversion, while satisfying the following two desirable normalizations: for all positive a, $u_a(0) = 0$ and $u'_a(0) = 1$. The second normalization achieves that, up to a first-order approximation, for small payments the utility coincides with the size of the payment - independent of the degree of risk aversion. This ensures that differences

1 One can think of default probabilities being publicly observable via rating agencies. Their rating should be based upon a detailed evaluation of all available information with particular emphasis on the measures an institution undertakes to ensure its solvency such as setting aside capital, having balanced trading books, etc. Another possibility how at least the ballpark of an institution's default probability can be common knowledge is through rumors in the market. For example, there were rumors on Lehman's insolvency weeks before it actually collapsed.

2 Otherwise $E[\tilde{x}]$ is a certain payment and consider the random variable $\tilde{x} - E[\tilde{x}]$ instead of $\tilde{x}$.

in risk aversion determine different attitudes towards large negative outcomes, but are irrelevant for very small payments.

Protection sellers. Protection sellers are risk-neutral and maximize profits. Protection seller $i \in \{1,2\}$ defaults with probability b_i in the bad endowment state $\underline{\theta}$ and with probability 0 in the good endowment state $\overline{\theta}$. He faces costs c for offering the derivative. As a starting point, the costs c do not vary with the default probability, and let $c = 0$.

Derivative contract. Dealers offer to fully insure clients against their endowment risk in exchange for a fixed payment γ. Dealers set $\gamma_i, i \in \{1,2\}$ in $t = 1$, which we interpret as fees necessary for establishing a client-dealer relationship. A contract (b, γ), sold by protection seller with default probability b, is called *derivative* (b, γ). A protection buyer with risk aversion parameter a derives the following utility from a derivative (b, γ):

$$U_a(b, \gamma) := (1 - bp)u_a(-\gamma) + bpu_a(\underline{\theta}). \tag{2}$$

From the perspective of the protection buyer, the derivative contract swaps the uncertain state-contingent endowment against a fixed payment of γ, unless the protection seller defaults which happens with probability (bp). In that case the protection buyer is left with the original bad endowment.

As shown in Appendix B, this is in fact the outcome of the following optimal contracting problem. Suppose clients in $t = 2$, after deciding from whom to buy, choose trade volumes, i.e. payments (y, z) *from* the protection buyer *to* the protection seller with [3] [4]

y due if $\tilde{x} = \overline{\theta}$ and the protection seller survives,

z due if $\tilde{x} = \underline{\theta}$ and the protection seller survives.

The payments (y, z) that maximize the expected utility of a protection buyer subject to the profit constraint of the dealer, are such that the risk-averse protection buyer receives a state-independent amount, namely $(-\gamma)$, unless the counterparty defaults.

3 All payments are due in $t = 3$. This includes γ, which, although set ex-ante, is also exchanged in $t = 3$ and hence only due if the protection seller survives.

4 $y, z < 0$ indicate that funds flow in the opposite direction, i.e. *from* the protection seller *to* the protection buyer.

Figure 1 summarizes the timing of events as described.

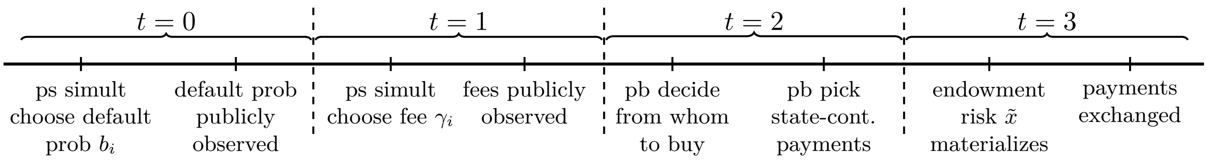

Figure 1: Timeline

2.2 Assumptions and discussion thereof

We impose the following parameter restrictions that in essence ensure that the set-up is such that risk aversion is sufficiently relevant.

Assumption A1. $p < \frac{1}{3}$.

This restricts the range of probabilities of the bad endowment and enables us to bound some objects. One should think of the bad endowment $\underline{\theta}$ as a large negative number and, subsequently, in order to keep the expected endowment zero, p is rather small. Hence, the restriction more or less clarifies the relevant range.

Assumption A2.

$$\underline{a}(-\underline{\theta}) > \log\left[\frac{1-\frac{1}{8}}{\exp(-2)-\frac{1}{8}}\right] \approx 4.4.$$

A lower bound on the degree of risk aversion times the absolute value of the bad endowment, $a(-\underline{\theta})$ for all $a \in [\underline{a}, \overline{a}]$, can be interpreted as follows. Recall that protection buyer a receives utility $u_a(\underline{\theta}) = 1/a \cdot (1 - \exp(-a\underline{\theta}))$ from the bad endowment. Hence, assumption A2 demands that even for the least risk-averse protection buyer the absolute value of the bad endowment and subsequently $\exp(-a\underline{\theta})$ is large enough, such that risk aversion kicks in. Demanding that risk aversion plays a role for all protection buyers is a condition on both the range of a and $\underline{\theta}$. For any large $\underline{\theta}$, one can find a small a such that assumption A2 is violated. Intuitively, for any large payment without limitations on a, one can find protection buyers whose utility is sufficiently close to a risk-neutral one (i.e. a close to 0) such that risk aversion barely kicks in. Assumption A2 rules out such almost risk-neutral protection buyers - relative to the bad endowment.

Finally, the range of fees and default probabilities is restricted as follows.

Assumption A3. For $i \in \{1, 2\}$: $b_i \in [0, \frac{1}{3}]$.

Demanding that dealers' default probabilities are below 30% seems innocuous, and, again, enables us to bound some objects.

Assumption A4.

$$\text{For } i \in \{1, 2\} : \gamma_i \in [0, \gamma^{max}] \text{ with } \gamma^{max} := (-\underline{\theta}) - \frac{2}{\underline{a}}.$$

Assumption A4 requires $-\underline{a}(\underline{\theta} + \gamma_i) > 2$ and, hence, demands that the difference between the fee and the absolute value of the bad endowment ($\gamma_i - (-\underline{\theta})$) is still relevant for risk aversion.[5]

3 Market segmentation

The model analysis starts by investigating Nash equilibria in prices in $t = 1$ for given choices of default probabilities. To fix roles, let $\Delta b := b_2 - b_1 > 0$. That is, protection seller 1 defaults with a lower probability (is the *safer dealer*) or, in other words, offers the product of *higher quality*. Consider two derivatives (b_1, γ_1) and (b_2, γ_2). How does a protection buyer decide between (b_1, γ_1) and (b_2, γ_2)? Key idea is that the degree of risk aversion, a, translates into an "intensity in taste for quality". Heterogeneity among clients in this dimension leads to market segmentation in the intuitive way: more risk-averse clients buy from the safer dealer (Lemma 6).

Let $\vec{b} := (b_1, b_2)$ and $\vec{\gamma} := (\gamma_1, \gamma_2)$ denote the pairs of default probabilities and fees.

Proposition 1. *For given $\vec{\gamma}$ and $\vec{b}$ with $\Delta b > 0$, a protection buyer with degree of risk aversion a is indifferent between the two contracts if*

$$g(a, \vec{\gamma}) := \frac{\exp(-a\Delta\gamma) - 1}{\exp(-a(\underline{\theta} + \gamma_2)) - 1} = \frac{p\Delta b}{1 - b_1 p} =: \tilde{g}(\vec{b}) \tag{3}$$

Proof. See Appendix A1. □

5 For $\gamma^{max} > 0$, $\underline{a}(-\underline{\theta}) > 2$ is needed and indeed ensured by assumption A2.

The following corollary follows as a direct consequence.

Corollary 2. *If $\Delta b = 0$, a consumer can be indifferent only if $\Delta\gamma := \gamma_2 - \gamma_1 = 0$. That is, if the default probabilities of the dealers coincide, pure price competition drives prices to marginal costs (which are set to zero here).*

Two observations follow. Firstly, for any two feasible contracts with $b_2 > b_1$, if there is a solution to (3), then $\Delta\gamma < 0$, i.e. $\gamma_1 > \gamma_2$.[6] In other words, as one would expect, the protection seller that offers the product of higher quality sets the higher price. Secondly, the marginal rate of substitution, i.e. the necessary reduction in the fee, γ, for an increase in default probability, b, in order to keep protection buyer a indifferent, is increasing in the degree of risk aversion.[7] It is intuitive that more risk-averse protection buyers have a larger willingness to pay for an increase in quality.

A derivative contract (b, γ) is called *feasible for a* if protection buyer a prefers the contract to none. This translates into the following condition

$$pu_a(\underline{\theta}) + (1-p)u_a(\overline{\theta}) \leq (1-bp)u_a(-\gamma) + bpu_a(\underline{\theta}) \tag{4}$$

$$\Leftrightarrow bp\left[u_a(-\gamma) - u_a(\underline{\theta})\right] + u_a(\overline{\theta}) - u_a(-\gamma) \leq p\left[u_a(\overline{\theta}) - u_a(-\gamma) + u_a(-\gamma) - u_a(\underline{\theta})\right] \tag{5}$$

$$\Leftrightarrow (1-p)\left[u_a(\overline{\theta}) - u_a(-\gamma)\right] \leq p(1-b)\left[u_a(-\gamma) - u_a(\underline{\theta})\right] \tag{6}$$

(6) admits an intuitive interpretation: Protection buyer a prefers the contract to no insurance, if the expected utility gain from avoiding the bad endowment in case the seller does not default (RHS) outweighs the expected utility loss from the fee if the good endowment materializes (LHS).[8]

The following proposition characterizes the protection buyer that is indifferent between derivative contract (b, γ) and no insurance.

6 To see this, note that with $\Delta b > 0$, the RHS of (3) is positive. The denominator of the LHS of (3) is positive. A positive nominator on the LHS necessitates $\Delta\gamma < 0$.

7 One can verify that $\partial MRS(a)/\partial a = -p/((1-bp)a)\left(\exp(-a(\underline{\theta}+\gamma))\left[1/a + \underline{\theta} + \gamma\right] - 1/a\right)$. To see that this expression is positive, note that for $(\underline{\theta}+\gamma) < -1/a$ it follows directly. For $0 > (\underline{\theta}+\gamma) > -1/a$ it follows, since for all $x \neq 0$ $\exp(x) > 1 + x$.

8 Note that from (6) we also know that for any feasible contract $(\underline{\theta}+\gamma) < 0$. (Since $-\gamma < 0 < \overline{\theta}$, the LHS of (6) is positive, hence, the RHS needs to be positive as well.) Indeed, we already restricted attention to $\gamma < (-\underline{\theta})$ by assumption A4.

Proposition 3. *Protection buyer a is indifferent between (b, γ) and no insurance, if*

$$\gamma = \gamma_a^{exit}(b) := (-\underline{\theta}) - \frac{1}{a} \ln \left(\frac{K(b) + 1}{K(b) + \exp(-a(\overline{\theta} - \underline{\theta}))} \right) \tag{7}$$

with $K(b) = (1 - b)p/(1 - p)$. $\gamma_a^{exit}(b)$ is strictly increasing in a and decreasing in b.

Proof. See Appendix A2. □

The result is intuitive: the fee at which a protection buyer is indifferent between the contract and no insurance is higher the more risk-averse he is. The next corollary follows as a direct consequence.

Corollary 4. *i) For fixed default probability b_i, a derivative contract (b_i, γ_i) is feasible for protection buyer a if $\gamma_i < \gamma_a^{exit}(b_i)$.*

ii) Let $a^{exit}(b_i, \gamma_i)$ be the protection buyer that is indifferent between contract (b_i, γ_i) and no insurance. For γ_i outside of $[\gamma_{\underline{a}}^{exit}(b_i), \gamma_{\overline{a}}^{exit}(b_i)]$, a^{exit} lies outside of the interval $[\underline{a}, \overline{a}]$ and is set to the respective boundary. Then protection buyers with $a < a^{exit}(b_i, \gamma_i)$ prefer no insurance.

iii) If the fee set by the unsafer dealer, γ_2, is smaller than $\gamma_{\underline{a}}^{exit}(b_2)$, $a^{exit} < \underline{a}$ and there is full market coverage.

In the following, the focus is on the case where the market is fully covered. Lifting this restriction does not alter the gist of the analysis, and necessitates case distinctions.[9]

The following main result of this section establishes that there is at most one protection buyer, characterized by some a^*, who is indifferent between the derivatives (b_1, γ_1) and (b_2, γ_2) and splits the market into a segment that buys from protection seller 1 and another segment that prefers protection seller 2.

Proposition 5. *For given choices of default probabilities $\vec{b}$ that satisfy assumption A3 and fees*

9 Later we will introduce $\gamma_2^*(\gamma^{max})$, that is, the best response by ps 2 to the largest possible fee set by ps 1. As shown later in Proposition 10, ps 2's reaction function is increasing. Hence $\gamma_2^*(\gamma^{max})$ is the largest fee possibly set by protection seller in equilibrium, and if $\gamma_2^*(\gamma^{max}) \leq \gamma_{\underline{a}}^{exit}(b_2)$ there is full market coverage anyways. Otherwise, ps 2's reaction function remains unaltered until $\gamma_{\underline{a}}^{exit}(b_2)$. Above that point, ps 2 potentially looses market share "from below" when increasing fees, which may induce him to set fees as best responses. Hence, we expect the reaction function to change above $\gamma_{\underline{a}}^{exit}(b_2)$, but it should leave the core of the analysis unchanged. Note that in the analysis of Appendix C, $\underline{\theta} = 0$ is assumed to avoid precisely this case distinction here and focus on finding a closed-form solution for the case where the low-quality firm indeed adjusts to the threat of loosing customers from below.

$\vec{\gamma}$ with $\gamma_i \in [0, \gamma^{max}]$, there is at most one $a^*(\vec{\gamma})$ satisfying

$$g(a^*(\vec{\gamma}), \vec{\gamma}) = \tilde{g}(\vec{b}) = \frac{p\Delta b}{1 - b_1 p}. \tag{8}$$

Such an $a^*(\vec{\gamma}) \in [\underline{a}, \overline{a}]$ indeed exists, if

$$g(\overline{a}, \vec{\gamma}) \leq \frac{p\Delta b}{1 - b_1 p} \leq g(\underline{a}, \vec{\gamma}). \tag{9}$$

Proof. See Appendix A3. □

The idea of the proof is to show that $g(a, \vec{\gamma})$ is strictly decreasing in a, while the RHS of (8) is fixed. Hence, there can be at most one solution, and (9) indeed guarantees existence of a unique indifferent client.[10] [11]

The next lemma on market segmentation follows as a direct consequence.

Lemma 6. *Suppose there exists an indifferent protection buyer, i.e. $a^*(\vec{\gamma}) \in [\underline{a}, \overline{a}]$ s.t. (8) holds. Then protection buyer a will choose protection seller 1 iff*

$$a \geq a^*(\vec{\gamma}). \tag{10}$$

In other words, protection seller 1 with $b_1 < b_2$ will face the protection buyers in the market segment $[a^*(\vec{\gamma}), \overline{a}]$, while protection seller 2 receives the market share $[\underline{a}, a^*(\vec{\gamma})]$.

Proof. See Appendix A4. □

Since a higher a corresponds to a higher degree of risk aversion, the clients of the safer dealer ($b_1 < b_2$) are the more risk-averse ones. The resulting market segmentation is depicted in Figure 2.

10 Proposition 5 does *not* claim that by enlarging $[\underline{a}, \overline{a}]$ one can necessarily achieve (9). To the contrary, it may well be that for a given pair of default probabilities, no choice of fees can achieve this - which would imply that one protection seller will "own" the whole market. If $g(\underline{a}, \vec{\gamma}) \leq \tilde{g}(\vec{b})$ dealer 1 "owns" the entire market, if $g(\overline{a}, \vec{\gamma}) \geq \tilde{g}(\vec{b})$ dealer 2 "owns" the entire market.

11 Note that in Proposition 5 not only the vector of default probabilities, but also the vector of fees was assumed to be fixed. For varying fees, a^* may assume any positive real value (fix γ_2 and let $\Delta\gamma$ go from zero to minus infinity). In the sequel, the possible range of values for a^* will be considered in different settings: For a given vector of fees, we require $a^* \in [\underline{a}, \overline{a}]$ for an interpretation of an indifferent consumer. But, as long as $\lim_{a \to 0} g(a, \vec{\gamma}) = \Delta\gamma/(\underline{\theta} + \gamma_2) \geq \tilde{g}(\vec{b})$, for some calculations using γ as a variable, a^* as defined by (8) is well-defined even if outside the admissible range $[\underline{a}, \overline{a}]$.

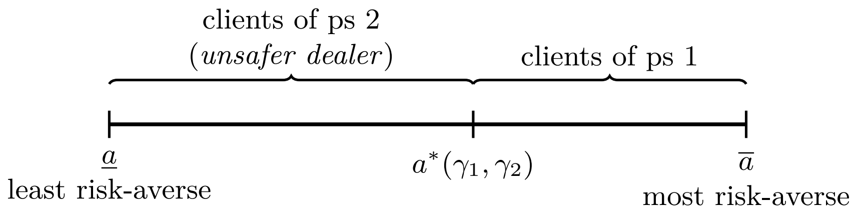

Figure 2: Market shares for protection seller 1 and 2 with $b_2 > b_1$

4 Existence of an equilibrium in prices

4.1 Continuous reaction function in prices of the "unsafer" dealer

The main result of this section (Proposition 10) establishes existence of a continuous and strictly increasing reaction function of protection seller 2. Since there is no closed-form solution of the indifferent client, this endeavor is a bit more involved. As before, fix a vector of default probabilities $\vec{b}$ with $\Delta b > 0$ that satisfies assumption A3.

We start by deriving some helpful properties of the indifferent consumer that also enable us to visualize the set-up. The following first lemma offers a second characterization of the indifferent consumer, symmetric to the one derived in Proposition 1. Exploiting this symmetry will be key in the sequel.

Lemma 7. *The protection buyer, a, that is indifferent between two derivatives (b_1, γ_1) and (b_2, γ_2) has a second characterization, symmetric to (3), namely*

$$h(a, \vec{\gamma}) := \frac{1 - \exp(-(-a\Delta\gamma))}{\exp(-a(\underline{\theta} + \gamma_1)) - 1} = \frac{p\Delta b}{1 - b_2 p}. \tag{11}$$

Proof. See Appendix A5. □

Some notation. To be able to succinctly state the following results, we introduce some notation. Define

$$\tilde{A} : [\underline{a}, \overline{a}] \times [0, -\underline{\theta})^2 \to \mathbb{R}, \quad (a, \vec{\gamma}) \mapsto \exp(-a\Delta\gamma) \tag{12}$$

$$\text{and} \quad \tilde{B}_i : [\underline{a}, \overline{a}] \times [0, -\underline{\theta}) \to \mathbb{R}, \quad (a, \gamma_i) \mapsto \exp(-a(\underline{\theta} + \gamma_i)). \tag{13}$$

Let

$$A(\vec{\gamma}) := \tilde{A}(a^*(\vec{\gamma}), \vec{\gamma}), \quad \text{and} \quad B_i(\vec{\gamma}) := \tilde{B}_i(a^*(\vec{\gamma}), \gamma_i) \tag{14}$$

be the two functions defined on $[0, -\underline{\theta})^2$ one obtains when inserting the indifferent consumer $a^*(\vec{\gamma})$ into (12) and (13). Whenever clear from the context, the explicit dependence on $\vec{\gamma}$ is omitted. The indifferent consumer $a^*(\vec{\gamma})$ has been characterized implicitly via (3) and (11). From the RHSs of (3) and (11) we infer that the respective LHSs, i.e.

$$g(a^*(\vec{\gamma}), \vec{\gamma}) = \frac{A(\vec{\gamma}) - 1}{B_2(\vec{\gamma}) - 1} \quad \text{and} \quad h(a^*(\vec{\gamma}), \vec{\gamma}) = \frac{1 - \frac{1}{A(\vec{\gamma})}}{B_1(\vec{\gamma}) - 1} \tag{15}$$

are constants, and call them g and h respectively. With the following last definitions,

$$\varphi_1 := (\underline{\theta} + \gamma_2) B_1 \quad \text{and} \quad \varphi_2 := (\underline{\theta} + \gamma_1) B_2, \tag{16}$$

one can state the auxiliary lemma.

Lemma 8. *With the notation introduced above, for firm 1 it holds that*

$$\partial_1 a^* > 0 \tag{f1-i}$$

$$\partial_1 a^* = \frac{a^*}{(\Delta\gamma - g\varphi_1)} \quad \textit{with} \quad (\Delta\gamma - g\varphi_1) > 0, \tag{f1-ii}$$

and analogously for firm 2

$$\partial_2 a^* < 0 \tag{f2-i}$$

$$\partial_2 a^* = \frac{-a^*}{(\Delta\gamma - h\varphi_2)} \quad \textit{with} \quad (\Delta\gamma - h\varphi_2) > 0. \tag{f2-ii}$$

Proof. See Appendix A6. □

The result is intuitive in light of the profits, Π_1 and Π_2, of the high- and low-quality dealer, respectively:

$$\Pi_1(\gamma_1, \gamma_2) = (\overline{a} - a^*(\gamma_1, \gamma_2))\, \gamma_1 \tag{17}$$

$$\Pi_2(\gamma_1, \gamma_2) = (a^*(\gamma_1, \gamma_2) - \underline{a})\, \gamma_2. \tag{18}$$

(f1-i) and (f2-i) state that both firms loose market share when increasing fees.

Figure 3 visualizes the set up with the admissible fees $[0, \gamma^{max}]$ of dealer 1 and 2 on the x- and y-axis respectively. With dealer 1 the safer dealer, fees lie below the diagonal. The green line

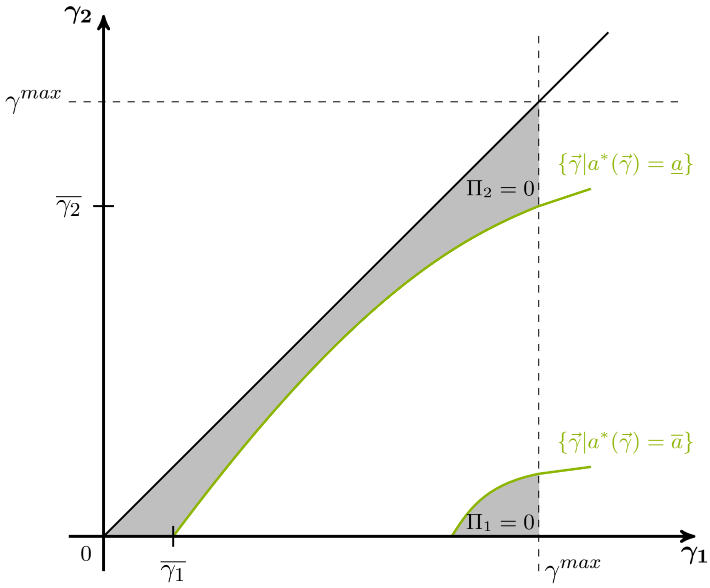

Figure 3: Illustration of the admissible fees and zero-profit regions

just below the diagonal depicts the pairs of fees for which the indifferent consumer a^* takes value $\underline{a}$. From Lemma 8 we know that contour lines of a^* qualitatively take this shape. Above this line the unsafer dealer has no market share and subsequently no profits. Denote by $\overline{\gamma_1}$ and $\overline{\gamma_2}$ the intercepts of this line with the x- and y-axis respectively.

We parameterize the pairs of fees for which the indifferent consumer takes the value $\underline{a}$, in terms of γ_1 as well as in terms of γ_2, namely

$$\{\vec{\gamma}|a^*(\vec{\gamma}) = \underline{a}\} = \{(\gamma_1^{\underline{a}}(\gamma_2), \gamma_2)|\gamma_2 \in [0, \overline{\gamma_2}]\} = \{(\gamma_1, \gamma_2^{\underline{a}}(\gamma_1))|\gamma_1 \in [\overline{\gamma_1}, \gamma^{max}]\}. \tag{19}$$

Ensuring that the set-up is interesting, i.e. that ps 1 does not a priori get the entire market, requires $\overline{\gamma_2} > 0$, which is exactly assumption A2. The following lemma formalizes the above.

Lemma 9. *i) For $\gamma_2 \in [0, \gamma^{max}]$ define*

$$\gamma_1^{\underline{a}}(\gamma_2) \textit{ such that } a^*(\gamma_1^{\underline{a}}(\gamma_2), \gamma_2) = \underline{a} \tag{20}$$

$$\gamma_1^{\overline{a}}(\gamma_2) \textit{ such that } a^*(\gamma_1^{\overline{a}}(\gamma_2), \gamma_2) = \overline{a}. \tag{21}$$

Then $\gamma_1^{\underline{a}} < \gamma_1^{\overline{a}}$ *and*

$$\gamma_1^{\underline{a}} \leq \gamma^{max} \quad \Leftrightarrow \quad \gamma_2 \leq \overline{\gamma_2} \tag{22}$$

$$\textit{with } \overline{\gamma_2} := \arg_\gamma\{a^*(\gamma^{max}, \gamma) = \underline{a}\} = (-\underline{\theta}) - \frac{1}{\underline{a}} \log\left[\frac{1 - \tilde{g}(\vec{b})}{\exp(-2) - \tilde{g}(\vec{b})}\right]. \tag{23}$$

ii) Analogously, for $\gamma_1 \in [0, \gamma^{max}]$ *define*

$$\gamma_2^{\underline{a}}(\gamma_1) \textit{ such that } a^*(\gamma_1, \gamma_2^{\underline{a}}(\gamma_1)) = \underline{a} \tag{24}$$

$$\gamma_2^{\overline{a}}(\gamma_1) \textit{ such that } a^*(\gamma_1, \gamma_2^{\overline{a}}(\gamma_1)) = \overline{a} \tag{25}$$

Then $\gamma_2^{\underline{a}} > \gamma_2^{\overline{a}}$ *and*

$$\gamma_2^{\underline{a}} \geq 0 \quad \Leftrightarrow \quad \gamma_1 \geq \overline{\gamma_1} \tag{26}$$

$$\textit{with } \overline{\gamma_1} := \arg_\gamma\{a^*(\gamma, 0) = \underline{a}\} = \frac{1}{\underline{a}} \log\left[1 + \tilde{g}(\vec{b})\left(\exp(\underline{a}(-\underline{\theta})) - 1\right)\right]. \tag{27}$$

iii) As one would expect from the picture $\gamma_2 \leq \overline{\gamma_2}$ *iff* $\gamma_1 \geq \overline{\gamma_1}$.

iv) Protection seller 1 gets the entire market if

$$\overline{\gamma_2} \leq 0 \quad \Leftrightarrow \quad \underline{a}(-\underline{\theta}) \leq \log\left[\frac{1 - \tilde{g}(\vec{b})}{\exp(-2) - \tilde{g}(\vec{b})}\right]. \tag{28}$$

With $\tilde{g}(\vec{b}) < 1/8$ *from assumption A1 and A3,*

$$\underline{a}(-\underline{\theta}) > \log\left[\frac{1 - \frac{1}{8}}{\exp(-2) - \frac{1}{8}}\right] \approx 4.4, \tag{29}$$

ensures that the set-up is interesting. This is exactly assumption A2.

Proof. See Appendix A7. □

The next proposition is the main result of this section and establishes existence of a differentiable and strictly increasing reaction function for protection seller 2.

Proposition 10. *Consider all pairs of fees* $(\gamma_1, \gamma_2) \in [0, \gamma^{max}]^2$ *with* $\Delta\gamma < 0$. *Then, i) for any* $\gamma_1 \in [0, \gamma^{max}]$, *there is a unique best response in fees for firm 2,* $\gamma_2^*(\gamma_1)$. *ii) This reaction* *function γ_2^* is continuously differentiable and strictly increasing in γ_1.*

Proof. See Appendix A8. □

The idea of the existence of firm 2's best response function is as follows: Fix some $\gamma_1 \in [0, \gamma^{max}]$. At an interior best response γ_2 we must have

$$d_2\Pi_2 = 0 \quad \Leftrightarrow \quad -\partial_2 a^* = \frac{(a^* - \underline{a})}{\gamma_2}. \tag{30}$$

One shows that the RHS of the equivalence (30) is strictly decreasing, while the LHS is strictly increasing. Hence, there can be at most one solution. Since Π_2 is a continuous function on a compact interval, a maximum exists, and since $\Pi_2(\gamma_1, 0) = \Pi_2(\gamma_1, \gamma^{max}) = 0$, it indeed lies in the interior. Figure 4 shows the qualitative shape.

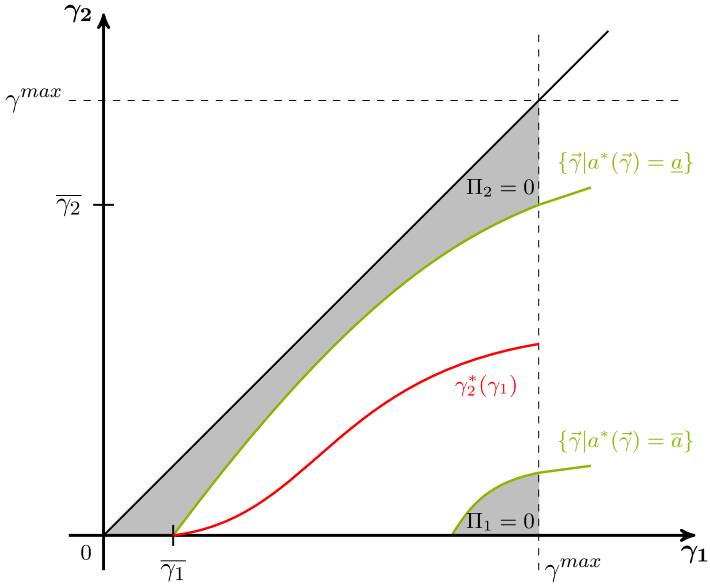

Figure 4: Protection seller 2's best response function

4.2 Existence of a Nash equilibrium in prices

The main results of this section (Proposition 12 and 13) show existence of a Nash equilibrium in prices, except in a non-generic case specified below. From now on we impose the following additional assumption.

Assumption A5. $d_1\Pi_1(\gamma^{max}, \overline{\gamma_2}) \geq 0$.

Lemma 11. *From assumption A5 it follows that $d_1\Pi_1(\gamma_1^{\underline{a}}(\gamma_2), \gamma_2) > 0$ for all $\gamma_2 \in [0, \overline{\gamma_2}]$.*

Proof. See Appendix A10. □

The consequence of assumption A5, captured in Lemma 11, has an intuitive interpretation. It demands that when protection seller 1 owns the entire market, there is no incentive for him to decrease prices. A reduction in prices entails lower prices on the existing market share, while increasing the market share. Since protection seller 1 already owns the entire market along $\gamma_1^{\overline{a}}(\cdot)$, there is no market share effect and, subsequently, decreasing prices should be unattractive. This natural condition is ensured for all price pairs at which ps 1 owns the entire market (i.e. along $\gamma_1^{\overline{a}}(\cdot)$), if it weakly holds just at the point where ps 1 is unable to increase prices any further, that is at $\gamma_1 = \gamma^{max}$. This is assumption A5.

Proposition 12 shows existence of a point on the reaction function of ps 2 that either satisfies the FOC of the profit maximization problem of ps 1 or is a boundary solution, i.e. lies on the boundary with a gradient pointing outside of the set.

Proposition 12. *There is a point $\vec{\gamma} \in [0, \gamma^{max}]^2$ on firm 2's reaction function that satisfies* *either*

$$d_1\Pi_1(\vec{\gamma}) = 0 = d_2\Pi_2(\vec{\gamma}) \tag{31}$$

or that lies on the boundary and fulfills the respective "local boundary condition", i.e.

$$d_2\Pi_2(\vec{\gamma}) = 0 < d_1\Pi_1(\vec{\gamma}) \quad \textit{if } \vec{\gamma} = (\gamma^{max}, \gamma_2^{\underline{a}}(\gamma^{max})) \tag{32}$$

Proof. See Appendix A9. □

Idea of the proof is to characterize points that satisfy either (31) or (32) as fixed points of a continuous mapping from $[0, \gamma^{max}]$ onto itself, and apply Brouwer's fixed point theorem.

Unfortunately, Proposition 12 does not prove existence of a Nash equilibrium for the following reason. With Lemma 10 we know that the FOC for firm 2 indeed describe best responses. For firm 1, however, the FOC *a priori* need not characterize local maxima, since firm 1's profit function need not be concave and some simple examples suggest it likely is not.[12] Alternatively,

12 Π_1 is strictly concave in γ_1 at the point (γ_1, γ_2) if $-a^*(\underline{\theta} + \gamma_2) < 2 \underbrace{(\Delta\gamma - g\varphi_1)/(-g\varphi_1)}_{\in(0,1)} \cdot \underbrace{1/(1 - gB_1)}_{>1}$. In light

one can demand that the second derivative of firm 1's profit function can switch sign only once. The following lemmata state such conditions and show that one can ensure existence of a Nash equilibrium.

Proposition 13. *Suppose the following condition holds.*

(T) For all $\gamma_2 \in [0, \gamma_2^*(\gamma^{max})]$ *such that* (γ_1, γ_2) *satisfies* (31) *or* (32) *for some* γ_1*, there exists* $\mu \in [0, \gamma^{max}]$ *such that for all* $\gamma \in [\gamma_1^{\underline{a}}(\gamma_2), \gamma^{max}]$

$$\text{for } \gamma < \mu \text{ local extrema are local minima, i.e. } d_1\Pi_1|_{(\gamma,\gamma_2)} = 0 \Rightarrow d_1^2\Pi_1|_{(\gamma,\gamma_2)} > 0,$$

$$\text{for } \gamma > \mu \text{ local extrema are local maxima, i.e. } d_1\Pi_1|_{(\gamma,\gamma_2)} = 0 \Rightarrow d_1^2\Pi_1|_{(\gamma,\gamma_2)} < 0.$$

Then (γ_1, γ_2) *satisfying* (31) *or* (32) *is* <u>*either*</u> *a saddle point, i.e.* $d_1\Pi_1(\gamma_1, \gamma_2) = 0 = d_1^2\Pi_2(\gamma_1, \gamma_2)$ *and* $d_1\Pi_1$ *and does not switch signs at* (γ_1, γ_2)*,* <u>*or*</u> *maximizes firm 1's profit function over the entire interval and hence is a Nash equilibrium.*

Proof. See Appendix A11. □

The idea of the proof is to show that condition (T) together with assumption A5 rules out local minima with FOC equal 0, while it also ensures that any local maximum is a global maximum. Hence, as long as the fixed point from Proposition 12 does not describe a saddle point of Π_1, it describes a best response on the entire set of feasible actions for both firms, hence is a Nash equilibrium.

Remark 1. It is a priori unsatisfactory that one cannot exclude the possibility that a fixed point from Proposition 12 is a saddle point for the profit function of ps 1. This case, however, is a non-generic one. There can only be finitely many saddle points of protection seller 1's profit function (shown in Appendix A12). By changing $\underline{a}$ a little bit, one can shift protection seller 2's reaction function slightly and can thus ensure that none of the saddle points lie on ps 2's reaction function. Changing the degree of risk aversion of the least risk-averse client, i.e. $\underline{a}$, a tiny bit does not alter the essence of the economic set-up and and one can thereby ensure existence of a Nash equilibrium.

of assumption A4 this cannot be ensured in general.

The following lemma clarifies when condition (T) holds.

Lemma 14. *Condition (T) holds if one of the following is satisfied*

$$a^* \geq \frac{2}{3}\overline{a} \tag{E2i}$$

$$or \quad -\frac{\partial_2 a^*}{\partial_1 a^*} < \frac{1}{2} \ \ and \ \ -a^*(\underline{\theta} + \gamma_2) > 4 \tag{E2ii}$$

$$or \quad 2\Delta\gamma - g\varphi_1 \leq 0 \tag{E2iii}$$

for all $\gamma_2 \in [0, \gamma_2^*(\gamma^{max})]$ *such that* (γ_1, γ_2) *satisfies* (31) *or* (32) *for some* γ_1, *and for all* $\gamma_1 \in [\gamma_1^{\underline{a}}(\gamma_2), \gamma^{max}]$.

Proof. See Appendix A13. □

One can ensure (E2i), for example, by restricting parameters further such that

$$\underline{a} \geq \frac{2}{3}\overline{a}. \tag{33}$$

5 Market discipline in quality choices

Until now, we have carried over key insights from the standard model of vertical product differentiation to a set-up where the good sold is insurance and clients' "taste for quality" is risk aversion. This section shows how the fact that the quality-leader enjoys larger profits in equilibrium, can produce market discipline to keep one's own default probabilities low. Such endogenous market discipline has not been the focus of the literature on vertical product differentiation and is, in fact, in disconnect with the standard model in Tirole (1988) on the topic. Thus, in Appendix C I revisit the standard model of vertical product differentiation, show a refined principle of product differentiation under less restrictive assumptions and demonstrate that upward pressure on qualities can arise already in the standard set-up.

This section starts by showing that no coordination issue can arise.

5.1 No coordination issue

Suppose assumption A5 and one of the conditions outlined in Lemma 14 holds. The following lemma shows that, for pre-assigned roles of quality-leader and quality-follower, a coordination issue stemming from multiple equilibria does not arise, since both protection sellers prefer the price equilibrium with higher prices.

Lemma 15. *Suppose* $\vec{\gamma}^1 = (\gamma_1^1, \gamma_2^2)$ *and* $\vec{\gamma}^2 = (\gamma_1^2, \gamma_2^2)$ *are both Nash equilibria with* $\gamma_1^1 < \gamma_1^2$*. Then both protection sellers prefer* $\vec{\gamma}^2$*. The Nash equilibrium with largest prices is called the "preferred Nash equilibrium".*

Proof. See Appendix A14. □

5.2 Upward pressure on qualities

Importantly, in any Nash equilibrium the quality-leader enjoys greater profits, as shown in the following lemma.

Lemma 16. *At any Nash equilibrium in prices* $\vec{\gamma}$*,*

i) the quality-leader has greater profits, i.e. $\Pi_1 > \Pi_2$*,*

ii) the quality-leader has the greater market share, i.e. $(\overline{a} - a^*) > (a^* - \underline{a})$*.*

Proof. See Appendix A15. □

Until now the vector of default probabilities was taken as given and were concerned with the resulting Nash equilibrium in prices. Now we let the default probabilities vary.

Lemma 17. *Consider quality choices* $\vec{b}^0 = (b_1^0, b_2^0) \in [0, 1/3]^2$ *with* $b_2^0 > b_1^0$ *and* $\tilde{g}(\vec{b}^0) = p(b_2^0 - b_1^0)/(1 - b_1^0 p)$*. Let* $\vec{\gamma}$ *be the corresponding preferred Nash equilibrium in prices. Then* $\vec{\gamma}$ *is also the preferred Nash equilibrium in prices for all* $(b_1, b_2) \in [0, 1/3]^2$ *with* $\tilde{g}(\vec{b}) = \tilde{g}(\vec{b}^0)$ *which is equivalent to* $b_2 = (1 - \tilde{g}(\vec{b}^0))b_1 + \tilde{g}(\vec{b}^0)/p$*.*

Proof. See Appendix A16. □

Lemma 17 shows that pairs of default probabilities that lead to the same price equilibrium lie on straight lines. This is visualized by Figure 5 with the admissible quality choices of dealer 1 and 2 on the x- and y-axis respectively. With dealer 1 the safer dealer, default probabilities lie above the diagonal. For (b_1^0, b_2^0) the blue line depicts all pairs of default probabilities that lead to the same value of $\tilde{g}$ and subsequently the same price equilibrium.

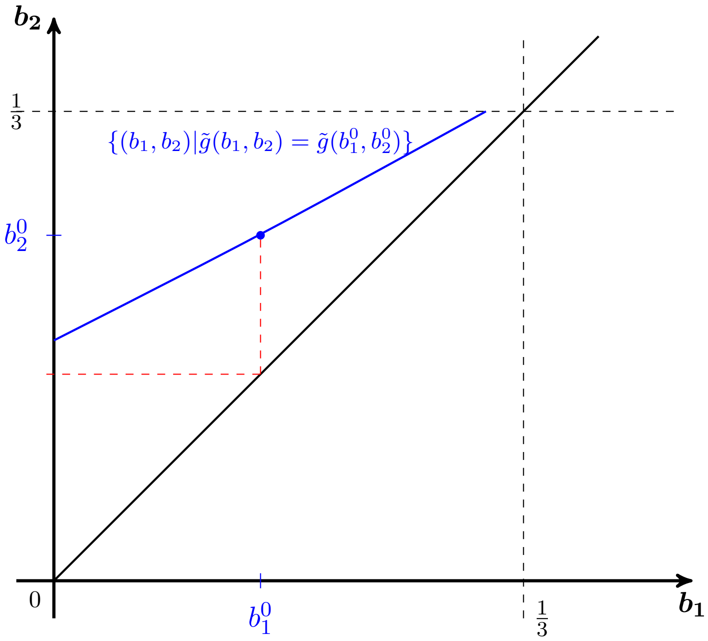

Figure 5: Qualities leading to the same price equilibrium lie on straight lines (blue line)

To demonstrate that there is upward pressure on qualities, we break the symmetry between the two firms not by assigning the role of quality-leader beforehand, but by making the quality choice sequential (*sequential game without assigned roles*). Namely, suppose that there is an additional time period $t = (-1)$ in which the first-mover, say protection seller 1, chooses its default probability, while protection seller 2 continues to choose its default probability in $t = 0$. The rest remains unaltered.

Proposition 18. *A necessary condition for some* (b_1, b_2) *to be a subgame-perfect Nash equilibrium in the sequential game without assigned roles, is that*

$$b_1 < \frac{8}{15} \cdot \frac{1}{3}. \tag{34}$$

Proof. See Appendix A17 for details. □

The intuition of the result is very similar to that of Proposition 24 in Appendix C. Suppose

$b_1 < b_2$ is a Nash equilibrium in the sequential game without assigned roles. Then is must not be profitable for the second-mover (protection seller 2) to take over the lead position in qualities. Lemma 17 characterizes the quality choices leading to the same price equilibrium as straight lines. Thus the natural candidate for a profitable deviation of the second-mover is to choose a quality pair which leads to the same price equilibrium, but with reversed roles. The above condition rules out this profitable deviation. In Figure 5 no such profitable deviation is possible, since the red dotted line and the blue line do not intersect.

As a consequence from the proposition, the default probability b_1 has to be almost in the lower half of the allowed range $[0, 1/3]$. In that sense the incentive to keep the leadership position quality-wise, produces the incentive to keep the own default probability low.

6 Conclusion

This paper models imperfect competition between dealers in derivatives markets. Key features of derivatives markets that such a model needs to accommodate are, firstly, the nature of competition between few dealers with market power, and, secondly, the two-tiered structure of the market with dealers at the core and clients at the periphery.

The way competition is modeled here resembles models from IO of vertical product differentiation: two dealers that choose their own default probability (i.e. quality) sell insurance against a common macro risk at a fee (i.e. price). The two-tiered market structure is embedded in the model since the two dealers sell derivatives to a continuum of clients that differ in their risk aversion. For two dealers with given, but differing, default probabilities I show existence of a price equilibrium that is preferred by both dealers. Since in this equilibrium the dealer with the lower default probability makes larger profits, this can produce market discipline among dealers by giving an incentive to keep the own default probability low.

This set-up is suitable for incorporating a CCP with client clearing: The two dealers become members of the CCP, while clients access clearing services via the dealers. Currently a CCP is added to the model only in an ad-hoc fashion as an innovation that removes the quality dimension of the competition between the dealers. I argue that a removal of the quality dimension is a likely effect of the introduction of CCPs, but, as the discussion of what a CCP does at its core is still ongoing, there are many more features of central clearing one could model.

The current approach demonstrates that it is fruitful to borrow from models of product differentiation in order to capture competition in derivatives markets. The economic forces at play can be carried over to the present (quite general) set-up with risk aversion that does not admit a closed-form solution. Modeling the CCP in more detail (e.g. with loss sharing mechanisms, margins, default probability of the CCP, etc) while keeping competition between dealers and the market structure in place is left for future research.

References

Abad, J., Aldasoro, I., Aymanns, C., D'Errico, M., Rousová, L. F., Hoffmann, P., Langfield, S., Neychev, M., and Roukny, T. (2016). Shedding light on dark markets: First insights from the new eu-wide otc derivatives dataset. *ESRB: Occasional Paper Series*, (2016/11).

Acharya, V. and Bisin, A. (2014). Counterparty risk externality: Centralized versus over-the-counter markets. *Journal of Economic Theory*, 149:153–182.

Belleflamme, P. and Peitz, M. (2015). *Industrial organization: markets and strategies.* Cambridge University Press, 2nd edition.

Biais, B., Heider, F., and Hoerova, M. (2012). Clearing, counterparty risk, and aggregate risk. *IMF Economic Review*, 60(2):193–222.

Biais, B., Heider, F., and Hoerova, M. (2016). Risk-sharing or risk-taking? Counterparty risk, incentives, and margins. *The Journal of Finance*, 71(4):1669–1698.

Braithwaite, J. P. and Murphy, D. (2020). Take on me: OTC derivatives client clearing in the european union. *LSE Legal Studies Working Paper.*

Capponi, A. and Cheng, W. A. (2018). Clearinghouse margin requirements. *Operations Research*, 66(6):1542–1558.

Carapella, F. and Monnet, C. (2020). Dealers' insurance, market structure, and liquidity. *Journal of Financial Economics*, 138(3):725–753.

Cucic, D. (forthcoming, 2022). Central clearing and loss allocation rules. *Journal of Financial Markets.*

Duffie, D., Gârleanu, N., and Pedersen, L. H. (2005). Over-the-counter markets. *Econometrica*, 73(6):1815–1847.

Duffie, D. and Zhu, H. (2011). Does a central clearing counterparty reduce counterparty risk? *The Review of Asset Pricing Studies*, 1(1):74–95.

Financial Stability Board (2018). Incentives to centrally clear over-the-counter (OTC) derivatives. *FSB and standard-setting bodies report.* https://www.bis.org/publ/othp29.pdf.

Gabszewicz, J. J. and Thisse, J.-F. (1979). Price competition, quality and income disparities. *Journal of Economic Theory*, 20(3):340–359.

Gabszewicz, J. J. and Thisse, J.-F. (1980). Entry (and exit) in a differentiated industry. *Journal of Economic Theory*, 22(2):327–338.

Gregory, J. (2014). *Central counterparties: mandatory central clearing and initial margin requirements for OTC derivatives*. John Wiley & Sons.

Group of Thirty Working Group on Treasury Market Liquidity (2021). U.S. treasury markets: Steps toward increased resilience. `https://group30.org/publications/detail/4950`.

Holmstrom, B. and Tirole, J. (1997). Financial intermediation, loanable funds, and the real sector. *The Quarterly Journal of Economics*, 112(3):663–691.

Huang, W. (2019). Central counterparty capitalization and misaligned incentives. *BIS Working Paper*.

Huang, W. and Zhu, H. (2021). CCP auction design. *BIS Working Paper*.

Kubitza, C., Pelizzon, L., and Getmansky Sherman, M. (2021). Loss sharing in central clearinghouses: Winners and losers. *SAFE Working Paper*.

Kuong, J. and Maurin, V. (2020). The design of a central counterparty. *Working Paper*.

Mas-Colell, A., Whinston, M. D., and Green, J. R. (1995). *Microeconomic theory*. Oxford university press New York, 1st edition.

Moorthy, K. S. (1988). Product and price competition in a duopoly. *Marketing Science*, 7(2):141–168.

Moorthy, K. S. (1991). Erratum to: Product and price competition in a duopoly. *Marketing Science*, 10(3):270.

Shaked, A. and Sutton, J. (1982). Relaxing price competition through product differentiation. *The Review of Economic Studies*, 49(1):3–13.

Tirole, J. (1988). *The theory of industrial organization*. MIT press, 1st edition.

Wang, J. J., Capponi, A., and Zhang, H. (2022). A theory of collateral requirements for central counterparties. *Management Science*.

A Appendix: Proofs

A1 Proof of Proposition 1

For the indifferent protection buyer we have

$$
\begin{aligned}
U_a(b_1, \gamma_1) &= U_a(b_2, \gamma_2) && \text{(A3)}\\
\Leftrightarrow (1 - b_1 p) u_a(-\gamma_1) + b_1 p u_a(\underline{\theta}) &= (1 - b_2 p) u_a(-\gamma_2) + b_2 p u_a(\underline{\theta}) && \text{(A4)}\\
\Leftrightarrow u_a(-\gamma_1) - u_a(-\gamma_2) + p\left[b_2 u_a(-\gamma_2) - b_1 u_a(-\gamma_1)\right] &= p \Delta b u_a(\underline{\theta}) && \text{(A5)}\\
\Leftrightarrow \left[u_a(-\gamma_1) - u_a(-\gamma_2)\right](1 - b_1 p) &= p \Delta b \left[u_a(\underline{\theta}) - u_a(-\gamma_2)\right] && \text{(A6)}\\
\Leftrightarrow \frac{u_a(-\gamma_1) - u_a(-\gamma_2)}{u_a(\underline{\theta}) - u_a(-\gamma_2)} &= \frac{p \Delta b}{1 - b_1 p} && \text{(A7)}\\
\Leftrightarrow \frac{\exp(-a\Delta\gamma) - 1}{\exp(-a(\underline{\theta} + \gamma_2)) - 1} &= \frac{p \Delta b}{1 - b_1 p}. && \text{(A8)}
\end{aligned}
$$

□

A2 Proof of Proposition 3

In light of (6), a protection buyer a is indifferent between buying contract (b, γ) and no insurance, if

$$
\begin{aligned}
\frac{u_a(\overline{\theta}) - u_a(-\gamma)}{u_a(-\gamma) - u_a(\underline{\theta})} &= \frac{p}{1-p}(1 - b) && \text{(A9)}\\
\Leftrightarrow \frac{\exp(-a\overline{\theta}) - \exp(a\gamma)}{\exp(a\gamma) - \exp(-a\underline{\theta})} &= K(b) && \text{(A10)}\\
\Leftrightarrow \frac{\exp(-a(\overline{\theta} + \gamma)) - 1}{1 - \exp(-a(\underline{\theta} + \gamma))} &= K(b) && \text{(A11)}\\
\Leftrightarrow \frac{\exp(-a\Delta\theta)\exp(-a(\underline{\theta} + \gamma)) - 1}{1 - \exp(-a(\underline{\theta} + \gamma))} &= K(b) && \text{(A12)}\\
\Leftrightarrow \exp(-a(\underline{\theta} + \gamma)) &= \frac{K(b) + 1}{K(b) + \exp(-a\Delta\theta)} && \text{(A13)}\\
\Leftrightarrow \gamma = \gamma_a^{exit}(b) &:= (-\underline{\theta}) - \frac{1}{a}\ln\left(\frac{K(b) + 1}{K(b) + \exp(-a\Delta\theta)}\right), && \text{(A14)}
\end{aligned}
$$

with $K(b) := (1 - b)p/(1 - p)$ and $\Delta\theta := (\overline{\theta} - \underline{\theta})$.

ad $\gamma_a^{exit}(b)$ increasing in a. We have

$$
\frac{\partial \gamma_a^{exit}}{\partial a} = \frac{1}{a}\left[\frac{1}{a}\log\left(\frac{K(b) + 1}{K(b) + \exp(-a\Delta\theta)}\right) - \frac{\exp(-a\Delta\theta)}{K(b) + \exp(-a\Delta\theta)}\Delta\theta\right]. \tag{A15}
$$

With

$$
y := \frac{1 - \exp(-a\Delta\theta)}{K(b) + \exp(-a\Delta\theta)} \tag{A16}
$$

this reads

$$
\frac{\partial \gamma_a^{exit}}{\partial a} = \frac{1}{a}\left[\frac{1}{a}\ln(1 + y) + \left(y - \frac{1}{K(b) + \exp(-a\Delta\theta)}\right)\Delta\theta\right] \tag{A17}
$$

$$= \frac{1}{a}\left[\frac{1}{a}y\left(\frac{\log(1+y)}{y}+a\Delta\theta\right)-\frac{1}{K(b)+\exp(-a\Delta\theta)}\Delta\theta\right] \tag{A18}$$

$$= \left(\frac{1}{a}\right)^2\frac{1}{K(b)+\exp(-a\Delta\theta)}\left[(1-\exp(-a\Delta\theta))\left(\frac{\log(1+y)}{y}+a\Delta\theta\right)-a\Delta\theta\right] \tag{A19}$$

$$= \left(\frac{1}{a}\right)^2\frac{1}{K(b)+\exp(-a\Delta\theta)}\left[(1-\exp(-a\Delta\theta))\frac{\log(1+y)}{y}-\exp(-a\Delta\theta)a\Delta\theta\right] \tag{A20}$$

With $x := a\Delta\theta$ this expression is positive if and only if

$$\frac{\exp(x)-1}{x} > \frac{y}{\log(1+y)} \tag{A21}$$

$$\Leftrightarrow \log(1+y) > y\frac{x}{\exp(x)-1} \tag{A22}$$

$$\Leftrightarrow \log\left(\frac{K(b)+1}{K(b)+\exp(-x)}\right) > \frac{x}{\exp(x)}\frac{1}{K(b)+exp(-x)}. \tag{A23}$$

For $x = 0$ the LHS and RHS are 0. For $x > 0$ the derivative w.r.t. x of the LHS reads

$$\frac{\partial LHS}{\partial x} = \frac{\exp(-x)}{K(b)+\exp(-x))}, \tag{A24}$$

while the derivative of the RHS reads

$$\frac{\partial RHS}{\partial x} = \frac{\exp(-x)}{K(b)+\exp(-x)}\underbrace{\left[(1-x)+\frac{1}{K(b)+\exp(-x)}\frac{x}{\exp(x)}\right]}_{<1}. \tag{A25}$$

To see why the expression in brackets is smaller one, note that

$$(1-x)+\frac{1}{K(b)+\exp(-x)}\frac{x}{\exp(x)} < 1 \Leftrightarrow \frac{1}{K(b)+\exp(-x)} < \exp(x) \Leftrightarrow 0 < K(b)\exp(x),$$

which always holds and proves the claim.

ad $\gamma_a^{exit}(b)$ increasing in b. Follows directly, since

$$\frac{\partial\gamma_{\underline{a}}^{exit}}{\partial K(b)} = \frac{1-\exp(-a\Delta\theta)}{(1+K(b))(K(b)+\exp(-a\Delta\theta))} > 0 \tag{A26}$$

$$\text{and } \partial_b K(b) < 0. \tag{A27}$$

□

A3 Proof of Proposition 5

ad i) The proof proceeds by showing that $\partial_a g < 0$. Suppose this was true. Then the LHS of (3) is monotonically decreasing, while the RHS of (3) is fixed, yielding at most one solution.

Claim. $\partial_a g < 0$.

Proof of claim. For the derivative of the function g with respect to a we get

$$\frac{\partial g(a)}{\partial a} = \frac{-\Delta\gamma\exp(-a\Delta\gamma)\left(\exp(-a(\underline{\theta}+\gamma_2))-1\right)}{\left(\exp(-a(\underline{\theta}+\gamma_2))-1\right)^2} + \frac{\left(\exp(-a\Delta\gamma)-1\right)\left(\underline{\theta}+\gamma_2\right)\exp(-a(\underline{\theta}+\gamma_2))}{\left(\exp(-a(\underline{\theta}+\gamma_2))-1\right)^2} \tag{A28}$$

$$= \frac{1}{(\exp(-a(\underline{\theta}+\gamma_2))-1)^2}\Big[\exp(-a\Delta\gamma)\Big(-\Delta\gamma\left(\exp(-a(\underline{\theta}+\gamma_2))-1\right) \tag{A29}$$

$$+(\underline{\theta}+\gamma_2)\exp(-a(\underline{\theta}+\gamma_2))\Big)-(\underline{\theta}+\gamma_2)\exp\left(-a(\underline{\theta}+\gamma_2)\right)\Big]$$

$$= \frac{1}{(\exp(-a(\underline{\theta}+\gamma_2))-1)^2} \tag{A30}$$

$$\Big[\exp(-a\Delta\gamma)\Big(\exp\left(-a(\underline{\theta}+\gamma_2)\right)(\underline{\theta}+\gamma_1)+\Delta\gamma\Big)-(\underline{\theta}+\gamma_2)\exp(-a(\underline{\theta}+\gamma_2))\Big]$$

$$= \underbrace{\frac{\exp(-a\Delta\gamma)}{(\exp(-a(\underline{\theta}+\gamma_2))-1)^2}}_{>0} \tag{A31}$$

$$\Big[\underbrace{\Delta\gamma}_{<0}+\underbrace{\exp\left(-a(\underline{\theta}+\gamma_1)\right)}_{>0}\underbrace{\Big(\exp(-a\Delta\gamma)(\underline{\theta}+\gamma_1)-(\underline{\theta}+\gamma_2)\Big)}_{:=f(a)}\Big]$$

using that

$$\exp(-a(\underline{\theta}+\gamma_2))=\exp(-a(\underline{\theta}+\gamma_1))\exp(-a\Delta\gamma). \tag{A32}$$

Then

$$f(a)<0 \Rightarrow \frac{\partial g(a)}{\partial a}<0. \tag{A33}$$

We have

$$f(a)=\exp(-a\Delta\gamma)(\underline{\theta}+\gamma_1)-(\underline{\theta}+\gamma_2)<0 \tag{A34}$$

$$\Leftrightarrow \exp(-a\Delta\gamma)(\underline{\theta}+\gamma_1)<(\underline{\theta}+\gamma_2) \tag{A35}$$

$$\Leftrightarrow \frac{\exp(-a(\underline{\theta}+\gamma_2))}{\exp(-a(\underline{\theta}+\gamma_1))}(\underline{\theta}+\gamma_1)<(\underline{\theta}+\gamma_2) \tag{A36}$$

$$\Leftrightarrow \frac{\exp(-a(\underline{\theta}+\gamma_2))}{(\underline{\theta}+\gamma_2)}<\frac{\exp(-a(\underline{\theta}+\gamma_1))}{(\underline{\theta}+\gamma_1)}. \tag{A37}$$

For $x<0$ the function

$$h(x):=\frac{\exp(-ax)}{x} \tag{A38}$$

is negative and

$$h'(x)=h(x)\left[-a-\frac{1}{x}\right]>0 \;\;\Leftrightarrow\;\; a+\frac{1}{x}>0 \;\;\Leftrightarrow\;\; a(-x)>1. \tag{A39}$$

For $x=\underline{\theta}+\gamma$ this is true from assumption A4. Since $\underline{\theta}+\gamma_2<\underline{\theta}+\gamma_1$, (A37) indeed holds and proves the claim.

ad ii) With $g(\cdot,\vec{\gamma})$ strictly decreasing, existence under (9) follows immediately.

□

A4 Proof of Lemma 6

A protection buyer with risk aversion parameter a chooses protection seller 1 if

$$U_a(b_1,\gamma_1) > U_a(b_2,\gamma_2) \tag{A40}$$
$$\Leftrightarrow (1-b_1p)u_a(-\gamma_1) + b_1pu_a(\underline{\theta}) > (1-b_2p)u_a(-\gamma_2) + b_2pu_a(\underline{\theta}) \tag{A41}$$
$$\Leftrightarrow [u_a(-\gamma_1) - u_a(-\gamma_2)]\,(1-b_1p) > p\Delta b\,\underbrace{(u_a(\underline{\theta}) - u_a(-\gamma_2))}_{<0} \tag{A42}$$
$$\Leftrightarrow \frac{u_a(-\gamma_1) - u_a(-\gamma_2)}{u_a(\underline{\theta}) - u_a(-\gamma_2)} < \frac{p\Delta b}{1-b_1p} \tag{A43}$$
$$\Leftrightarrow g(a) < g(a^*) \tag{A44}$$
$$\Leftrightarrow a > a^*(\gamma_1,\gamma_2). \tag{A45}$$

□

A5 Proof of Lemma 7

The idea is to proceed analogously to the proof of Proposition 1, but add and subtract $b_2u_a(-\gamma_1)$ instead of $b_1u_a(-\gamma_2)$. Namely, for the indifferent protection buyer we have

$$U_a(b_1,\gamma_1) = U_a(b_2,\gamma_2) \tag{A46}$$
$$\Leftrightarrow u_a(-\gamma_1) - u_a(-\gamma_2) + p\,[b_2u_a(-\gamma_2) - b_1u_a(-\gamma_1)] = p\Delta bu_a(\underline{\theta}) \tag{A47}$$
$$\Leftrightarrow [u_a(-\gamma_1) - u_a(-\gamma_2)]\,(1-b_2p) = p\Delta b\,[u_a(\underline{\theta}) - u_a(-\gamma_1)] \tag{A48}$$
$$\Leftrightarrow \frac{u_a(-\gamma_1) - u_a(-\gamma_2)}{u_a(\underline{\theta}) - u_a(-\gamma_1)} = \frac{p\Delta b}{1-b_2p} \tag{A49}$$
$$\Leftrightarrow \frac{1-\exp(-(-a\Delta\gamma))}{\exp(-a(\underline{\theta}+\gamma_1))-1} = \frac{p\Delta b}{1-b_2p}. \tag{A50}$$

□

A6 Proof of Lemma 8

For firm 1. For the function $g(a^*(\vec{\gamma}),\vec{\gamma})$, as defined in (3), we have from Proposition 1

$$0 = d_1g = \partial_1 g|_{a=a^*} + \partial_a g|_{a=a^*}\cdot\partial_1 a^* \tag{A51}$$
$$\Leftrightarrow \partial_1 a^* = \frac{-\partial_1 g|_{a=a^*}}{\partial_a g|_{a=a^*}}. \tag{A52}$$

In the following write $\partial_i g$ shorthand for $\partial_i g|_{a=a^*}$. We have

$$\partial_1 g = a^*\frac{A}{B_2-1} > 0. \tag{A53}$$

and from Proposition 5 we know that $\partial_a g < 0$. Hence, in light of (A52), we have $\partial_1 a^* > 0$.

For the proof of (f1-ii), note that the expression for $\partial_a g$, derived in the proof of Proposition 5, can be written in short-hand notation as follows

$$\partial_a g = \frac{A}{(B_2-1)}\Big[-\Delta\gamma + (\underline{\theta}+\gamma_2)\underbrace{\frac{(A-1)}{(B_2-1)}}_{=g}\underbrace{\frac{B_2}{A}}_{=B_1}\Big] \overset{(A53)}{=} \frac{\partial_1 g}{a^*}\left[-\Delta\gamma + g\varphi_1\right]. \tag{A54}$$

Inserted into (A51) this yields

$$0 = \partial_1 g + \frac{\partial_1 g}{a^*}(-\Delta\gamma + g\varphi_1)\partial_1 a^* \tag{A55}$$

$$= \underbrace{\frac{\partial_1 g}{a^*}}_{>0}\left[a^* + (-\Delta\gamma + g\varphi_1)\partial_1 a^*\right]. \tag{A56}$$

Hence

$$a^* = (\Delta\gamma - g\varphi_1)\underbrace{\partial_1 a^*}_{>0}, \tag{A57}$$

and subsequently

$$(\Delta\gamma - g\varphi_1) > 0. \tag{A58}$$

For firm 2. Analogously, for the function $h(a^*(\vec{\gamma}), \vec{\gamma})$, as defined in (11), we have

$$0 = d_2 h = \partial_2 h|_{a=a^*} + \partial_a h|_{a=a^*} \cdot \partial_2 a^* \tag{A59}$$

$$\Leftrightarrow \partial_2 a^* = \frac{-\partial_2 h|_{a=a^*}}{\partial_a h|_{a=a^*}}. \tag{A60}$$

Similar to before we write $\partial_i h$ shorthand for $\partial_i h|_{a=a^*}$. Then we have

$$\partial_2 h = (-a^*)\frac{1}{A(B_1 - 1)} < 0, \tag{A61}$$

and

$$\partial_a h = (-\Delta\gamma)\frac{1}{A(B_1 - 1)} + (\underline{\theta} + \gamma_1)\frac{(1 - \frac{1}{A})B_1}{(B_1 - 1)^2} \tag{A62}$$

$$= \frac{1}{A(B_1 - 1)^2}\left[\Delta\gamma - \Delta\gamma B_1 - (\underline{\theta} + \gamma_1)B_1 + (\underline{\theta} + \gamma_1)AB_1\right] \tag{A63}$$

$$= \underbrace{\frac{1}{A(B_1 - 1)^2}}_{>0}\left[\underbrace{\Delta\gamma}_{<0} + \underbrace{B_1}_{>0}\left(A(\underline{\theta} + \gamma_1) - (\underline{\theta} + \gamma_2)\right)\right]. \tag{A64}$$

From the proof of Proposition 5 we know that $A(\underline{\theta} + \gamma_1) - (\underline{\theta} + \gamma_2)$ is negative, hence $\partial_a h < 0$. Then from (A60) we get $\partial_2 a^* < 0$.

For the remaining part, note that $AB_1 = B_2$ and hence (A62) can also be written as

$$\partial_a h = \frac{1}{A(B_1 - 1)}\left[-\Delta\gamma + (\underline{\theta} + \gamma_1)AB_1\frac{(1 - \frac{1}{A})B_1}{(B_1 - 1)}\right] \tag{A65}$$

$$= \frac{\partial_a h}{a^*}\left[\Delta\gamma - \varphi_2 h\right]. \tag{A66}$$

Inserted into (A59) this yields

$$0 = \underbrace{\frac{\partial_2 h}{a^*}}_{<0}\left[a^* + (\Delta\gamma - \varphi_2 h)\partial_2 a^*\right]. \tag{A67}$$

Hence,

$$a^* = -(\Delta\gamma - \varphi_2 h)\underbrace{\partial_2 a^*}_{<0}, \tag{A68}$$

and subsequently

$$(\Delta\gamma - \varphi_2 h) > 0, \tag{A69}$$

which concludes the proof.

□

A7 Proof of Lemma 9

ad i). First of all, we show that for a fixed $\gamma_2 \in [0, \gamma^{max}]$ such $\gamma_1^{\underline{a}}(\gamma_2), \gamma_1^{\overline{a}}(\gamma_2)$ indeed exist. Whenever clear form the context we suppress the dependence on γ_2. Note that for $a \in [\underline{a}, \overline{a}]$ $g(a, \gamma_2, \gamma_2) = 0$, while $\lim_{\gamma\to\infty} g(a, \gamma, \gamma_2) = \lim_{\gamma\to\infty} \frac{1}{c_1}(\exp(a\gamma)c_2 - 1) = \infty$ with $c_1 := \exp(-a(\underline{\theta} + \gamma_2))$ and $c_2 := \exp(-a\gamma_2)$ independent of γ. Hence, from continuity such $\gamma_1^{\underline{a}}, \gamma_1^{\overline{a}}$ exist and, since $\partial_1 g > 0$, they are also unique.

Claim. $\gamma_1^{\underline{a}} < \gamma_1^{\overline{a}}$

Proof of claim. Since $\partial_a g < 0$ we have $\tilde{g}(\vec{b}) = g(\overline{a}, \gamma_1^{\overline{a}}, \gamma_2) = g(\underline{a}, \gamma_1^{\underline{a}}, \gamma_2) > g(\overline{a}, \gamma_1^{\underline{a}}, \gamma_2)$. With $\partial_1 g > 0$ this implies $\gamma_1^{\underline{a}} < \gamma_1^{\overline{a}}$.

For the last part of the statement we have

$$\gamma_1^{\underline{a}} \leq \gamma^{max} \tag{A70}$$

$$\Leftrightarrow g(\underline{a}, \gamma^{max}, \gamma_2) \geq \tilde{g}(\vec{b}) \tag{A71}$$

$$\frac{\exp(-\underline{a}(\underline{\theta} + \gamma_2))\exp(-2) - 1}{\exp(-\underline{a}(\underline{\theta} + \gamma_2)) - 1} \geq \tilde{g}(\vec{b}) \tag{A72}$$

$$\exp(-\underline{a}(\underline{\theta} + \gamma_2))\left(\exp(-2) - \tilde{g}(\vec{b})\right) \geq 1 - \tilde{g}(\vec{b}) \tag{A73}$$

$$-\underline{a}(\underline{\theta} + \gamma_2) \geq \log\left[\frac{1 - \tilde{g}(\vec{b})}{\exp(-2) - \tilde{g}(\vec{b})}\right] \tag{A74}$$

$$\gamma_2 \leq (-\underline{\theta}) - \frac{1}{\underline{a}}\log\left[\frac{1 - \tilde{g}(\vec{b})}{\exp(-2) - \tilde{g}(\vec{b})}\right]. \tag{A75}$$

Note that we use $\tilde{g}(\vec{b}) < \exp(-2)$ here, which is ensured by assumptions A1 and A3.

ad ii). The argument for existence is analogous to before, so is the argument for $\gamma_2^{\underline{a}} > \gamma_2^{\overline{a}}$ except that now $\partial_2 g < 0$. For the last part we have

$$\gamma_2^{\underline{a}} \geq 0 \tag{A76}$$

$$\Leftrightarrow g(\underline{a}, \gamma_1, 0) \geq \tilde{g}(\vec{b}) \tag{A77}$$

$$\Leftrightarrow \frac{\exp(\underline{a}\gamma_1) - 1}{\exp(\underline{a}(-\underline{\theta})) - 1} \geq \tilde{g}(\vec{b}) \tag{A78}$$

$$\Leftrightarrow \exp(\underline{a}\gamma_1) \geq 1 + \tilde{g}(\vec{b})\left(\exp(\underline{a}(-\underline{\theta})) - 1\right) \tag{A79}$$

$$\Leftrightarrow \gamma_1 \geq \overline{\gamma_1} =: \frac{1}{\underline{a}}\log\left[1 + \tilde{g}(\vec{b})\left(\exp(\underline{a}(-\underline{\theta}) - 1\right)\right]. \tag{A80}$$

ad iii). We have

$$\overline{\gamma_2} \geq 0 \tag{A81}$$

$$\Leftrightarrow (-\underline{\theta}) - \frac{1}{\underline{a}} \log \left[\frac{1 - \tilde{g}(\vec{b})}{\exp(-2) - \tilde{g}(\vec{b})} \right] \geq 0 \tag{A82}$$

$$\Leftrightarrow \log \left[\frac{1 - \tilde{g}(\vec{b})}{\exp(-2) - \tilde{g}(\vec{b})} \right] \leq \underline{a}(-\underline{\theta}). \tag{A83}$$

At the same time

$$\overline{\gamma_1} \leq \gamma^{max} \tag{A84}$$

$$\Leftrightarrow 2 + \log \left[1 + \tilde{g}(\vec{b}) \left(\exp(\underline{a}(-\underline{\theta})) - 1 \right) \right] \leq \underline{a}(-\underline{\theta}) \tag{A85}$$

$$\Leftrightarrow \log \left[\exp(2) \left(1 + \tilde{g}(\vec{b}) \left(\exp(\underline{a}(-\underline{\theta})) - 1 \right) \right) \right] \leq \underline{a}(-\underline{\theta}) \tag{A86}$$

$$\Leftrightarrow \exp(2) \left(1 + \tilde{g}(\vec{b}) \left(\exp(\underline{a}(-\underline{\theta})) - 1 \right) \right) \leq \exp(\underline{a}(-\underline{\theta})) \tag{A87}$$

$$\Leftrightarrow \exp(2) \left(1 - \tilde{g}(\vec{b}) \right) \leq \exp(\underline{a}(-\underline{\theta})) \left(1 - \exp(2)\tilde{g}(\vec{b}) \right) \tag{A88}$$

$$\Leftrightarrow \frac{1 - \tilde{g}(\vec{b})}{\exp(-2) - \tilde{g}(\vec{b})} \leq \exp(\underline{a}(-\underline{\theta})) \tag{A89}$$

$$\Leftrightarrow \overline{\gamma_2} \geq 0. \tag{A90}$$

ad iv). Since the LHS of (A83) is increasing in $\tilde{g}(\vec{b})$ and under assumption (A3) $\tilde{g}(\vec{b}) < 1/8$,

$$\underline{a}(-\underline{\theta}) > \log \left[\frac{1 - \frac{1}{8}}{\exp(-2) - \frac{1}{8}} \right] \approx 4.4 \tag{A91}$$

ensures $\overline{\gamma_2} \geq 0$ for all admissible parameters and hence renders the set-up interesting.

□

A8 Proof of Proposition 10

We first prove some preliminary claims.

Claim 1. $\partial_2^2 a^* < 0$.

Claim 2. $\alpha := (-\partial_2 a^*)/(\partial_1 a^*) = (1 - gB_1)$.

Claim 3. $\partial_1 \partial_2 a^* > 0$ and reads

$$\partial_1 \partial_2 a^* = \frac{(\partial_1 a^*)(h\varphi_2)}{(\Delta\gamma - h\varphi_2)^2} \left[a^*(\underline{\theta} + \gamma_2) + 2\frac{(\Delta\gamma - h\varphi_2)}{(-h\varphi_2)} \right]. \tag{A92}$$

Proof of claim 1. The proof proceeds by showing that the logarithmic derivative w.r.t. the second argument of $\partial_2 a^*$ is positive. Here, the logarithmic derivative w.r.t. the second argument reads

$$dlog_2(\partial_2 a^*) := \frac{\partial_2^2 a^*}{\partial_2 a^*}. \tag{A93}$$

Since from Lemma 8 $\partial_2 a^* < 0$, $dlog_2(\partial_2 a^*) > 0$ yields $\partial_2^2 a^* < 0$, i.e. $\partial_2 a^*$ strictly decreasing.

With (f2-ii) from Lemma 8 we have

$$
\begin{aligned}
dlog_2(\partial_2 a^*) &= dlog_2(-a^*) - dlog_2(\Delta\gamma - h\varphi_2) && \text{(A94)}\\
&= \frac{-\partial_2 a^*}{-a^*} - \frac{1 - hd_2\varphi_2}{(\Delta\gamma - h\varphi_2)} && \text{(A95)}\\
&= -\frac{1}{(\Delta\gamma - h\varphi_2)} \frac{(-2 + h\varphi_2(\Delta\gamma - h\varphi_2 - (\underline{\theta} + \gamma_2))\partial_2 a^*)}{(\Delta\gamma - h\varphi_2)} && \text{(A96)}\\
&= \frac{1}{(\Delta\gamma - h\varphi_2)^2}\left[-2(\Delta\gamma - h\varphi_2) + h\varphi_2\left[(\Delta\gamma - h\varphi_2) - (\underline{\theta} + \gamma_2)(-a^*)\right]\right] && \text{(A97)}\\
&= \frac{1}{(\Delta\gamma - h\varphi_2)^2}\left[(\Delta\gamma - h\varphi_2)(-2 - h\varphi_2 a^*) - h\varphi_2(-a^*\underline{\theta} + \gamma_2)\right] && \text{(A98)}\\
&= \frac{1}{(\Delta\gamma - h\varphi_2)^2}\Big[-a^*(\underline{\theta} + \gamma_1)\underbrace{\left[\Delta\gamma - h\varphi_2 - (\underline{\theta} + \gamma_2)\right]hB_2}_{=-h\varphi_2(1+hB_2)} - 2(\Delta\gamma - h\varphi_2)\Big] && \text{(A99)}
\end{aligned}
$$

Hence, $dlog_2(\partial_2 a^*)$ is positive iff

$$
-a^*(\underline{\theta} + \gamma_1) > 2\,\underbrace{\frac{(\Delta\gamma - h\varphi_2)}{-h\varphi_2}}_{\in(0,1)}\underbrace{\frac{1}{1 + hB_2}}_{<1}, \qquad \text{(A100)}
$$

which is ensured if $-\underline{a}(\underline{\theta} + \gamma_1) > 2$ (assumption (A2)) and proves the claim.

Proof of claim 2. We first establish that

$$
(1 - gB_1) = \frac{B_1 - 1}{B_2 - 1} = \frac{1}{(1 + hB_2)}. \qquad \text{(A101)}
$$

This follows, since from the definition

$$
\begin{aligned}
1 - gB_1 &= 1 - \frac{A - 1}{B_2 - 1}\frac{B_2}{A} = \frac{B_2 - A}{A(B_2 - 1)} = \frac{B_1 - 1}{B_2 - 1} && \text{(A102)}\\
1 + hB_2 &= 1 + \frac{1 - \frac{1}{A}}{B_1 - 1}B_2 = \frac{B_1 - 1 + B_2 - \frac{B_2}{A}}{B_1 - 1} = \frac{B_2 - 1}{B_1 - 1}. && \text{(A103)}
\end{aligned}
$$

In light of (f1-ii) and (f2-ii) we have

$$
\begin{aligned}
\alpha &= \frac{\Delta\gamma - g\varphi_1}{\Delta\gamma - h\varphi_2} && \text{(A104)}\\
&= \frac{(\underline{\theta} + \gamma_2) - (\underline{\theta} + \gamma_1) - g\varphi_1}{(\underline{\theta} + \gamma_2) - (\underline{\theta} + \gamma_1) - h\varphi_2} && \text{(A105)}\\
&= \frac{(\underline{\theta} + \gamma_2)(1 - gB_1) - (\underline{\theta} + \gamma_1)}{-(\underline{\theta} + \gamma_1)(1 + hB_2) + (\underline{\theta} + \gamma_2)} && \text{(A106)}\\
&= \frac{(1 - gB_1)\left((\underline{\theta} + \gamma_2) - \frac{1}{(1-gB_1)}(\underline{\theta} + \gamma_1)\right)}{(\underline{\theta} + \gamma_2) - (1 + hB_2)(\underline{\theta} + \gamma_1)} && \text{(A107)}\\
&= (1 - gB_1), && \text{(A108)}
\end{aligned}
$$

which proves the claim.

Proof of claim 3. We have

$$
d_1\varphi_2 = B_2 + (\underline{\theta} + \gamma_1)d_1 B_2 \qquad \text{(A109)}
$$

$$= B_2 + (\underline{\theta} + \gamma_1)B_2(-\partial_2 a^*)(\underline{\theta} + \gamma_2) \tag{A110}$$
$$= B_2 \left((1 - (\underline{\theta} + \gamma_1)(\underline{\theta} + \gamma_2)\partial_1 a^*\right) \tag{A111}$$

Then

$$\partial_1\partial_2 a^* = -\partial_1 \left[\frac{a^*}{\Delta\gamma - h\varphi_2}\right] \tag{A112}$$
$$= -\frac{\partial_1 a^*(\Delta\gamma - h\varphi_2) - a^*(-1 - hB_2(1 - (\underline{\theta} + \gamma_1)(\underline{\theta} + \gamma_2)\partial_1 a^*)}{(\Delta\gamma - h\varphi_2)^2} \tag{A113}$$
$$= \frac{-\partial_1 a^* \left[\Delta\gamma - h\varphi_2 - a^*(\underline{\theta} + \gamma_1)(\underline{\theta} + \gamma_2)hB_2\right] + a^*(1 + hB_2)}{(\Delta\gamma - h\varphi_2)^2} \tag{A114}$$
$$= \underbrace{\frac{-\partial_1 a^*}{(\Delta\gamma - h\varphi_2)^2}}_{<0} \underbrace{\left[\Delta\gamma - h\varphi_2 - a^*(\underline{\theta} + \gamma_1)(\underline{\theta} + \gamma_2)hB_2 + (\Delta\gamma - g\varphi_1)(1 + hB_2)\right]}_{=:W} \tag{A115}$$

Hence, $\partial_1\partial_2 a^* > 0$ if the expression in brackets is negative. This is indeed the case, since

$$W = 2\Delta\gamma + hB_2\left[\Delta\gamma - g\varphi_1 - (\underline{\theta} + \gamma_1) - a^*(\underline{\theta} + \gamma_1)(\underline{\theta} + \gamma_2)\right] - g\varphi_1 \tag{A116}$$
$$= \Delta\gamma(2 + hB_2) - (\underline{\theta} + \gamma_1)hB_2 \underbrace{(1 + a^*(\underline{\theta} + \gamma_2))}_{=-1+(2+a^*(\underline{\theta}+\gamma_2))} - g\varphi_1(1 + hB_2) \tag{A117}$$
$$= hB_2\left[(\underline{\theta} + \gamma_1) + \Delta\gamma\right] - g(\underline{\theta} + \gamma_2)B_1(1 + hB_2) - (\underline{\theta} + \gamma_1)hB_2(2 + a^*(\underline{\theta} + \gamma_2)) + 2\Delta\gamma \tag{A118}$$
$$= (\underline{\theta} + \gamma_2)\underbrace{(hB_2 - gB_1(1 + hB_2))}_{=\frac{\Delta B}{B_1 - 1} - \frac{\Delta B}{B_2 - 1}\frac{B_1 - 1}{B_2 - 1} = 0} - (\underline{\theta} + \gamma_1)hB_2(2 + a^*(\underline{\theta} + \gamma_2)) + 2\Delta\gamma \tag{A119}$$
$$= \underbrace{-(\underline{\theta} + \gamma_1)}_{>0} hB_2 \underbrace{(2 + a^*(\underline{\theta} + \gamma_2))}_{<0} + \underbrace{2\Delta\gamma}_{<0} \tag{A120}$$
$$< 0, \tag{A121}$$

which together yields

$$\partial_1\partial_2 a^* = \frac{(\partial_1 a^*)h\varphi_2}{(\Delta\gamma - h\varphi_2)^2}\left[a^*(\underline{\theta} + \gamma_2) + 2\frac{(\Delta\gamma - h\varphi_2)}{(-h\varphi_2)}\right], \tag{A122}$$

and proves the claim.

ad i). Fix some $\gamma_1 \in [0, \gamma^{max}]$. Then

$$d_2\Pi_2 = (a^* - \underline{a}) + \gamma_2\partial_2 a^* = 0 \tag{A123}$$
$$\Leftrightarrow -\partial_2 a^* = \frac{(a^* - \underline{a})}{\gamma_2} \tag{A124}$$

The RHS of (A124) is strictly decreasing in γ_2 and from claim 1 we know that $-\partial_2 a^*$ is strictly increasing. Hence, with the RHS of (A124) strictly increasing and the RHS strictly decreasing, there is maximally one γ_2 s.t. $d_2\Pi_2 = 0$. For existence, note that $\gamma_2 \mapsto \Pi_2(\gamma_1, \gamma_2)$ as continuous function on a compact interval, assumes its maximum. But $\Pi_2(\gamma_1, 0) = \Pi_2(\gamma_1, \gamma^{max}) = 0$, hence the maximum is assumed in the interior.

ad ii). To see that $\gamma_2^* \in \mathcal{C}^1$, we again make use of the implicit function theorem. Note that $d_2\Pi_2 \in \mathcal{C}^1$ and

$$d_2^2\Pi_2 = 2\partial_2 a^* + \gamma_2\partial_2^2 a^* < 0. \tag{A125}$$

Hence, from the implicit function theorem the mapping

$$\gamma_1 \mapsto \gamma_2^*(\gamma_1) = \arg_{\gamma_2}\{d_2\Pi_2(\gamma_1, \gamma_2) = 0\} \tag{A126}$$

is continuously differentiable.

For monotonicity of γ_2^*

$$d_2\Pi_2(\gamma_1, \gamma_2^*(\gamma_1)) \equiv 0 \quad \Leftrightarrow \quad -\partial_2 a^*(\gamma_1, \gamma_2^*(\gamma_1)) \equiv \frac{a^*(\gamma_1, \gamma_2^*(\gamma_1)) - \underline{a}}{\gamma_2^*}. \tag{A127}$$

To simplify notation we define

$$a^{*1} = a^*(\gamma_1, \gamma_2^*(\gamma_1)). \tag{A128}$$

Then

$$\partial\gamma_2^* = \frac{(\partial_1 a^* + (\partial_2 a^*)(\partial\gamma_2^*))(-\partial_2 a^{*1}) + (a^{*1} - \underline{a})(\partial_1\partial_2 a^* + (\partial_2^2 a^*)\partial\gamma_2^*)}{(-\partial_2 a^{*1})^2} \tag{A129}$$

$$\Leftrightarrow \partial\gamma_2^*[\underbrace{(-\partial_2 a^{*1})^2}_{>0} + \underbrace{(\partial_2 a^*)(\partial_2 a^{*1})}_{>0} - \underbrace{(a^{*1} - \underline{a})}_{>0}\underbrace{(\partial_2^2 a^*)}_{<0}] = \underbrace{(\partial_1 a^*)}_{>0}\underbrace{(-\partial_2 a^{*^1})}_{>0} + \underbrace{(a^{*1} - \underline{a})}_{>0}\partial_1\partial_2 a^* \tag{A130}$$

Hence, $\partial_1\partial_2 a^* > 0$ is sufficient for $\partial\gamma_2^* > 0$ and this holds by claim 3.

□

A9 Proof of Proposition 12

At a point $\vec{\gamma}$ with $d_1\Pi_1(\vec{\gamma}) = 0$ we have

$$d_1\Pi_1(\vec{\gamma}) = 0 \;\Leftrightarrow\; (\overline{a} - a^*) - \gamma_1\frac{a^*}{\Delta\gamma - g\varphi_1} = 0 \tag{A131}$$

$$\Leftrightarrow\; \overline{a} - a^*\left[1 + \frac{\gamma_1}{\Delta\gamma - g\varphi_1}\right] = 0 \tag{A132}$$

$$\Leftrightarrow\; a^* = \overline{a}\left[\frac{\Delta\gamma - g\varphi_1}{\gamma_2 - g\varphi_1}\right] \tag{A133}$$

$$\Leftrightarrow\; a^* = \overline{a}\left[1 - \frac{\gamma_1}{\gamma_2 - g\varphi_1}\right] =: f(\vec{\gamma}) \tag{A134}$$

Subsequently, at a point with $0 = d_1\Pi_1(\gamma_1, \gamma_2^*(\gamma_1))$ there is only one candidate for the value of $a^*(\gamma_1, \gamma_2^*(\gamma_1))$, namely

$$f^*(\gamma_1) := f(\gamma_1, \gamma_2^*(\gamma_1)) = \overline{a}\left[1 - \frac{\gamma_1}{\gamma_2^*(\gamma_1) - g\varphi_1(\gamma_1, \gamma_2^*(\gamma_1))}\right]. \tag{A135}$$

We define the following functions

$$\Psi_1 : [0, \gamma^{max}] \to [0, \gamma_2^*(\gamma^{max})] \times [\underline{a}, \overline{a}] \tag{A136}$$

$$\gamma_1 \mapsto (\gamma_2^*(\gamma_1), \min\{\overline{a}, \max\{\underline{a}, f^*(\gamma_1)\}\}) \tag{A137}$$

and

$$\Psi_2 : [0, \gamma_2^*(\gamma^{max})] \times [\underline{a}, \overline{a}] \to [0, \gamma^{max}] \tag{A138}$$

$$(\gamma_2, a) \mapsto \begin{cases} \arg_\gamma\{a^*(\gamma, \gamma_2) - a = 0\} & \text{if } a^*(\gamma^{max}, \gamma_2) \geq a \\ \gamma^{max} & \text{if } a^*(\gamma^{max}, \gamma_2) < a \end{cases} \tag{A139}$$

Claim. Ψ_1 and Ψ_2 are continuous functions.

Proof of claim. Continuity of Ψ_1 follows directly, since γ_2^* is a continuous function from Lemma 10 and subsequently also f^*. For continuity of Ψ_2 we first consider the case if $a^*(\gamma^{max}, \gamma_2) \geq a$. Since $\partial_1 a^* > 0$ there exists a $\gamma \in [0, \gamma^{max}]$ then s.t. $a^*(\gamma, \gamma_2) = a$. Then from the implicit function theorem the mapping

$$(\gamma_2, a) \mapsto \gamma \text{ with } a^*(\gamma, \gamma_2) = a \tag{A140}$$

is continuous. Continuity for $a^*(\gamma^{max}, \gamma_2) < a$ is trivial. It remains to show that the function value at the piecewise-defined function coincides at $a^*(\gamma^{max}, \gamma_2) = a$. But then by definition $\arg_\gamma\{a^*(\gamma, \gamma_2) - a = 0\} = \gamma^{max}$, which proves the claim.

From the claim it follows that the function

$$(\Psi_2 \circ \Psi_1) : [0, \gamma^{max}] \to [0, \gamma^{max}] \tag{A141}$$

is a continuous self-mapping on a nonempty, compact and convex set and, hence, by Brouwer's fixed point theorem (rf Mas-Colell et al. (1995, p. 952)) there exists a fixed point. By construction a fixed point of $(\Psi_2 \circ \Psi_1)$ either satisfies both FOCs or lies at the boundary.

If the point lies at the boundary, consider first the case that $(\gamma^{max}, \gamma_2^{\underline{a}}(\gamma^{max}))$ is a fixed point of $(\Psi_2 \circ \Psi_1)$. Since by construction it lies on firm 2's reaction function, we have $d_2\Pi_2(\gamma^{max}, \gamma_2^{\underline{a}}(\gamma^{max})) = 0$. From (A139) and (A137) we know

$$f(\gamma^{max}, \gamma_2^*(\gamma^{max})) \geq a^*(\gamma^{max}, \gamma_2^*(\gamma^{max})). \tag{A142}$$

But this means, by (A131) - (A134) read backwards, that at $\vec{\gamma} = (\gamma^{max}, \gamma_2^*(\gamma^{max}))$

$$a^* \leq f(\vec{\gamma}) \Leftrightarrow a^* \leq \overline{a}\left[1 - \frac{\gamma_1}{\gamma_2 - g\varphi_1}\right]\Big|_{\vec{\gamma}} \tag{A143}$$

$$\Leftrightarrow d_1\Pi_1(\vec{\gamma}) \geq 0. \tag{A144}$$

In the same spirit, consider the case that $(\gamma_1^{\underline{a}}(0), 0)$ is a fixed point of $(\Psi_2 \circ \Psi_1)$. At that point, by construction $a^*(\vec{\gamma}) = \underline{a}$, hence from (A139) and (A137)

$$f(\vec{\gamma}) \leq \underline{a} = a^* \tag{A145}$$

and, again by (A131) - (A134) read backwards,

$$a^* \geq f(\vec{\gamma}) \Leftrightarrow a^* \geq \overline{a}\left[1 - \frac{\gamma_1}{\gamma_2 - g\varphi_1}\right]\Big|_{\vec{\gamma}} \tag{A146}$$

$$\Leftrightarrow d_1\Pi_1(\vec{\gamma}) \leq 0. \tag{A147}$$

Since we have $d_1\Pi_1(\gamma_1^{\underline{a}}(0), 0) > 0$ by assumption A5, the only admissible boundary point remains $(\gamma^{max}, \gamma_2^{\underline{a}}(\gamma^{max}))$.

□

A10 Proof of Lemma 11

It is to show that assumption A5 implies $d_1\Pi_1(\gamma_1^{\underline{a}}(\gamma_2), \gamma_2) > 0$ for all $\gamma_2 \in [0, \overline{\gamma_2}]$. First of all recall that $\gamma^{max} = \gamma_1^{\underline{a}}(\overline{\gamma_2})$. We know from (f2-i) and (20) that

$$0 \leq d_1\Pi_1|_{(\gamma^{max},\overline{\gamma_2})} = (\overline{a} - a^*(\gamma^{max}, \overline{\gamma_2})) - \gamma^{max}\partial_1 a^*|_{(\gamma^{max},\overline{\gamma_2})} \tag{A148}$$

$$\Leftrightarrow (\overline{a} - \underbrace{a^*(\gamma_1^{\underline{a}}(\overline{\gamma_2}), \overline{\gamma_2})}_{\underline{a}}) \geq \gamma_1^{\underline{a}}(\overline{\gamma_2}) \frac{a^*(\gamma_1^{\underline{a}}(\overline{\gamma_2}), \overline{\gamma_2})}{(\overline{\gamma_2} - \gamma_1^{\underline{a}}(\overline{\gamma_2})) - g\varphi_1^{\underline{a}}(\overline{\gamma_2})} \tag{A149}$$

$$\Leftrightarrow \frac{(\overline{a} - \underline{a})}{\underline{a}} \geq \frac{\gamma_1^{\underline{a}}(\overline{\gamma_2})}{(\overline{\gamma_2} - \gamma_1^{\underline{a}}(\overline{\gamma_2})) - g\varphi_1^{\underline{a}}(\overline{\gamma_2})} \tag{A150}$$

with

$$\varphi_1^{\underline{a}}(\gamma_2) := \varphi_1((\gamma_1^{\underline{a}}(\gamma_2), \gamma_2)). \tag{A151}$$

The idea is to show that the RHS of (A150) as a function of γ_2 is strictly increasing, and, hence, $d_1\Pi_1(\gamma_1^{\underline{a}}(\gamma_2), \gamma_2) > 0$ for $\gamma_2 < \overline{\gamma_2}$. Define the RHS of (A150) and its denominator as a function of γ_2, i.e.

$$F(\gamma_2) := \frac{\gamma_1^{\underline{a}}(\gamma_2)}{(\gamma_2 - \gamma_1^{\underline{a}}(\gamma_2)) - g\varphi_1^{\underline{a}}(\gamma_2)} \tag{A152}$$

$$N(\gamma_2) := (\gamma_2 - \gamma_1^{\underline{a}}(\gamma_2)) - g\varphi_1^{\underline{a}}(\gamma_2). \tag{A153}$$

We know $N > 0$ by (f1-ii). Note that, since $\gamma_1^{\underline{a}}$ is defined via $a^*(\gamma_1^{\underline{a}}(\gamma_2), \gamma_2) = \underline{a}$, the chain rule and definition of α in claim 2 of Appendix A8 directly yields

$$d\gamma_1^{\underline{a}} = \frac{\partial \gamma_1^{\underline{a}}(\gamma_2)}{\partial \gamma_2} = \frac{-\partial_2 a^*}{\partial_1 a^*} = \alpha. \tag{A154}$$

Subsequently, if $dN < 0$ holds for all $\gamma_2 \in [0, \overline{\gamma_2}]$, we have

$$dF = \frac{d\gamma_1^{\underline{a}}}{N} - \gamma_1^{\underline{a}}\frac{dN}{N^2} = \underbrace{\frac{\alpha}{N}}_{>0} \underbrace{-\gamma_1^{\underline{a}}\frac{dN}{N^2}}_{>0} > 0 \tag{A155}$$

and therefore $d_1\Pi_1(\gamma_1^{\underline{a}}(\gamma_2), \gamma_2) > 0$ for all $\gamma_2 < \overline{\gamma_2}$.

It remains to show $dN < 0$ for all $\gamma_2 \in [0, \overline{\gamma_2}]$. To that end first note that, by analogously to (A151) defining $B_1^{\underline{a}}(\gamma_2) := B_1(\gamma_1^{\underline{a}}(\gamma_2), \gamma_2)$, we get

$$dB_1^{\underline{a}}(\gamma_2) = d\big[\exp(-a^*(\gamma_1^{\underline{a}}(\gamma_2), \gamma_2)(\underline{\theta} + \gamma_1^{\underline{a}}(\gamma_2))\big] \tag{A156}$$

$$= B_1^{\underline{a}}(\gamma_2)\big[\underbrace{(-\partial_1 a^* d\gamma_1^{\underline{a}} - \partial_2 a^*)}_{=0 \text{ by (A154)}}(\underline{\theta} + \gamma_1^{\underline{a}}(\gamma_2)) - \underbrace{a^*(\gamma_1^{\underline{a}}(\gamma_2), \gamma_2)}_{\underline{a}}\, d\gamma_1^{\underline{a}}\big] \tag{A157}$$

$$= B_1^{\underline{a}}(\gamma_2) \cdot (-\underline{a}\alpha) \tag{A158}$$

and subsequently

$$d\varphi_1^{\underline{a}}(\gamma_2) = d\left[(\underline{\theta} + \gamma_2)B_1^{\underline{a}}(\gamma_2)\right] \tag{A159}$$

$$= B_1^{\underline{a}}(\gamma_2)\left[(\underline{\theta} + \gamma_2)(-\underline{a}\alpha) + 1\right]. \tag{A160}$$

Together this implies

$$
\begin{aligned}
dN &= d\left[(\gamma_2 - \gamma_1^{\underline{a}}(\gamma_2)) - g\varphi_1^{\underline{a}}(\gamma_2)\right] && \text{(A161)}\\
&= 1 - d\gamma_1^{\underline{a}} - gd\varphi_1^{\underline{a}} && \text{(A162)}\\
&= (1-\alpha) - gB_1^{\underline{a}}(\gamma_2)\left[(\underline{\theta}+\gamma_2)(-\underline{a}\alpha)+1\right] && \text{(A163)}\\
&= (1-\alpha)\underbrace{(1-gB_1^{\underline{a}}(\gamma_2))}_{=\alpha} - \alpha\underbrace{gB_1^{\underline{a}}(\gamma_2)}_{=(1-\alpha)}(1-\underline{a}(\underline{\theta}+\gamma_2)) && \text{(A164)}\\
&= \underbrace{(1-\alpha)\alpha}_{>0}\underbrace{\underline{a}(\underline{\theta}+\gamma_2)}_{<0} && \text{(A165)}\\
&< 0. && \text{(A166)}
\end{aligned}
$$

□

A11 Proof of Proposition 13

Consider $\gamma_2 \in [0, \gamma_2^*(\gamma^{max})]$ such that (γ_1, γ_2) satisfies (31) or (32) for some γ_1. We proceed by showing the following claims.

Claim 1. Π_1 is monotonically increasing on $[\gamma_1^{\underline{a}}(\gamma_2), \mu]$.

Claim 2. There cannot be local minima with FOC $= 0$ on $[\gamma_1^{\underline{a}}(\gamma_2), \gamma^{max}]$.

Claim 3. A local maximum with FOC $= 0$ is also the global maximum and need to lie in $[\mu, \gamma^{max}]$.

A consequence from the claims is that only one of the following cases can occur: (i) Π_1 has a saddle point at μ and a global maximum on $(\mu, \gamma^{max}]$, (ii) Π_1 has exactly one local maximum in $[\mu, \gamma^{max}]$ which is also the global maximum and no saddle point, (iii) Π_1 does not satisfy the FOC anywhere, but has a global maximum at γ^{max}.

In case (ii) and (iii) the point (γ_1, γ_2) is indeed a Nash equilibrium.

Proof of claim 1. From Lemma 11 we know that Π_1 is increasing at $(\gamma_1^{\underline{a}}(\gamma_2), \gamma_2)$. From condition (T), we know that there is no local maximum with FOC $= 0$ on $[\gamma_1^{\underline{a}}(\gamma_2), \mu)$, which proves the claim.

Proof of claim 2. From condition (T) we know that there is no local minimum with FOC $=$ 0 on $(\mu, \gamma^{max}]$. Suppose a local minimum with FOC $= 0$, call it γ^{min}, exists on $(\gamma_1^{\underline{a}}(\gamma_2), \mu]$. From condition (T) we know that on $(\gamma_1^{\underline{a}}(\gamma_2), \mu]$ there can be at most one local minimum, for otherwise a local maximum would lie in between - contradiction. Subsequently, Π_1 would be monotonically decreasing on $[\gamma_1^{\underline{a}}(\gamma_2), \gamma^{min}]$. But this contradicts claim 1.

Proof of claim 3. From condition (T) we know that for $\gamma < \mu$, Π_1 does not have a local maximum with FOC $= 0$. There cannot be a local maximum at the lower boundary from Lemma 11. For $\gamma \geq \mu$, Π_1 has at most one local maximum, since if there were two local maxima $\gamma_1, \gamma_2 \geq \mu$, a local minimum would need to lie in between - contradiction. Hence, there can be at most one, and hence there is exactly one, local maximum (potentially at the upper boundary) and this local maximum has to be the global one.

□

A12 Proof of claim in Remark 1

In Remark 1 we made the following claim, which we prove in the sequel.

Claim 1. There are only finitely many saddle points for the profit function of ps 1, i.e. $\vec{\gamma} \in \mathcal{M}_{\text{saddle}} := \{d_1\Pi_1 = 0 = d_1^2\Pi_1\}$, on $\mathcal{I} := \{\vec{\gamma} \in [0, \gamma^{max}]^2 | \gamma_1 > \gamma_2, a^*(\vec{\gamma}) \in [\underline{a}, \overline{a}]\}$.

We first show the following claims.

Claim 2. There is no point $\vec{\gamma} \in \mathcal{I}$ with $d_1^2\Pi_1(\vec{\gamma}) = 0 = d_1 d_2 \Pi_1(\vec{\gamma})$.

Claim 3. Saddle points on $\mathcal{I}$ are isolated points, i.e. for each saddle point $\vec{\gamma} \in \mathcal{M}_{\text{saddle}} \cap \mathcal{I}$ there is an open set U such that $U \cap \{d_1\Pi_1 = 0 = d_1^2\Pi_1\} = \{\vec{\gamma}\}$.

Proof of claim 2. Idea of the proof is to consider the derivative in the direction in which a^* stays constant and derive at a contradiction. Define $d_a := d_1 + (1/\alpha) \cdot d_2$. Since

$$d_a a^* = d_1 a^* + \frac{d_1 a^*}{-d_2 a^*} d_2 a^* = 0, \tag{A167}$$

this is the directional derivative in the direction in which a^* stays constant. We get

$$\begin{aligned} d_a \gamma_1 &= 1 && \text{(A168)} \\ d_1 d_a &= d_1^2 + d_1\left(\frac{1}{\alpha} d_2\right) && \text{(A169)} \\ &= \underbrace{d_1^2 + \frac{1}{\alpha} d_1 d_2}_{= d_a d_1} + d_1\left(\frac{1}{\alpha}\right) \cdot d_2. && \text{(A170)} \end{aligned}$$

Since

$$\begin{aligned} d_1 \alpha &= d_1(1 - gB_1) && \text{(A171)} \\ &= -g d_1 B_1 && \text{(A172)} \\ &= -gB_1(d_1 a^*)\left[-(\underline{\theta} + \gamma_1) - (\Delta\gamma - g\varphi_1)\right] && \text{(A173)} \\ &= -gB_1(d_1 a^*)(-(\underline{\theta} + \gamma_2))(1 - gB_1) && \text{(A174)} \\ &< 0, && \text{(A175)} \end{aligned}$$

we have

$$d_1\left(\frac{1}{\alpha}\right) > 0. \tag{A176}$$

With

$$\begin{aligned} d_a \Pi_1 &= d_a(\overline{a} - a^*) \cdot \gamma_1 + (\overline{a} - a^*) d_a \gamma_1 && \text{(A177)} \\ &\overset{\text{(A167),(A168)}}{=} 0 + (\overline{a} - a^*) && \text{(A178)} \\ d_1(d_a \Pi_1) &= d_1(\overline{a} - a^*) = -d_1 a^* < 0, && \text{(A179)} \end{aligned}$$

and (A169) - (A170), we have

$$d_a d_1 \Pi_1 = \underbrace{d_1 d_a \Pi_1}_{<0 \text{ from (A179)}} - \underbrace{d_1\left(\frac{1}{\alpha}\right)}_{>0 \text{ from (A176)}} \cdot \underbrace{d_2 \Pi_1}_{>0} < 0. \tag{A180}$$

Now suppose at some point $\vec{\gamma} \in \mathcal{I}$ we had $d_1^2\Pi_1 = 0 = d_1 d_2 \Pi_1$. This directly yields

$$d_a(d_1\Pi_1) = d_1^2\Pi_1 + \frac{1}{\alpha} d_1 d_2 \Pi_1 = 0, \tag{A181}$$

which contradicts (A180) and, hence, proves the claim.

Proof of claim 3. We know the set $\mathcal{M}_{\text{saddle}}$ consisting of saddle points is a subset of the intersection of the sets $\mathcal{M}_{d_1\Pi_1} := \{d_1\Pi_1 = 0\}$ and $\mathcal{M}_{d_1^2\Pi_1} := \{d_1^2\Pi_1 = 0\}$, but also of $\mathcal{M}_{d_1\Pi_1}$ and $\mathcal{M}_R := \{R - 2\bar{a} = 0\}$, with R defined in (A196) in Appendix A13, since by (A191) - (A195) $\mathcal{M}_{d_1\Pi_1} \cap \mathcal{M}_{d_1^2\Pi_1} = \mathcal{M}_{d_1\Pi_1} \cap \mathcal{M}_R$.
Take a point $\vec{\gamma} \in \mathcal{M}_{d_1\Pi_1} \cap \mathcal{M}_R$. On the one hand, from Appendix A13 we know that at $\vec{\gamma}$, $d_1(R - 2\bar{a}) = d_1 R < 0$. Subsequently, from the implicit function theorem, near $\vec{\gamma}$, $\mathcal{M}_R$ ia a smooth submanifold whose tangent vector in $\vec{\gamma}$ is not horizontal. On the other hand, from claim 2, we know that at $\vec{\gamma}$, $d_1 d_2 \Pi_1$ is non-zero. Hence, from the implicit function theorem, near $\vec{\gamma}$, $\mathcal{M}_{d_1\Pi_1}$ is a smooth submanifold whose tangent vector is indeed horizontal, since by definition $d_1^2\Pi_1 = 0$ in $\vec{\gamma}$. Subsequently, the tangent vectors to $\mathcal{M}_{d_1\Pi_1}$ and $\mathcal{M}_R$ at $\vec{\gamma}$ do not coincide and, hence, $\vec{\gamma}$ is an isolated point in $\mathcal{M}_{d_1\Pi_1} \cap \mathcal{M}_R$ and, thus, $\vec{\gamma}$ is also an isolated point of the set $\mathcal{M}_{\text{saddle}}$ of saddle points of ps 1's profit function.

Proof of claim 1. By claim 3, $\mathcal{M}_{\text{saddle}}$ contains only isolated points and is thus closed, and as a subset of $\mathcal{I}$ bounded, hence, compact. Each one-point subset of $\mathcal{M}_{\text{saddle}}$ is open in $\mathcal{M}_{\text{saddle}}$, so, by compactness, a finite set of points of $\mathcal{M}_{\text{saddle}}$ covers $\mathcal{M}_{\text{saddle}}$, i.e. $\mathcal{M}_{\text{saddle}}$ is finite.

□

A13 Proof of Lemma 14

Take some $\gamma_2 \in [0, \gamma_2^*(\gamma^{max})]$ such that (γ_1, γ_2) satisfies (31) or (32) for some γ_1. We need to show existence of a μ s.t. condition (T) holds. We know

$$d_1\Pi_1 = (\bar{a} - a^*) - \gamma_1 \partial_1 a^* = 0 \Leftrightarrow \partial_1 a^* = \frac{(\bar{a} - a^*)}{\gamma_1} \tag{A182}$$

$$d_1^2\Pi_1 = -(2\partial_1 a^* + \gamma_1 \partial_1^2 a^*) \tag{A183}$$

$$d_1\varphi_1 = \varphi_1(\underline{\theta} + \gamma_2)(1 - gB_1)\partial_1 a^* \tag{A184}$$

and

$$\partial_1^2 a^* = \partial_1 \left[\frac{a^*}{(\Delta\gamma - g\varphi_1)} \right] \tag{A185}$$

$$= \frac{\partial_1 a^*}{(\Delta\gamma - g\varphi_1)} + a^* d_1 \left[\frac{1}{(\Delta\gamma - g\varphi_1)} \right] \tag{A186}$$

$$\overset{(A184)}{=} \frac{\partial_1 a^*}{(\Delta\gamma - g\varphi_1)} - a^* \frac{1}{(\Delta\gamma - g\varphi_1)^2} \left[-1 + g\varphi_1(\underline{\theta} + \gamma_2)(1 - gB_1)\partial_1 a^* \right] \tag{A187}$$

$$\overset{(A182)}{=} \frac{\partial_1 a^*}{(\Delta\gamma - g\varphi_1)} \left[2 - g\varphi_1(\underline{\theta} + \gamma_2)(1 - gB_1)\partial_1 a^* \right] \tag{A188}$$

$$= (\partial_1 a^*) \frac{a^*}{(\Delta\gamma - g\varphi_1)} \left[\frac{2}{a^*} - g\varphi_1(\underline{\theta} + \gamma_2)(1 - gB_1) \frac{\partial_1 a^*}{a^*} \right] \tag{A189}$$

$$= (\partial_1 a^*)^2 \left[\frac{2}{a^*} - g\varphi_1(\underline{\theta} + \gamma_2)(1 - gB_1) \frac{1}{(\Delta\gamma - g\varphi_1)} \right]. \tag{A190}$$

Hence, if $d_1\Pi_1 = 0$,

$$d_1^2\Pi_1 \geq 0 \tag{A191}$$

$$\Leftrightarrow 2\partial_1 a^* \leq -\gamma_1(\partial_1 a^*)^2\left[\frac{2}{a^*} - g\varphi_1\frac{(\underline{\theta}+\gamma_2)(1-gB_1)}{\Delta\gamma - g\varphi_1}\right] \tag{A192}$$

$$\overset{d_1\Pi_1=0,(A182)}{\Leftrightarrow} 2 \leq -\gamma_1\frac{(\overline{a}-a^*)}{\gamma_1}\left[\frac{2}{a^*} - g\varphi_1\frac{(\underline{\theta}+\gamma_2)(1-gB_1)}{\Delta\gamma - g\varphi_1}\right] \tag{A193}$$

$$\Leftrightarrow 2a^* \leq (\overline{a}-a^*)\left[-2 + a^* g\varphi_1\frac{(\underline{\theta}+\gamma_2)(1-gB_1)}{\Delta\gamma - g\varphi_1}\right] \tag{A194}$$

$$\Leftrightarrow 2\overline{a} \leq (\overline{a}-a^*)g\varphi_1(\underline{\theta}+\gamma_2)\underbrace{a^*\frac{(1-gB_1)}{\Delta\gamma - g\varphi_1}}_{=(-\partial_2 a^*)}. \tag{A195}$$

Assuming $d_1\Pi_1 = 0$ and if equality holds for one of (A191) - (A195), it holds for all. Defining the RHS of (A195) as

$$R := (\overline{a}-a^*)g\varphi_1(\underline{\theta}+\gamma_2)(-\partial_2 a^*), \tag{A196}$$

we know that there is at most one solution μ for which equality holds in (A195) if $d_1 R < 0$ for all $\gamma_1 \in [\gamma_1^{\underline{a}}(\gamma_2), \gamma^{max}]$. In this case, for all γ with $d_1\Pi_1|_{(\gamma,\gamma_2)} = 0$ and $\gamma > \mu$ (resp. $\gamma < \mu$) one has $d_1^2\Pi_1 > 0$ (resp. $d_1^2\Pi_1 < 0$). Hence, it remains to show $d_1 R < 0$.

Using $\alpha = (-\partial_2 a^*)/(\partial_1 a^*) = (1-gB_1)$ and (A184), we get

$$d_1 R = (-\partial_1 a^*)\frac{R}{(\overline{a}-a^*)} + d_1\varphi_1\frac{R}{\varphi_1} - (\partial_1\partial_2 a^*)\frac{R}{(-\partial_2 a^*)} \tag{A197}$$

$$= R(\partial_1 a^*)\cdot\left[-\frac{1}{(\overline{a}-a^*)} - \alpha(\underline{\theta}+\gamma_2) - \frac{1}{\alpha}\frac{(\partial_1\partial_2 a^*)}{(\partial_1 a^*)^2}\right] \tag{A198}$$

From claim 3 of Appendix A8 and using $1/\alpha = (\Delta\gamma - h\varphi_2)/(\Delta\gamma - g\varphi_1)$ from (A104) we know that

$$\partial_1\partial_2 a^* = \frac{(\partial_1 a^*)(h\varphi_2)}{(\Delta\gamma - h\varphi_2)^2}\left[a^*(\underline{\theta}+\gamma_2) + 2\frac{(\Delta\gamma - h\varphi_2)}{(-h\varphi_2)}\right] \tag{A199}$$

$$\Leftrightarrow \partial_1\partial_2 a^* = (\partial_1 a^*)\left[\frac{h\varphi_2}{(\Delta\gamma - h\varphi_2)^2}a^*(\underline{\theta}+\gamma_2) - 2\frac{1}{(\Delta\gamma - h\varphi_2)}\right] \tag{A200}$$

$$\Leftrightarrow \frac{1}{\alpha}\frac{(\partial_1\partial_2 a^*)}{(\partial_1 a^*)^2} = \frac{1}{\alpha}\frac{1}{(\partial_1 a^*)}\left[\frac{h\varphi_2}{(\Delta\gamma - h\varphi_2)^2}a^*(\underline{\theta}+\gamma_2) - 2\frac{1}{(\Delta\gamma - h\varphi_2)}\right] \tag{A201}$$

$$\Leftrightarrow \frac{1}{\alpha}\frac{(\partial_1\partial_2 a^*)}{(\partial_1 a^*)^2} = \frac{(\Delta\gamma - h\varphi_2)}{(\Delta\gamma - g\varphi_1)}\frac{1}{(\partial_1 a^*)}\frac{1}{(\Delta\gamma - h\varphi_2)}\left[\frac{h\varphi_2}{(\Delta\gamma - h\varphi_2)}a^*(\underline{\theta}+\gamma_2) - 2\right] \tag{A202}$$

$$\Leftrightarrow \frac{1}{\alpha}\frac{(\partial_1\partial_2 a^*)}{(\partial_1 a^*)^2} = \frac{1}{a^*}\left[\frac{h\varphi_2}{(\Delta\gamma - h\varphi_2)}a^*(\underline{\theta}+\gamma_2) - 2\right] \tag{A203}$$

Hence, since $R > 0$,

$$d_1 R = \underbrace{R(\partial_1 a^*)}_{>0}\cdot\left[-\frac{1}{(\overline{a}-a^*)} + \frac{2}{a^*} - (\underline{\theta}+\gamma_2)\left(\alpha + \frac{h\varphi_2}{(\Delta\gamma - h\varphi_2)}\right)\right] \tag{A204}$$

and subsequently

$$d_1 R < 0 \tag{A205}$$

$$\Leftrightarrow \frac{2\overline{a}-3a^*}{(\overline{a}-a^*)a^*} < (\underline{\theta}+\gamma_2)\left(\alpha + \frac{h\varphi_2}{(\Delta\gamma - h\varphi_2)}\right) \tag{A206}$$

$$\Leftrightarrow \underbrace{\frac{2\overline{a}-3a^*}{(\overline{a}-a^*)}}_{<2} < \underbrace{a^*(\underline{\theta}+\gamma_2)}_{<(-2)}(\underbrace{\alpha}_{\in(0,1)} + \underbrace{\frac{h\varphi_2}{(\Delta\gamma - h\varphi_2)}}_{<(-1)}). \tag{A207}$$

(A207) holds if the LHS is negative, i.e.

$$a^* \geq \frac{2}{3}\overline{a},$$

or if the RHS is larger than 2, which is true if

$$-a^*(\underline{\theta}+\gamma_2) > 4 \text{ and } \alpha < \frac{1}{2},$$

or

$$-1 \geq \alpha + \frac{h\varphi_2}{(\Delta\gamma - h\varphi_2)} = \frac{(\Delta\gamma - g\varphi_1)+h\varphi_1}{(\Delta\gamma - h\varphi_2)} = \frac{2\Delta\gamma - g\varphi_1}{(\Delta\gamma - h\varphi_2)} - 1 \;\Leftrightarrow\; 2\Delta\gamma - g\varphi_1 \geq 0.$$

□

A14 Proof of Lemma 15

First of all note that since the reaction function γ_2^* is strictly increasing by Proposition 10, we have $\gamma_2^1 = \gamma_2^*(\gamma_1^1) < \gamma_2^*(\gamma_1^1) = \gamma_2^2$. It is to show that

$$\Pi_1(\vec{\gamma}^2) > \Pi_1(\vec{\gamma}^1) \tag{A208}$$

$$\text{and } \Pi_2(\vec{\gamma}^2) > \Pi_2(\vec{\gamma}^1). \tag{A209}$$

For protection seller 1. Since $d_2\Pi_1 = (-\partial_2 a^*)\gamma_1 > 0$ we know

$$\Pi_1(\gamma_1^1,\gamma_2^1) < \Pi_1(\gamma_1^1,\gamma_2^2) < \Pi_1(\gamma_1^2,\gamma_2^2), \tag{A210}$$

where the last inequality follows since $\vec{\gamma}^2$ is a best response for protection seller 1.

For protection seller 2. It suffices to show that $d_1\Pi_2(\gamma_1,\gamma_2^*(\gamma_1)) > 0$. We have

$$d_1\Pi_2|_{(\gamma_1,\gamma_2^*(\gamma_1))} = d_1\Pi_2|_{(\gamma_1,\gamma_2^*(\gamma_1))} + \underbrace{d_2\Pi_2|_{(\gamma_1,\gamma_2^*(\gamma_1))}}_{=0 \text{ from optimality}} \frac{\partial\gamma_2^*(\gamma_1)}{\partial\gamma_2} \tag{A211}$$

$$= d_1\Pi_2|_{(\gamma_1,\gamma_2^*(\gamma_1))} \tag{A212}$$

$$> 0, \tag{A213}$$

since $d_1\Pi_2 = (\partial_1 a^*)\gamma_2 > 0$.

□

A15 Proof of Lemma 16

First, at a Nash equilibrium $\vec{\gamma}$ one has $d_1\Pi_1(\vec{\gamma}) \geq 0 = d_2\Pi_2(\vec{\gamma})$ with $d_1\Pi_1(\vec{\gamma}) > 0$ only if $\gamma_1 = \gamma^{max}$. Note furthermore that

$$d_1\Pi_1 \geq 0 \Leftrightarrow (\overline{a}-a^*) - \gamma_1\partial_1 a^* \geq 0 \tag{A214}$$

$$d_2\Pi_2 = 0 \Leftrightarrow (a^*-\underline{a}) + \gamma_2\partial_2 a^* = 0. \tag{A215}$$

Using that by claim 2 in the proof of Lemma 10, $\alpha < 1$, it thus follows that at a point $\vec{\gamma}$ with $d_1\Pi_1(\vec{\gamma}) \geq 0 = d_2\Pi_2(\vec{\gamma})$ we have

$$1 > \alpha = \frac{-\partial_2 a^*}{\partial_1 a^*} \geq \frac{-\partial_2 a^*}{(\bar{a} - a^*)}\gamma_1 = \frac{(a^* - \underline{a})}{(\bar{a} - a^*)}\frac{\gamma_1}{\gamma_2} = \frac{\Pi_2}{\Pi_1}\frac{\gamma_1^2}{\gamma_2^2} > \frac{\Pi_2}{\Pi_1}, \tag{A216}$$

where the last inequality follows since $\Delta\gamma < 0 \Leftrightarrow \gamma_1/\gamma_2 > 1$. Hence, (A216) yields $(a^* - \underline{a}) < (\bar{a} - a^*)$ and $\Pi_2 < \Pi_1$.

□

A16 Proof of Lemma 17

The optimization problem for a given vector of default probabilities $\vec{b}^0$ depends only on $\tilde{g}(\vec{b}^0) = p(b_2^0 - b_1^0)/(1 - b_1^0 p)$. Hence, vectors of default probabilities with the same $\tilde{g}$ yield the same Nash equilibria.

Claim. For a given pair of default probabilities $(b_1^0, b_2^0) = \vec{b}^0$ with $\tilde{g}(\vec{b}^0)$, the set of default probabilities $\vec{b}$ with the same $\tilde{g}$ is

$$\left\{\left(b_1^0 - \alpha, b_2^0 - (1 - \tilde{g}(b_1^0, b_2^0))\alpha\right) | \alpha \in \left[b_1^0 - \frac{1}{3}, b_1^0\right]\right\}. \tag{A217}$$

Proof of claim. We have

$$\partial_{b_2}\tilde{g}|_{\vec{b}^0} = \frac{p}{1 - b_1^0 p} \tag{A218}$$

$$\partial_{b_1}\tilde{g}\Big|_{\vec{b}^0} = \frac{-p(1 - b_1^0 p) + p(b_2^0 - b_1^0)p}{(1 - b_1^0 p)^2} \tag{A219}$$

$$= -p\frac{(1 - b_2^0 p)}{(1 - b_1^0 p)^2} \tag{A220}$$

$$= -\frac{p}{(1 - b_1^0 p)}[1 - \underbrace{\frac{p\Delta b}{(1 - b_1^0 p)}}_{=\tilde{g}(\vec{b}^0)}] \tag{A221}$$

$$-\frac{\partial_{b_1}\tilde{g}}{\partial_{b_2}\tilde{g}}\Big|_{\vec{b}^0} = (1 - \tilde{g}(\vec{b}^0)) \in (0, 1) \tag{A222}$$

Hence, from the implicit function theorem we know that sets $\{\vec{\gamma}|\tilde{g}(\vec{\gamma}) = c\}$ are submanifolds that have (for a given c) the same slope $(1 - \tilde{g})$ at each point. Hence they are straight lines.

□

A17 Proof of Lemma 18

Suppose $b_1^0 < b_2^0$ is a Nash equilibrium in the sequential game without assigned roles. In that case one must not be able to find a profitable deviation for the unsafer dealer, that is, no b_2^1 with $b_2^1 < b_1^0 < b_2^0$ such that the profit when taking the lead position in quality, exceeds the profit when choosing the optimal quality as unsafer dealer, i.e. no b_2^1 with $\Pi_1(b_2^1, b_1^0) > \Pi_2(b_1^0, b_2^0)$. From Lemma 17 we know that pairs of default probabilities (b_1, b_2) with

$$(b_1, b_2) = (b_1^0 - \alpha, b_2^0 - (1 - \tilde{g}(b_1^0, b_2^0))\alpha) \tag{A223}$$

$\alpha \in [b_1^0 - \frac{1}{3}, b_1^0]$, lead to the same Nash equilibria in prices. Hence, if the unsafer dealer, protection seller 2, has the option to choose a quality $b_2^1 < b_1^0$ with

$$(b_2^1, b_1^0) = \left(\frac{1}{(1-\tilde{g})} \left[(1-\tilde{g}) b_1^0 - (b_2^0 - b_1^0) \right], b_1^0 \right) \tag{A224}$$

it leads to the same Nash equilibrium in prices, but with reversed roles. By Lemma 16, we know that the protection seller 2 makes greater profits than before, hence, this is a profitable deviation. This deviation is infeasible if

$$b_2^0 - b_1^0 > (1 - \tilde{g}(b_1^0, b_2^0)) b_1^0 \tag{A225}$$

$$\Leftrightarrow b_2^0 > (2 - \tilde{g}(b_1^0, b_2^0)) b_1^0 \tag{A226}$$

$$\Leftrightarrow b_1^0 < \underbrace{\frac{1}{(2-\tilde{g})}}_{<2-1/8 \text{ from Lemma 9}} \underbrace{b_2^0}_{<1/3}. \tag{A227}$$

□

B Appendix: Optimal choice of state-contingent payments

B1 Optimality

This section shows that the derivative (b, γ) is the outcome of the optimal contracting problem described in the text. Consider a protection buyer who is deciding whether to buy a derivative (b, γ). Upon entering the derivative contract, the client agrees to pay a fixed rate γ for establishing the client-dealer relationship, before the volume of the derivative is determined endogenously and the dealer offers the actuarially fair price. In particular, the protection buyer chooses payments (y, z) to maximize expected utility

$$(1-p)u(\overline{\theta}-y)+p(1-b)u(\underline{\theta}-z)+bpu(\underline{\theta}) \tag{B3}$$

subject to the constraint

$$(1-p)y+p(1-b)z-\left[\gamma-\frac{bp\underline{\theta}}{(1-bp)}\right](1-bp)\geq 0 \tag{B4}$$

$$\Leftrightarrow (1-p)y+p(1-b)z\geq\gamma(1-bp)-bp\underline{\theta}. \tag{B5}$$

(B4) and (B5) offer two views on the constraint. (B5) demands that the expected cash flows to the protection seller (LHS) must be at least as high as the expected fee already agreed upon minus the expected endowment if the protection seller survives. To see the latter part note that

$$E\left[\tilde{x}|\text{ps survives}\right]P[\text{ps survives}]=(1-p)\overline{\theta}+p(1-b)\underline{\theta}\overset{E[\tilde{x}]=0}{=}-bp\underline{\theta} \tag{B6}$$

$$\Leftrightarrow E\left[\tilde{x}|\text{ps survives}\right]=\frac{-bp\underline{\theta}}{(1-bp)}>0. \tag{B7}$$

The risk-averse protection buyer passes the risky endowment to the protection seller unless the protection seller defaults.

(B4) offers an alternative explanation. Let γ^{nom} be the expression in brackets, i.e.

$$\gamma^{nom}:=\gamma-\frac{bp\underline{\theta}}{(1-bp)}. \tag{B8}$$

Then the third term on the LHS of (B4) is the "nominal" fee per client-dealer relationship, γ^{nom}, times the survival probability of the protection seller, since only in that case the payment is actually exchanged. It is subtracted because this fee for establishing the client-dealer relationship has already been agreed upon, so the dealer already "mentally set it aside" and subsequently wants to break even in $t=3$. Compared to γ, from the definition we have $\gamma=\gamma^{nom}+bp\underline{\theta}/(1-bp)<\gamma^{nom}$. In view of (B7) the adjustment term, $bp\underline{\theta}/(1-bp)$, is precisely the expected endowment conditional on the survival of the protection seller. Since it is positive, the protection buyer claims this extra revenue for himself, rendering γ the "true" fees for the protection seller. In the formulation of the protection seller's constraint in (B4) one assumes that the protection seller chooses "true" fees γ instead of "nominal" ones γ^{nom}. This reparametrization will make subsequent calculations tractable as we will see, while simplifying the intuition.

Proposition 19. *For a given (b, γ), the protection buyer optimally chooses*

$$y^*(b,\gamma)=\gamma+\frac{p(1-b)\overline{\theta}-p\underline{\theta}}{(1-bp)} \tag{B9}$$

$$z^*(b,\gamma) = \gamma - \frac{(1-p)}{(1-bp)}\overline{\theta} - \frac{b(1-bp)-(1-p)}{(1-b)(1-bp)}\underline{\theta}. \tag{B10}$$

Let $r^*(b,\gamma)$ be the payoff a protection buyer is left with in an optimal derivative contract unless the counterparty defaults (residual endowment), i.e. $r^*(b,\gamma) := \overline{\theta} - y^*(b,\gamma) = \underline{\theta} + z^*(b,\gamma)$. Then, as one would expect from risk aversion, $r^*(b,\gamma)$ does not depend on the endowment state, namely

$$r^*(b,\gamma) = -\gamma. \tag{B11}$$

Proof. See section B2 for a formal proof, below for the intuition. □

The intuition of the result is as follows: Let us rewrite the constraint (B4) under equality,

$$(1-p)y + p(1-b)z = \gamma^{nom}(1-bp) \tag{B12}$$

$$\Leftrightarrow (1-p)(y-\gamma^{nom}) + p(1-b)(z-\gamma^{nom}) = 0. \tag{B13}$$

The risk-averse protection buyer chooses payments (y,z) to equalize his outcome across states, i.e. payments (y,z) such that

$$\overline{\theta} - y = \underline{\theta} - z \tag{B14}$$

$$\Leftrightarrow \qquad y = \overline{\theta} + k \text{ and } z = \underline{\theta} + k \quad \text{for some } k \in \mathbb{R}. \tag{B15}$$

Then the derivative contract can be interpreted as follows: the protection seller offsets the endowment for the protection buyer in each state in exchange for a fixed payment k, leaving the protection buyer with $(-k)$ unless the protection seller defaults. In other words, $-k = \overline{\theta} - y = \underline{\theta} - z$ is the *residual endowment* of the protection buyer. Plugging (B15) into (B13) yields

$$(1-p)(\overline{\theta} + k - \gamma^{nom}) + p(1-b)(\underline{\theta} + k - \gamma^{nom}) = 0 \tag{B16}$$

$$\Leftrightarrow \underbrace{\left[(1-p)\overline{\theta} + p(1-b)\underline{\theta}\right]}_{=-bp\underline{\theta}} + (1-bp)(k-\gamma^{nom}) = 0 \tag{B17}$$

$$\Leftrightarrow k = \gamma^{nom} + \frac{bp\underline{\theta}}{(1-bp)}. \tag{B18}$$

Hence, $-k > -\gamma^{nom}$, i.e. the protection buyer pays *less* than the nominal profit per contract. As explained above, this is because the expected endowment conditional on the protection seller's survival is positive and the protection buyer claims this extra revenue for himself, rendering the "true" profits $k = \gamma = \gamma^{nom} + bp\underline{\theta}/(1-bp)$.

B2 Proof of Proposition 19

The protection buyer solves the following optimization problem

$$\max_{y,z} \left\{(1-p)u(\overline{\theta}-y) + p(1-b)u(\underline{\theta}-z) + bpu(\underline{\theta}) \middle| (1-p)y + p(1-b)z = \gamma(1-bp) - bp\underline{\theta}\right\} \tag{B19}$$

which is equivalent to the unconstrained problem

$$\max_{y} \left\{(1-p)u(\overline{\theta}-y) + p(1-b)u\left(\underline{\theta} - \frac{(1-bp)}{p(1-b)}\gamma + \frac{b}{(1-b)}\underline{\theta} + \frac{(1-p)}{p(1-b)}y\right) + bpu(\underline{\theta})\right\}.$$

With $\Delta\theta := \overline{\theta} - \underline{\theta}$, the resulting first-order condition reads

$$0 = -(1-p)u'(\overline{\theta}-y) + p(1-b)u'\left(\underline{\theta} - \frac{(1-bp)}{p(1-b)}\gamma + \frac{b}{(1-b)}\underline{\theta} + \frac{(1-p)}{p(1-b)}y\right)\frac{(1-p)}{p(1-b)}$$

$$\begin{aligned}
\Leftrightarrow u'(\overline{\theta}-y) &= u'\left(\underline{\theta} - \frac{(1-bp)}{p(1-b)}\gamma + \frac{b}{(1-b)}\underline{\theta} + \frac{(1-p)}{p(1-b)}y\right) \\
\Leftrightarrow \Delta\theta - y\left[1 + \frac{(1-p)}{p(1-b)}\right] &= -\gamma\frac{(1-bp)}{p(1-b)} + \frac{b}{(1-b)}\underline{\theta} \\
\Leftrightarrow y\frac{(1-bp)}{p(1-b)} &= \Delta\theta + \gamma\frac{(1-bp)}{p(1-b)} - \frac{bp}{p(1-b)}\underline{\theta} \\
\Leftrightarrow y &= \frac{p(1-b)}{(1-bp)}\Delta\theta + \gamma - \frac{bp}{(1-bp)}\underline{\theta} \\
\Leftrightarrow y &= \gamma + \frac{p(1-b)(\overline{\theta}-\underline{\theta}) - bp\underline{\theta}}{(1-bp)} \qquad \text{(B20)} \\
\Leftrightarrow y^*(b,\gamma) &= \gamma + \frac{p(1-b)\overline{\theta} - p\underline{\theta}}{(1-bp)}. \qquad \text{(B21)}
\end{aligned}$$

With (B21) plugged into

$$z^*(b,\gamma) = \frac{1-bp}{p(1-b)}\gamma - \frac{pb}{p(1-b)}\underline{\theta} - \frac{(1-p)}{p(1-b)}y^*(b,\gamma) \qquad \text{(B22)}$$

from the constraint in (B19), some simple rearranging yields the formula for $z^*(b,\gamma)$. Using (B20) we confirm that

$$\begin{aligned}
\overline{\theta} - y^*(b,\gamma) &= -\gamma + \frac{\overline{\theta} - bp\overline{\theta} - \left[p(1-b)(\overline{\theta}-\underline{\theta}) - bp\underline{\theta}\right]}{(1-bp)} \qquad \text{(B23)} \\
&= -\gamma + \frac{(1-p)\overline{\theta} + p\underline{\theta}}{(1-bp)} \qquad \text{(B24)} \\
&\overset{E[\tilde{x}]=0}{=} -\gamma \qquad \text{(B25)}
\end{aligned}$$

as well as

$$\begin{aligned}
\underline{\theta} - z^*(b,\gamma) &= \underline{\theta} - \left(\gamma - \frac{(1-p)}{(1-bp)}\overline{\theta} - \frac{b(1-bp)-(1-p)}{(1-b)(1-bp)}\underline{\theta}\right) \qquad \text{(B26)} \\
&= -\gamma + \frac{p\underline{\theta} + (1-p)\overline{\theta}}{(1-bp)} \qquad \text{(B27)} \\
&\overset{E[\tilde{x}]=0}{=} -\gamma. \qquad \text{(B28)}
\end{aligned}$$

□

C Appendix: Standard model of vertical product differentiation revisited

This section clarifies which assumption in the standard model of vertical product differentiation need to be relaxed to yield endogenous market discipline. I revisit the standard model (see e.g. Tirole (1988, section 7.5.1) or Belleflamme and Peitz (2015, chapter 5.3)) and lift the assumptions of full market coverage and quality-invariant costs. The section then shows a refined principle of product differentiation and in how far upward pressure on qualities emerges.

C1 Set-up

Agents. There are two firms that produce the same good, but of different qualities $s_i, i \in 1, 2$ taken from some interval $[\underline{s}, \overline{s}]$, $\underline{s} \geq 0$. There is a continuum of consumers who each demand one unit of the good. Consumers differ in their preference for quality captured by a taste parameter θ. Specifically, a consumer with taste parameter θ derives linear utility $U(p, s) = \theta s - p$ from a good of quality s sold at price p. The taste parameter is assumed to be uniformly distributed over some interval $[\underline{\theta}, \overline{\theta}]$, $\underline{\theta} \geq 0$.

Timing. There are three points in time, $t \in \{0, 1, 2\}$. At date 0, firms simultaneously choose qualities s_i. In $t = 1$, firms simultaneously choose prices p_i upon the publicly observed quality decisions in the previous period. Lastly, consumers decide from whom to buy in $t = 2$. Figure 6 summarizes the simple timing of events.

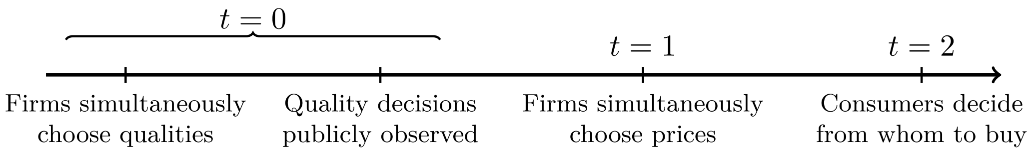

Figure 6: Timeline

If the firms choose the same level of quality, their products can potentially only differ in the price. Since consumers prefer a lower price, competition solely in prices drives the profit margins (or markups) to zero. In order to soften price competition, firms have an incentive to differentiate their products in quality. Since firms are ex-ante symmetric and do not choose the same qualities in equilibrium, if (s_1^*, s_2^*) is an equilibrium in qualities, so is (s_2^*, s_1^*). Without loss of generality we assume that firm 1 is the *low-quality* firm while firm 2 is the *high-quality* firm, that is, suppose $\Delta s := s_2 - s_1 > 0$.[13] I am interested in subgame-perfect Nash equilibria.

C2 Maximal differentiation under full market coverage and constant costs

We briefly review the driving forces at play under the standard assumptions.[14] The standard model assumes that per-unit costs c are the same for all qualities. Additionally the following restrictions on parameters are imposed:

$$\overline{\theta} = \underline{\theta} + 1 \tag{A0}$$

$$\overline{\theta} > 2\underline{\theta} \tag{A1}$$

$$c + \frac{1}{3}(\overline{s} - \underline{s})(\overline{\theta} - 2\underline{\theta}) \leq \underline{\theta}\underline{s}. \tag{A2}$$

Since (A0) and (A1) together imply $\underline{\theta} \in [0, 1)$, they can be understood as demanding that, relative to $\underline{\theta}$, there is sufficient consumer heterogeneity. As will become clear from the prices derived below, the LHS of (A2) is the highest price the low-quality firm might set in equilibrium. The RHS is the lowest possible valuation a consumer can have for the low-quality product. Hence, (A2) ensures that all consumers buy the good (*full market coverage*).

The standard result states that given quality choices $s_1 < s_2$ made in $t = 0$, the prices

$$p_1(s_1, s_2) = c + \frac{1}{3}\Delta s(\overline{\theta} - 2\underline{\theta}) \quad \text{and} \quad p_2(s_1, s_2) = c + \frac{1}{3}\Delta s(2\overline{\theta} - \underline{\theta}) \tag{C3}$$

form a Nash equilibrium in $t = 1$. In $t = 0$, there are two pure-strategy Nash equilibria in the choice of qualities and both exhibit maximal product differentiation. Specifically, for $s_1 < s_2$, firm 1 chooses the lowest possible quality $\underline{s}$ and firm 2 chooses the highest possible quality $\overline{s}$. Reversing the role of the two firms yields the other equilibrium.

The intuition of the result is as follows: In $t = 1$, when qualities $s_1 < s_2$ are already chosen, the consumer who is indifferent between the two firms is characterized by a taste parameter $\hat{\theta}$ such that $\hat{\theta}s_1 - p_1 = \hat{\theta}s_2 - p_2$, hence $\hat{\theta} = (p_2 - p_1)/\Delta s$. Firm 1 receives the consumers with θ below

13 In the presence of multiple equilibria, a coordination issue emerges and one needs to break the symmetry between the two firms somehow. Here, the symmetry is broken by assigning the role of quality-leader ex-ante. Later we choose the other way of breaking the symmetry, i.e. making the quality choice sequential with one firm as first-mover.

14 as in section 7.5.1 in Tirole (1988)

the threshold $\hat{\theta}$, while firm 2 receives those with $\theta > \hat{\theta}$. Firm's profits Π_1 and Π_2 take the form

$$\Pi_1(p_1, p_2) = \underbrace{(p_1 - c)}_{\text{profit margin}} \cdot \underbrace{\left[\frac{(p_2 - p_1)}{\Delta s} - \underline{\theta}\right]}_{\text{market share}}, \quad \Pi_2(p_1, p_2) = (p_2 - c) \cdot \left[\overline{\theta} - \frac{(p_2 - p_1)}{\Delta s}\right]. \tag{C4}$$

In $t = 1$, each firm chooses a price, taking the price of the other firm as given, in order to maximize profits. In $t = 0$, each firm takes into account the Nash equilibrium in prices in the next period, which gives rise to profits as a function of quality choices, specifically $\Pi_1(s_1, s_2) = \frac{1}{9}\Delta s(\overline{\theta} - 2\underline{\theta})^2$ and $\Pi_2(s_1, s_2) = \frac{1}{9}\Delta s(2\overline{\theta} - \underline{\theta})^2$. As profits are increasing in the quality differential, firm 1 chooses the lowest possible quality, while firm 2 chooses the highest possible quality. Note that as a direct consequence the quality-leader enjoys the larger profits - an important observation for later.

The driving forces behind the result of maximal product differentiation are twofold. Firstly, assumption (A2) ensures that the entire market is always covered. Whatever quality choices firms make in $t = 0$ under (A2), they will always be able to optimally respond with their price choices in such a way that the indifferent consumer is left unchanged.[15] This implies that the quantity effect cancels out and only the margin effect is left. For firm 1, for example, we have

$$\frac{\partial \Pi_1(s_1)}{\partial s_1} = \underbrace{\frac{\partial (p_1(s_1) - c)}{\partial s_1}}_{\textit{margin effect}} \underbrace{[\hat{\theta}(s_1) - \underline{\theta}]}_{>0} + (p_1(s_1) - c) \underbrace{\frac{\partial [\hat{\theta}(s_1) - \underline{\theta}]}{\partial s_1}}_{=0,\ \textit{quantity effect}} . \tag{C5}$$

Since prices positively depend on the amount of product differentiation, both firms have an incentive to implement maximal product differentiation. Crucial for this result is that there is no upper limit on the price. For both firms it is optimal to increase prices in response to more product differentiation, keeping the indifferent consumer and as a result the market shares constant. Especially for the high-quality firm which charges the higher price, this means that potentially very large (also relative to costs) prices are set without the risk of loosing customers. Secondly, higher quality is not associated with higher costs.

15 Formally, this can be seen when we insert equilibrium prices into the formula for the indifferent consumer and obtain $\hat{\theta}(s_1, s_2) = \frac{1}{3}(\underline{\theta} + \overline{\theta})$, independent of s_1, s_2.

C3 No full market coverage and costs varying with quality

Let's consider the following generalized set-up. Suppose costs are increasing in quality, that is, suppose there is a smooth "cost" function $c : \mathbb{R}_+ \to \mathbb{R}_+$ with $c' \geq 0$ and $c'' \geq 0$ where the argument is thought of as quality. A firm incurs higher costs when choosing a higher quality, and, at a higher level of quality, increasing quality even further is even more costly.

We lift the assumption that the entire market is covered, i.e. we do not assume (A0), (A1) and (A2) anymore. In the absence of (A2), the symmetry between the two firms vanishes, since firm 1 needs to take into account that at too unfavorable quality and price choices, some consumers might not buy at all. Specifically, a consumer θ_0 is indifferent between not buying at all and buying from the low-quality firm if $p_1 = \theta_0 s_1$. Firm 1 faces only the market segment from θ_0 upwards, which alters its optimization problem to

$$\max_{p_1} \left\{ (p_1 - c(s_1)) \left[\frac{(p_2 - p_1)}{\Delta s} - \max \left\{ \underline{\theta}, \frac{p_1}{s_1} \right\} \right] \right\}. \tag{C6}$$

In order to avoid cumbersome case distinctions that do not seem to carry further intuition, we ensure that $p_1/s_1 \geq \underline{\theta}$ by assuming $\underline{\theta} = 0$.

Attention is restricted to pairs of qualities (s_1, s_2) that satisfy the following assumptions.

Assumption B1. $c(s_1)/s_1 < \overline{\theta}/2$

Assumption B2. $c(s_2)/s_2 < 2\overline{\theta}$

Assumption B3.

$$\frac{\Delta c}{\Delta s} := \frac{c(s_2) - c(s_1)}{\Delta s} \in \left(2\frac{c(s_1)}{s_1} - \overline{\theta}, 2\overline{\theta} - \frac{c(s_2)}{s_2} \right) \tag{C7}$$

Assumption B3 ensures that the markups of both firms are positive. In particular, as will become clear from the equilibrium prices derived below, firm 1's markup will be positive if and only if $\Delta c/\Delta s > 2c(s_1)/s_1 - \overline{\theta}$, while firms 2's markup will be positive if and only if $\Delta c/\Delta s < 2\overline{\theta} - c(s_2)/s_2$. Assumption B3 is a condition on the difference in costs relative to the difference in quality chosen by the two firms. It means that some combinations of (s_1, s_2) kick one firm out of the market, which makes it plausible how a firm may exert a "pull effect" on the quality decisions of the other firm, as shown below. Assumptions B1 and B2 mandate that the upper and lower boundary of the admissible interval in assumption B3 are positive and negative, respectively. Since $\Delta c/\Delta s$ is positive, assumption B2 is a necessary condition, while assumption B1 is only a sufficient condition for positive profit margins of firm 2 and 1 respectively.[16] [17] Assumptions B1 - B3 can be ensured by a large enough $\overline{\theta}$, hence sufficient consumer heterogeneity.

Refined principle of product differentiation

The Nash equilibrium in prices takes the following form.

Proposition 20. *Given quality choices (s_1, s_2) that satisfy assumptions B1 - B3, the following is a Nash equilibrium in prices in $t = 1$:*

$$p_1(s_1, s_2) = \frac{s_1}{3s_2 + \Delta s}\left[c(s_2) + 2\frac{s_2}{s_1}c(s_1) + \overline{\theta}\Delta s\right] \tag{C8}$$

$$= c(s_1) + \frac{s_1}{3s_2 + \Delta s}\left[\Delta c + \Delta s\left(-2\frac{c(s_1)}{s_1} + \overline{\theta}\right)\right] \tag{C9}$$

$$p_2(s_1, s_2) = \frac{s_2}{3s_2 + \Delta s}\left[2c(s_2) + c(s_1) + 2\overline{\theta}\Delta s\right] \tag{C10}$$

$$= c(s_2) + \frac{s_2}{3s_2 + \Delta s}\left[-\Delta c + \Delta s\left(-\frac{c(s_2)}{s_2} + 2\overline{\theta}\right)\right] \tag{C11}$$

Proof. The idea of the proof is analogous to the proof of the standard result presented above in the text. The details are presented in Appendix D1. □

As before, we are interested in whether the quality-leader has higher profits than the low-quality firm. The following corollary shows that this is the case as long as $\Delta c/\Delta s$ lies closer to the lower than to the upper boundary of the admissible interval.

Corollary 21. *i) Firm 2 enjoys larger profit margins than firm 1, i.e. $p_1 - c(s_1) < p_2 - c(s_2)$ if and only if*

$$s_1\left[\frac{\Delta c}{\Delta s} - \left(2\frac{c(s_1)}{s_1} - \overline{\theta}\right)\right] < s_2\left[\left(2\overline{\theta} - \frac{c(s_2)}{s_2}\right) - \frac{\Delta c}{\Delta s}\right].$$

16 If $c(0)$ is normalized to zero, the function $x \mapsto c(x)/x$ is increasing for positive x, since for $x > 0$ we have $\frac{\partial}{\partial x}\left(\frac{c(x)}{x}\right) = \frac{1}{x}\left[c'(x) - \frac{c(x)-c(0)}{(x-0)}\right] \geq 0$ from convexity. But we do not make this assumption here in general as it would rule out fixed costs.

17 Constant costs imply $\Delta c/\Delta s = 0$, hence, satisfy assumption B3 under assumptions B1 and B2.

ii) Firm 2 enjoys larger market shares than firm 1, i.e. $\hat{\theta} - \theta_0 < \overline{\theta} - \hat{\theta}$ *if and only if*

$$\left[\frac{\Delta c}{\Delta s} - \left(2\frac{c(s_1)}{s_1} - \overline{\theta}\right)\right] < \left[\left(2\overline{\theta} - \frac{c(s_2)}{s_2}\right) - \frac{\Delta c}{\Delta s}\right]. \tag{B4}$$

iii) If firm 2 has the higher market share, i.e. if (B4) *is satisfied, it also has the higher profit margin and, as a result, higher profits.*

Proof. Follows directly from plugging in the respective formulas. □

When both firms anticipate the equilibrium in prices for given quality choices, one can express profits as a function of quality choices:

$$\Pi_1(s_1, s_2) = \Delta s \frac{s_2}{s_1}\left[\frac{s_1}{3s_2 + \Delta s}\left(\frac{c(s_2) - c(s_1)}{\Delta s} - \left(2\frac{c(s_1)}{s_1} - \overline{\theta}\right)\right)\right]^2 \tag{C12}$$

$$\Pi_2(s_1, s_2) = \Delta s\left[\frac{s_2}{3s_2 + \Delta s}\left(2\overline{\theta} - \frac{c(s_2)}{s_2} - \frac{c(s_2) - c(s_1)}{\Delta s}\right)\right]^2. \tag{C13}$$

In the original set-up, profits were increasing in the quality differential. Here, in (C12) as well as in (C13), the first factor increases as products become more differentiated, but the effect on the expressions in brackets is unclear. Hence, an interior Nash equilibrium in qualities may be possible. Specifying conditions on the functional form of $c(\cdot)$ that ensure existence of an interior Nash equilibrium does not promise interesting economic results because of lenghty and tedious expressions, and I do not have a general existence proof. The following result, however, derives properties of a Nash equilibrium in qualities and shows a *refined principle of product differentiation.*

Proposition 22. *a) At any point* (s_1, s_2) *that satisfies assumption B1 - B3*

ii) if marginal costs for extra quality are small for firm 1, firm 1 wants to increase quality. Specifically,

$$c'(s_1) < 2\frac{c(s_1)}{s_1} - \overline{\theta} \qquad \Rightarrow \qquad \frac{\partial \Pi_1(s_1, s_2)}{\partial s_1} > 0. \tag{C14}$$

iiii) For firm 2, if marginal costs for extra quality are large, decreasing quality increases *profits. Specifically,*

$$2\overline{\theta} - \frac{c(s_2)}{s_2} < c'(s_2) \qquad \Rightarrow \qquad \frac{\partial \Pi_2(s_1, s_2)}{\partial s_2} < 0. \tag{C15}$$

b) For a sequence of (s_1, s_2) where each pair of qualities satisfies assumptions B1 - B3 and stays distinct while converging to some s_0, i.e. Δs going to zero, we have

$$\lim_{s_1,s_2\to s_0} \frac{\partial \Pi_1(s_1, s_2)}{\partial s_1} = -\frac{1}{9}\left(c'(s_0) - 2\frac{c(s_0)}{s_0} + \overline{\theta}\right)^2 \leq 0, \tag{C16}$$

$$\lim_{s_1,s_2\to s_0} \frac{\partial \Pi_2(s_1, s_2)}{\partial s_2} = \frac{1}{9}\left(2\overline{\theta} - \frac{c(s_0)}{s_0} - c'(s_0)\right)^2 \geq 0. \tag{C17}$$

Proof. See Appendix D2. □

The following observation follows. The threshold $c(s_1)/s_1 - \overline{\theta}$ in (C14) indeed also depends on s_1. It can be meaningfully interpreted, since by assumption B3, $\Delta c/\Delta s$ needs to lie above this threshold. Analogously for the threshold in (C15).

Proposition 4 part b) shows that, if qualities are very close together, i.e. when Δs is small, firms want to differentiate qualities. In other words, the same effect as in the original model prevails, but now it is only an "infinitesimal" effect as it holds for small differences in quality. At the same time, Proposition 4 part a) demonstrates that high or low marginal costs for firm 2 or 1, respectively, can be the driver behind a tendency to move qualities closer together. From Proposition 4 part aii) the quality-leader wants to provide only as much quality as "necessary", while from part ai) the low-quality firm provides "as much quality as feasible" with respect to the increasing marginal costs of quality. Together the forces from part a) and b) act like pull and push factors keeping the qualities of the two firms somewhat close together, but never equal, as illustrated in Figure 7.

C4 Upward pressure on qualities

Two questions arise naturally. Firstly, since the low-quality firm now experiences competition from above (the high-quality firm) and below (the option not to buy), does that exert a pull effect on the quality choice of firm 1? Secondly, when the leadership position in quality is the more attractive one, can the threat to be overtaken by the other firm induce the quality-leader

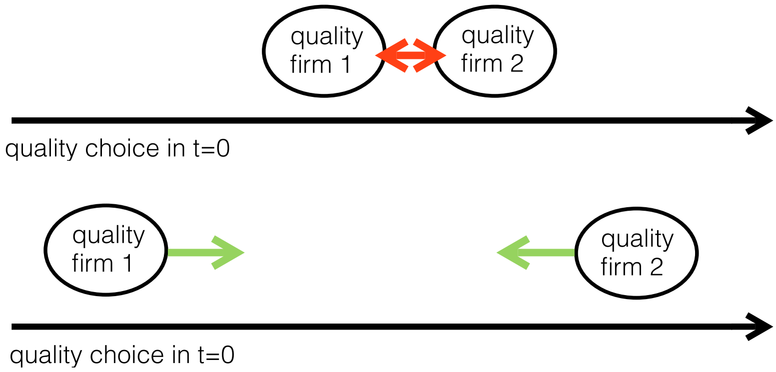

Figure 7: Without full market coverage and without quality-invariant costs there are push *and* pull factors keeping the quality choices somewhat close, but never equal.

to set high qualities whatsoever? The interplay of these forces would produce upward pressure on qualities.

The following proposition and subsequent discussion clarifies in how far there may be a pull effect on the quality choice of the low-quality firm.

Proposition 23. *At any point* (s_1, s_2) *that satisfies assumptions B1 - B3, if*

$$K := \overline{\theta}\underbrace{(s_2 - 2s_1)}_{=:A} + \underbrace{\left(2s_2 - \Delta s\frac{s_1}{s_2}\right)}_{>0}\underbrace{\left[\frac{c(s_1)}{s_1} - c'(s_1)\right]}_{=:B} + \underbrace{\Delta s\frac{s_1}{s_2}}_{>0}\underbrace{\left[\frac{\Delta c}{\Delta s} - c'(s_1)\right]}_{=:C} + \underbrace{2\Delta s}_{>0}\underbrace{(-c'(s_1))}_{=:D}$$

is non-negative, then $\partial\Pi_1/\partial s_1 > 0$ *and subsequently the point can not be an equilibrium.*

Proof. See Appendix D3. □

We discuss the consequences for the special case of constant costs $c \in \mathbb{R}_+$, quadratic costs and the general case. For constant costs, K reduces to $\overline{\theta}(s_2 - 2s_1) + (2s_2 - (\Delta s)s_1/s_2)\, c/s_1$. For $s_2 \geq 2s_1$ this expression is positive, subsequently the point can not be an equilibrium. This admits the following interpretation: In order for an equilibrium to exist, the low-quality firm needs to choose s_1 sufficiently close to the quality of firm 2, i.e. larger than $0.5\, s_2$ (*pull effect*).[18] For quadratic costs, which play a prominent role in the literature on the subject, say $c(s) = \tau s^2$, K reduces to $K = \overline{\theta}(s_2 - 2s_1) - \tau s_1 \underbrace{(5s_2 - 3s_1)}_{>0}$. So $K \geq 0$ requires $(s_2 - 2s_1) > 0$ and is fulfilled if

18 In the case of constant costs, one can easily show that firm 2 chooses the maximal quality. This is intuitive, as higher quality is not associated with higher costs in this case. The simplification of constant costs helps show the key idea of a "pull" effect exerted on the low-quality firm most clearly, but it also eliminates the force that previously counteracted the quality-leader's incentive to choose the extreme quality.

$0 \leq \tau < \overline{\theta}(s_2 - 2s_1)/(5s_2 - 3s_1)$. This again has an intuitive interpretation when we think of the costs $c(s) = \tau s^2$ as a quadratic "error term" to zero costs with "intensity" τ. A non-negative K requires that the condition $s_2 \geq 2s_1$, which precludes an equilibrium for zero costs, still suffices to preclude existence for quadratic costs provided the "intensity" τ of the "error term" is below some threshold.
For the general case, K consists of "drivers" A, B, C and D, as defined above, with positive weights. For a fixed s_2, each driver is monotone in s_1 and the level of s_1 determines whether the corresponding driver increases or decreases K, i.e. whether it exerts upward pressure or not. Specifically, A is positive iff $s_1 < 1/2\, s_2$, B is positive iff s_1 is smaller than s_0 with s_0 such that $c'(s_0) = c(s_0)/s_0$, C is positive for $s_1 \neq s_2$ and D is always negative.

That the quality-leader exerts a "pull effect" on the low-quality firm upwards rather than the other way around is intuitive also from a different point of view. Already in the original model the quality-leader enjoys greater profits. Albeit the fact that the low-quality firm will subsequently choose the lowest quality there, this indicates that there is room for a race for the "pole position in quality", as also noted in Tirole (1988, p. 297). Corollary 21 shows that this result persists in the generalized set-up under the condition that the relation $\Delta c/\Delta s$ may not be too large. Specifically, if $\Delta c/\Delta s$ lies closer to the lower than to the upper boundary of the admissible interval, the lead position in quality is the more attractive one and the quality-leader will try to keep this "pole position". It seems plausible that the quality-leader is aware of the risk of being overtaken by the other firm at too low quality choices. Then the risk of being overtaken may exert upward pressure on the quality choices when moving qualities closer together. This is shown formally in the sequel.

To capture this, suppose we break the symmetry between the two firms not, as done so far, by assigning the roles of quality-leader and quality-follower ex-ante, but instead by making the quality choice sequential. We call the new set-up *sequential game without assigned roles* and assume firm 2 has a first-mover advantage in the choice of quality. Specifically, we introduce an additional time period $t = (-1)$ in which firm 2 chooses its quality, while firm 1, upon observing firm 2's decision, continues to choose its quality in $t = 0$. The rest remains as before.

In $t = 0$, firm 1 can either "adapt" by actually becoming the quality-follower or overtake firm 2's leadership position by choosing a higher quality. We ensure assumptions B1 - B3 and (B4) for all quality pairs by assuming that for all s in $[\underline{s}, \overline{s}]$

$$\frac{c(s)}{s} < \frac{\overline{\theta}}{2} \tag{B1'}$$

$$c'(s) \in \left(2 \sup_t \frac{c(t)}{t} - \overline{\theta}, 2\overline{\theta} - \inf_t \frac{c(t)}{t}\right) \tag{B3'}$$

$$\overline{\theta} - 2 \inf_t \frac{c(t)}{t} + c'(s) < 2\overline{\theta} - \sup_t \frac{c(t)}{t} - c'(s). \tag{B4'}$$

(B1') - (B4') relate marginal costs of a further quality improvement to $\overline{\theta}$, the marginal willingness to pay of the most quality-sensitive consumer for a quality improvement. Note that with $c(s_2) - c(s_1) = \int_{s_1}^{s_2} c'(t)dt$, (B3') yields assumption B3 for all $s \in [\underline{s}, \overline{s}]$, while (B4') ensures condition (B4) for all qualities. Conditions (B1'), (B3') and (B4') can be ensured if $\overline{\theta}$ is large enough.[19]

Hence, the quality-leader always enjoys larger profits, which enables us to derive the following proposition.

Proposition 24. *A necessary condition for some (s_1, s_2) to be a subgame-perfect Nash equilibrium in the sequential game without assigned roles, is that*

$$s_2 > \frac{4}{5}\overline{s}. \tag{C18}$$

Proof. As before, the main idea is presented below in the text, while some calculations are relegated to Appendix D4. □

The intuition of the result is as follows: Suppose $s_1 < s_2$ is a Nash equilibrium in the sequential game without assigned roles. In that case one must not be able to find a profitable deviation for the quality-follower, that is, no s_3 with $s_2 < s_3 \leq \overline{s}$ such that the profit when taking the lead position in quality, exceeds the profit when choosing the optimal quality as quality-follower, that is no s_3 $\Pi_2(s_2, s_3) > \Pi_1(s_1, s_2)$. As shown in the appendix, $s_3 = (s_2^2 + s_1 s_2 - s_1^2)/s_2$ is such an profitable deviation, which is infeasible if $(4/5)\overline{s} < s_2$.

Proposition 24 shows that in the sequential game without assigned roles, a necessary condition

19 In the same spirit as in the original model, this can be interpreted as a condition on sufficient consumer heterogeneity, and thereby neatly connects to the set of assumptions made in the original model. There, (A0) and (A1) demand sufficient consumer heterogeneity while (A2) demands full market coverage; here, only sufficient consumer heterogeneity is needed.

for a Nash equilibrium to exist is that the quality-leader chooses a quality at least as high as 80% of the maximal quality, as illustrated in Figure 8. In other words, the threat of being overtaken and loosing the leadership position in quality induces the first-mover to pick a high quality even in an environment where costs are increasing and convex in the level of quality.

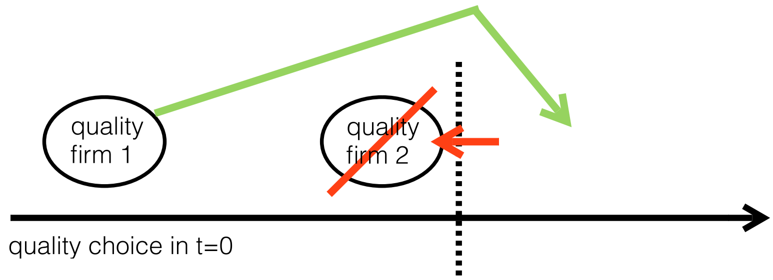

Figure 8: The first-mover wants to keep the leadership position in quality, exerting upward pressure on the qualities.

The interplay between a pull effect on the quality choice of the low-quality firm and pressure on the high-quality firm not to leave too much room quality-wise above, gives rise to upward pressure on the quality choices.

D Appendix: Proofs of appendix C

D1 Proof of Proposition 20

The full maximization problem reads

$$\max_{p_1} \Pi_1(p_1, p_2) = \max_{p_1} \left\{ (p_1 - c(s_1)) \left[\frac{(p_2 - p_1)}{\Delta s} - \frac{p_1}{s_1} \right] \right\} \tag{D3}$$

$$\max_{p_2} \Pi_2(p_1, p_2) = \max_{p_2} \left\{ (p_2 - c(s_2)) \left[\overline{\theta} - \frac{(p_2 - p_1)}{\Delta s} \right] \right\}, \tag{D4}$$

with the additional conditions

$$(p_1 - c(s_1)) \geq 0 \quad \text{positive profit margin of firm 1} \tag{Bi}$$

$$(p_2 - c(s_2)) \geq 0 \quad \text{positive profit margin of firm 2} \tag{Bii}$$

$$\frac{p_2 - p_1}{\Delta s} \geq \frac{p_1}{s_1} \quad \text{positive market share of firm 1} \tag{Biii}$$

$$\overline{\theta} \geq \frac{p_2 - p_1}{\Delta s} \quad \text{positive market share of firm 2} \tag{Biv}$$

$$\frac{p_1}{s_1} \geq \underline{\theta} \quad \text{firm 1's market share takes the form } \frac{(p_2 - p_1)}{\Delta s} - \frac{p_1}{s_1} \tag{Bv}$$

$$\frac{p_2 - p_1}{\Delta s} \geq \frac{p_2}{s_2} \quad \text{firm 2's market share takes the form } \overline{\theta} - \frac{(p_2 - p_1)}{\Delta s}. \tag{Bvi}$$

I first solve the unconstrained maximization problem and then verify that the (unique) solution satisfies (Bi) - (Bvi). Solving the reaction functions

$$p_1 = R_1(p_2) := \frac{1}{2} \left[p_2 \frac{s_1}{s_2} + c(s_1) \right] \tag{D5}$$

$$p_2 = R_2(p_1) := \frac{1}{2} \left[p_1 + c(s_2) + \overline{\theta} \Delta s \right] \tag{D6}$$

yields the formula for the prices.

It remains to check whether conditions (Bi) - (Bvi) hold. (Bi) and (Bii) are ensured by (B3) as argued in the text. Since plugging in the respective formulas directly yields

$$\hat{\theta} - \theta_0 = \frac{s_2}{(3s_2 + \Delta s)} \left[\frac{\Delta c}{\Delta s} - \left(2 \frac{c(s_1)}{s_1} - \overline{\theta} \right) \right] \tag{D7}$$

$$\overline{\theta} - \hat{\theta} = \frac{s_2}{(3s_2 + \Delta s)} \left[\left(2\overline{\theta} - \frac{c(s_2)}{s_2} \right) - \frac{\Delta c}{\Delta s} \right], \tag{D8}$$

(B3) ensures (Biii) and (Biv). (Bv) follows directly from the assumption $\underline{\theta} = 0$, since prices are positive. It remains to show (Bvi), which is a little more cumbersome. As a first step note that (Bvi) follows if we know that

$$\frac{p_2}{p_1} \geq \frac{s_2}{s_1}, \tag{D9}$$

since then

$$p_2 s_1 \geq p_1 s_2 \tag{D10}$$

$$\Leftrightarrow p_2 s_1 - p_2 s_2 + p_2 s_2 \geq p_1 s_2 \tag{D11}$$

$$\Leftrightarrow -p_2 \Delta s + s_2 \Delta p \geq 0 \tag{D12}$$

$$\Leftrightarrow \frac{\Delta p}{\Delta s} \geq \frac{p_2}{s_2}. \tag{D13}$$

It remains to show that (D9) holds. To that end we have

$$\frac{p_2}{p_1} \geq \frac{s_2}{s_1} \tag{D14}$$

$$\Leftrightarrow \frac{\frac{s_2}{3s_2+\Delta s}\left[2c(s_2)+c(s_1)+2\overline{\theta}\Delta s\right]}{\frac{s_1}{3s_2+\Delta s}\left[c(s_2)+2\frac{s_2}{s_1}c(s_1)+\overline{\theta}\Delta s\right]} \geq \frac{s_2}{s_1} \tag{D15}$$

$$\Leftrightarrow \frac{2c(s_2)+c(s_1)+2\overline{\theta}\Delta s}{c(s_2)+2\frac{s_2}{s_1}c(s_1)+\overline{\theta}\Delta s} \geq 1 \tag{D16}$$

$$\Leftrightarrow c(s_2)+c(s_1)\underbrace{\left(1-2\frac{s_2}{s_1}\right)}_{=-\frac{(s_2+\Delta s)}{s_1}}+\overline{\theta}\Delta s \geq 0 \tag{D17}$$

$$\Leftrightarrow c(s_2)-\frac{s_2}{s_1}c(s_1)+\frac{c(s_1)}{s_1}(-\Delta s+2\Delta s)+\Delta s\left[\overline{\theta}-2\frac{c(s_1)}{s_1}\right] \geq 0 \tag{D18}$$

$$\Leftrightarrow \frac{\Delta c}{\Delta s} \geq 2\frac{c(s_1)}{s_1}-\overline{\theta}, \tag{D19}$$

which is ensured by (B3).

□

D2 Proof of Proposition 22

Part a) follows immediately, if we know the following expressions for the derivatives of the profits. With α and β the expressions inside the squared brackets in (C12) and (C13), namely

$$\alpha(s_1,s_2) := \frac{s_1}{3s_2+\Delta s}\left(\frac{\Delta c}{\Delta s}-\left(2\frac{c(s_1)}{s_1}-\overline{\theta}\right)\right) \tag{D20}$$

$$\beta(s_1,s_2) := \frac{s_2}{3s_2+\Delta s}\left(2\overline{\theta}-\frac{c(s_2)}{s_2}-\frac{\Delta c}{\Delta s}\right). \tag{D21}$$

we claim that

$$\frac{\partial \Pi_1(s_1,s_2)}{\partial s_1} = -\frac{s_2^2}{s_1^2}\alpha^2+2\alpha\frac{s_2}{s_1}\Delta s\frac{\partial \alpha(s_1,s_2)}{\partial s_1} \tag{D22}$$

$$= \underbrace{\frac{s_2}{s_1}\frac{\alpha}{(3s_2+\Delta s)^2}}_{>0}\Bigg[(2\Delta s(3s_2+\Delta s)+3s_1s_2)\underbrace{\left(\frac{\Delta c}{\Delta s}-c'(s_1)\right)}_{\geq 0 \text{ from convexity}} \tag{D23}$$

$$+s_2(3s_2+\Delta s)\left(2\frac{c(s_1)}{s_1}-\overline{\theta}-c'(s_1)\right)+4s_2\Delta s\underbrace{\left(2\overline{\theta}-\frac{c(s_2)}{s_2}-c'(s_1)\right)}_{>0 \text{ from (B3)}}\Bigg],$$

$$\frac{\partial \Pi_2(s_1,s_2)}{\partial s_2} = \beta^2+2\beta\Delta s\frac{\partial \beta(s_1,s_2)}{\partial s_2} \tag{D24}$$

$$= \underbrace{\frac{\beta}{(3s_2+\Delta s)^2}}_{>0}\Bigg[(3s_2+\Delta s)(s_2+2\Delta s)\underbrace{\left(\frac{\Delta c}{\Delta s}-c'(s_2)\right)}_{\leq 0 \text{ from convexity}} \tag{D25}$$

$$+4s_1\Delta s\underbrace{\left(2\frac{c(s_1)}{s_1}-\overline{\theta}-\frac{\Delta c}{\Delta s}\right)}_{<0 \text{ from (B3)}}+(3s_2+\Delta s)s_2\left(2\overline{\theta}-\frac{c(s_2)}{s_2}-c'(s_2)\right)\Bigg].$$

To show this, note that for firm 1 the derivative can be written as follows

$$\frac{\partial\Pi_1(s_1,s_2)}{\partial s_1} = -\frac{s_2^2}{s_1^2}\alpha^2+2\alpha\frac{s_2}{s_1}\Delta s\frac{\partial\alpha(s_1,s_2)}{\partial s_1} \tag{D26}$$

$$\overset{\text{Def }\alpha}{=} \frac{s_2}{s_1}\frac{\alpha}{(3s_2+\Delta s)^2}\Bigg[-\frac{s_2}{s_1}(3s_2+\Delta s)s_1\left(\frac{\Delta c}{\Delta s}-\left(2\frac{c(s_1)}{s_1}-\overline{\theta}\right)\right)$$
$$+2\Delta s\frac{\partial\alpha(s_1,s_2)}{\partial s_1}(3s_2+\Delta s)^2\Bigg]. \tag{D27}$$

For the derivative of $\alpha(s_1,s_2)$ w.r.t. s_1 it proves helpful to use two versions of the formula for α when applying the product rule, namely

$$\alpha(s_1,s_2) := \frac{1}{3s_2+\Delta s}\left(s_1\frac{\Delta c}{\Delta s}-\left(2c(s_1)-s_1\overline{\theta}\right)\right) \tag{D28}$$

$$= \frac{1}{3s_2+\Delta s}\left(\frac{s_1}{\Delta s}c(s_2)-\frac{s_2+\Delta s}{\Delta s}c(s_1)+s_1\overline{\theta}\right). \tag{D29}$$

Then

$$\frac{\partial\alpha(s_1,s_2)}{\partial s_1}(3s_2+\Delta s)^2 = \left[\overline{\theta}+\frac{s_2}{(\Delta s)^2}c(s_2)-\frac{s_2}{(\Delta s)^2}c(s_1)-\frac{s_2+\Delta s}{\Delta s}c'(s_1)\right](3s_2+\Delta s)$$
$$+s_1\overline{\theta}-2c(s_1)+s_1\frac{\Delta c}{\Delta s}$$
$$= 4\overline{\theta}s_2-2c(s_1)+\frac{(s_2+\Delta s)(3s_2+\Delta s)}{\Delta s}\left[\frac{\Delta c}{\Delta s}-c'(s_1)\right]$$
$$-(3s_2+\Delta s)\frac{\Delta c}{\Delta s}+s_1\frac{\Delta c}{\Delta s}$$
$$= \frac{(s_2+\Delta s)(3s_2+\Delta s)}{\Delta s}\left[\frac{\Delta c}{\Delta s}-c'(s_1)\right]$$
$$+\underbrace{\left[-(3s_2+\Delta s)+s_1+2\Delta s\right]}_{=-2s_2}\frac{\Delta c}{\Delta s}+2s_2\left(2\overline{\theta}-\frac{c(s_2)}{s_2}\right)$$
$$= \frac{(s_2+\Delta s)(3s_2+\Delta s)}{\Delta s}\left[\frac{\Delta c}{\Delta s}-c'(s_1)\right]+2s_2\left[2\overline{\theta}-\frac{c(s_2)}{s_2}-\frac{\Delta c}{\Delta s}\right].$$

Hence together with (D27)

$$\frac{\partial\Pi_1(s_1,s_2)}{\partial s_1} = \frac{s_2}{s_1}\frac{\alpha}{(3s_2+\Delta s)^2}\Bigg[s_2(3s_2+\Delta s)\left(2\frac{c(s_1)}{s_1}-\overline{\theta}-\frac{\Delta c}{\Delta s}\right) \tag{D30}$$
$$+2(s_2+\Delta s)(3s_2+\Delta s)\left(\frac{\Delta c}{\Delta s}-c'(s_1)\right)$$
$$+(2\Delta s)2s_2\left(2\overline{\theta}-\frac{c(s_2)}{s_2}-\frac{\Delta c}{\Delta s}\right)\Bigg]$$

$$= \frac{s_2}{s_1}\frac{\alpha}{(3s_2+\Delta s)^2}\Bigg[s_2\Delta s\left(2\frac{c(s_1)}{s_1}-\overline{\theta}\right)+s_2\Delta s\left(2\overline{\theta}-\frac{c(s_2)}{s_2}\right) \tag{D31}$$
$$-s_2(3s_2+\Delta s)\frac{\Delta c}{\Delta s}+2s_2\left[3s_2\left(2\frac{c(s_1)}{s_1}-\overline{\theta}\right)+3\Delta s\left(2\overline{\theta}-\frac{c(s_2)}{s_2}\right)\right]$$

$$
\begin{aligned}
& -4s_2\Delta s\frac{\Delta c}{\Delta s} - 2(s_2+\Delta s)(3s_2+\Delta s)c'(s_1) \\
& +2s_2(3s_2+\Delta s)\frac{\Delta c}{\Delta s} + 2\Delta s(3s_2+\Delta s)\frac{\Delta c}{\Delta s}\Bigg] \\
= \; & \frac{s_2}{s_1}\frac{\alpha}{(3s_2+\Delta s)^2}\Bigg[2\Delta s(3s_2+\Delta s)\left(\frac{\Delta c}{\Delta s}-c'(s_1)\right) \qquad \text{(D32)} \\
& -2s_2(3s_2+\Delta s)c'(s_1) + s_2(3s_2+\Delta s)\left(2\frac{c(s_1)}{s_1}-\overline{\theta}\right) \\
& +s_2\left(3\Delta s+\Delta s\right)\left(2\overline{\theta}-\frac{c(s_2)}{s_2}\right) - 4s_2\Delta s\frac{\Delta c}{\Delta s} + s_2(3s_2+\Delta s)\frac{\Delta c}{\Delta s}\Bigg] \\
= \; & \frac{s_2}{s_1}\frac{\alpha}{(3s_2+\Delta s)^2}\Bigg[2\Delta s(3s_2+\Delta s)\left(\frac{\Delta c}{\Delta s}-c'(s_1)\right) \qquad \text{(D33)} \\
& +s_2(3s_2+\Delta s)\left(2\frac{c(s_1)}{s_1}-\overline{\theta}-c'(s_1)\right) \\
& +s_2(3s_2+\Delta s)\left(2\overline{\theta}-\frac{c(s_2)}{s_2}-c'(s_1)\right) \\
& +\underbrace{(4s_2\Delta s - s_2(3s_2+\Delta s))}_{=-3s_1s_2}\left(2\overline{\theta}-\frac{c(s_2)}{s_2}-\frac{\Delta c}{\Delta s}\right)\Bigg] \\
= \; & \frac{s_2}{s_1}\frac{\alpha}{(3s_2+\Delta s)^2}\Bigg[2\Delta s(3s_2+\Delta s)\left(\frac{\Delta c}{\Delta s}-c'(s_1)\right) \qquad \text{(D34)} \\
& +s_2(3s_2+\Delta s)\left(2\frac{c(s_1)}{s_1}-\overline{\theta}-c'(s_1)\right) \\
& +s_2(3s_2+\Delta s)\left(2\overline{\theta}-\frac{c(s_2)}{s_2}-c'(s_1)\right) \\
& -3s_1s_2\left(2\overline{\theta}-\frac{c(s_2)}{s_2}-c'(s_1)\right) + 3s_1s_2\left(\frac{\Delta c}{\Delta s}-c'(s_1)\right)\Bigg] \\
= \; & \frac{s_2}{s_1}\frac{\alpha}{(3s_2+\Delta s)^2}\Bigg[\left(2\Delta s(3s_2+\Delta s)+3s_1s_2\right)\left(\frac{\Delta c}{\Delta s}-c'(s_1)\right) \qquad \text{(D35)} \\
& +s_2(3s_2+\Delta s)\left(2\frac{c(s_1)}{s_1}-\overline{\theta}-c'(s_1)\right) \\
& +\underbrace{\left(3s_2^2+s_2\Delta s-3s_1s_2\right)}_{=4s_2\Delta s}\left(2\overline{\theta}-\frac{c(s_2)}{s_2}-c'(s_1)\right)\Bigg].
\end{aligned}
$$

For firm 2 the proof follows analogous steps but now

$$
\begin{aligned}
\frac{\partial \Pi_2(s_1,s_2)}{\partial s_2} \; &= \; \beta^2 + 2\beta\Delta s\frac{\partial\beta(s_1,s_2)}{\partial s_2} \qquad \text{(D36)} \\
&\overset{\text{Def }\beta}{=} \; \frac{\beta}{(3s_2+\Delta s)^2}\Bigg[(3s_2+\Delta s)s_2\left(2\overline{\theta}-\frac{c(s_2)}{s_2}-\frac{\Delta c}{\Delta s}\right) \\
& \qquad +2\Delta s\frac{\partial\beta(s_1,s_2)}{\partial s_2}(3s_2+\Delta s)^2\Bigg],
\end{aligned}
$$

the two versions of β read

$$
\begin{aligned}
\beta(s_1,s_2) &:= \frac{1}{3s_2+\Delta s}\left(2s_2\overline{\theta} - c(s_2) - s_2\frac{c(s_2)-c(s_1)}{\Delta s}\right) \qquad \text{(D37)} \\
&= \frac{1}{3s_2+\Delta s}\left(2s_2\overline{\theta} + \frac{s_2}{\Delta s}c(s_1) - \frac{s_2+\Delta s}{\Delta s}c(s_2)\right), \qquad \text{(D38)}
\end{aligned}
$$

for the derivative of β w.r.t. s_2 we have

$$\begin{aligned}
\frac{\partial\beta(s_1,s_2)}{\partial s_2}(3s_2+\Delta s)^2 &= \left[2\overline{\theta}-\frac{s_1}{(\Delta s)^2}c(s_1)+\frac{s_1}{(\Delta s)^2}c(s_2)-\frac{s_2+\Delta s}{\Delta s}c'(s_2)\right](3s_2+\Delta s)\\
&\quad -4\left(2\overline{\theta}s_2-c(s_2)-s_2\frac{\Delta c}{\Delta s}\right)\\
&= -\frac{(s_2+\Delta s)(3s_2+\Delta s)}{\Delta s}c'(s_2)+\underbrace{\frac{s_1(3s_2+\Delta s)+4s_2\Delta s}{\Delta s}}_{=\frac{(s_2+\Delta s)(3s_2-\Delta s)}{\Delta s}}\left(\frac{\Delta c}{\Delta s}\right)\\
&\quad -2\overline{\theta}s_1+4c(s_2)\\
&= \frac{(s_2+\Delta s)(3s_2+\Delta s)}{\Delta s}\left(\frac{\Delta c}{\Delta s}-c'(s_2)\right)-2\frac{(s_2+\Delta s)\Delta s}{\Delta s}\left(\frac{\Delta c}{\Delta s}\right)\\
&\quad -2\overline{\theta}s_1+4c(s_2)\\
&= \frac{(s_2+\Delta s)(3s_2+\Delta s)}{\Delta s}\left(\frac{\Delta c}{\Delta s}-c'(s_2)\right)\\
&\quad -2s_1\left(\overline{\theta}+\frac{\Delta c}{\Delta s}\right)\underbrace{+2(2s_1)\frac{\Delta c}{\Delta s}-4s_2\frac{\Delta c}{\Delta s}+4c(s_2)}_{=-4\Delta c+4c(s_2)=4c(s_1)}\\
&= \frac{(s_2+\Delta s)(3s_2+\Delta s)}{\Delta s}\left(\frac{\Delta c}{\Delta s}-c'(s_2)\right)+2s_1\left(2\frac{c(s_1)}{s_1}-\overline{\theta}-\frac{\Delta c}{\Delta s}\right)
\end{aligned}$$

and together with (D37) this yields

$$\begin{aligned}
\frac{\partial\Pi_2(s_1,s_2)}{\partial s_2} &= \frac{\beta}{(3s_2+\Delta s)^2}\Bigg[(3s_2+\Delta s)s_2\left(2\overline{\theta}-\frac{c(s_2)}{s_2}-\frac{\Delta c}{\Delta s}\right) && \text{(D39)}\\
&\quad -2(s_2+\Delta s)(3s_2+\Delta s)\left(c'(s_2)-\frac{\Delta c}{\Delta s}\right)+(2\Delta s)2s_1\left(2\frac{c(s_1)}{s_1}-\overline{\theta}-\frac{\Delta c}{\Delta s}\right)\Bigg]\\
&= \frac{\beta}{(3s_2+\Delta s)^2}\Bigg[(3s_2+\Delta s)s_2\left(2\overline{\theta}-\frac{c(s_2)}{s_2}-c'(s_2)\right) && \text{(D40)}\\
&\quad -s_2(3s_2+\Delta s)\frac{\Delta c}{\Delta s}+s_2(3s_2+\Delta s)c'(s_2)\\
&\quad +(2s_2+2\Delta s)(3s_2+\Delta s)\left(\frac{\Delta c}{\Delta s}-c'(s_2)\right)+4s_1\Delta s\left(2\frac{c(s_1)}{s_1}-\overline{\theta}-\frac{\Delta c}{\Delta s}\right)\Bigg]\\
&= \frac{\beta}{(3s_2+\Delta s)^2}\Bigg[(3s_2+\Delta s)s_2\left(2\overline{\theta}-\frac{c(s_2)}{s_2}-c'(s_2)\right) && \text{(D41)}\\
&\quad +(3s_2+\Delta s)(s_2+2\Delta s)\left(\frac{\Delta c}{\Delta s}-c'(s_2)\right)+4s_1\Delta s\left(2\frac{c(s_1)}{s_1}-\overline{\theta}-\frac{\Delta c}{\Delta s}\right)\Bigg].
\end{aligned}$$

For the limits in part b) note that for firm 2, if the limit exists, (D22) implies

$$\lim_{s_1,s_2\to s_0}\frac{\partial\Pi_1(s_1,s_2)}{\partial s_1}=-\lim_{s_1,s_2\to s_0}\alpha(s_1,s_2)^2+\lim_{s_1,s_2\to s_0}2\alpha\Delta s\frac{\partial\alpha(s_1,s_2)}{\partial s_1}, \qquad \text{(D42)}$$

with

$$\begin{aligned}
\lim_{s_1,s_2\to s_0}\alpha(s_1,s_2) &= \frac{1}{3}\left(c'(s_0)-2\frac{c(s_0)}{s_0}+\overline{\theta}\right)=:K_1\\
\lim_{s_1,s_2\to s_0}\frac{\partial\alpha(s_1,s_2)}{\partial s_1} &= \lim_{s_1,s_2\to s_0}\Bigg[\underbrace{\frac{(s_2+\Delta s)(3s_2+\Delta s)}{\Delta s}\left(\frac{\Delta c}{\Delta s}-c'(s_1)\right)}_{\to 3s_2^2\cdot 0}+2s_2\underbrace{\left(2\overline{\theta}-\frac{c(s_2)}{s_2}-\frac{\Delta c}{\Delta s}\right)}_{\to 2\overline{\theta}-\frac{c(s_0)}{s_0}-c'(s_0)}\Bigg]
\end{aligned}$$

$$= 2s_0 \left(2\overline{\theta} - \frac{c(s_0)}{s_0} - c'(s_0)\right) =: K_2.$$

Plugged into (D42) this yields

$$\lim_{s_1,s_2 \to s_0} \frac{\partial \Pi_1(s_1,s_2)}{\partial s_1} = -K_1^2 + 2K_1K_2 \lim_{s_1,s_2 \to s_0} \Delta s = -K_1^2. \tag{D43}$$

Analogously for firm 2 we know from (D24) that, if the limit exists,

$$\lim_{s_1,s_2 \to s_0} \frac{\partial \Pi_2(s_1,s_2)}{\partial s_2} = \lim_{s_1,s_2 \to s_0} \beta(s_1,s_2)^2 + \lim_{s_1,s_2 \to s_0} 2\beta \Delta s \frac{\partial \beta(s_1,s_2)}{\partial s_2}, \tag{D44}$$

with

$$\lim_{s_1,s_2 \to s_0} \beta(s_1,s_2) = \frac{1}{3}\left(2\overline{\theta} - \frac{c(s_0)}{s_0} - c'(s_0)\right) =: K_3,$$

$$\lim_{s_1,s_2 \to s_0} \frac{\partial \beta(s_1,s_2)}{\partial s_2} = \lim_{s_1,s_2 \to s_0} \Bigg[\underbrace{\frac{(s_2+\Delta s)(3s_2+\Delta s)}{\Delta s}\left(\frac{\Delta c}{\Delta s} - c'(s_2)\right)}_{\to 3s_2^2 \cdot 0} + 2s_1 \underbrace{\left(2\frac{c(s_1)}{s_1} - \overline{\theta} - \frac{\Delta c}{\Delta s}\right)}_{\to 2\frac{c(s_0)}{s_0} - \overline{\theta} - c'(s_0)} \Bigg]$$

$$= 2s_0 \left(2\frac{c(s_0)}{s_0} - \overline{\theta} - c'(s_0)\right) =: K_4.$$

Plugged into (D44) this yields

$$\lim_{s_1,s_2 \to s_0} \frac{\partial \Pi_2(s_1,s_2)}{\partial s_2} = K_3^2 + 2K_3K_4 \lim_{s_1,s_2 \to s_0} \Delta s = K_3^2, \tag{D45}$$

which concludes the proof.

□

D3 Proof of Proposition 23

The proposition is a direct consequence of the following claim.

Claim. $\partial \Pi_1 / \partial s_1$ can be bounded from below as follows

$$\frac{\partial \Pi_1(s_1,s_2)}{\partial s_1} \geq \underbrace{\frac{s_2}{s_1}\frac{\alpha}{(3s_2+\Delta s)^2}}_{>0} \Bigg[\underbrace{(3s_2+\Delta s)}_{>0} K + \underbrace{s_1 \Delta s \left(2\overline{\theta} - \frac{c(s_2)}{s_2} - \frac{\Delta c}{\Delta s}\right)}_{>0} \Bigg] \tag{D46}$$

with K as defined in the proposition.

Proof of claim. For the lower bound of $\partial \Pi_1 / \partial s_1$, we start with (D33) to obtain

$$\begin{aligned}
\frac{\partial \Pi_1(s_1,s_2)}{\partial s_1} \overset{\text{(D33)}}{=} & \frac{s_2}{s_1}\frac{\alpha}{(3s_2+\Delta s)^2} \Bigg[2\Delta s(3s_2+\Delta s)\left(\frac{\Delta c}{\Delta s} - c'(s_1)\right) \\
& + s_2(3s_2+\Delta s)\left(2\frac{c(s_1)}{s_1} - \overline{\theta} - c'(s_1)\right) \\
& + s_2(3s_2+\Delta s)\left(2\overline{\theta} - \frac{c(s_2)}{s_2} - c'(s_1)\right) - 3s_1s_2\left(2\overline{\theta} - \frac{c(s_2)}{s_2} - \frac{\Delta c}{\Delta s}\right) \Bigg] \\
= & \frac{s_2}{s_1}\frac{\alpha}{(3s_2+\Delta s)^2} \Bigg[2\Delta s(3s_2+\Delta s)\left(\frac{\Delta c}{\Delta s} - c'(s_1)\right)
\end{aligned} \tag{D47}$$

$$
\begin{aligned}
& +s_2(3s_2+\Delta s)\left(2\frac{c(s_1)}{s_1}-\overline{\theta}-c'(s_1)\right) \\
& +s_2(3s_2+\Delta s)\left(2\overline{\theta}-\frac{c(s_2)}{s_2}-c'(s_1)\right) \\
& -(3s_2+\Delta s)\left(2\overline{\theta}-\frac{c(s_2)}{s_2}-\frac{\Delta c}{\Delta s}\right)+s_1\Delta s\left(2\overline{\theta}-\frac{c(s_2)}{s_2}-\frac{\Delta c}{\Delta s}\right)\Bigg] \\
= \; & \frac{s_2}{s_1}\frac{\alpha}{(3s_2+\Delta s)^2}\left[(3s_2+\Delta s)K+s_1\Delta s\left(2\overline{\theta}-\frac{c(s_2)}{s_2}-\frac{\Delta c}{\Delta s}\right)\right] \qquad \text{(D48)}
\end{aligned}
$$

with

$$
\begin{aligned}
K \; = \; & \overline{\theta}(s_2-2s_1)+c(s_2)\left(\frac{s_1}{s_2}-1\right)+2\frac{s_2}{s_1}c(s_1)-2(s_2+\Delta s)c'(s_1)+\Delta c\left(2+\frac{s_1}{\Delta s}\right) \qquad \text{(D49)} \\
= \; & \overline{\theta}(s_2-2s_1)-2(s_2+\Delta s)c'(s_1) \qquad \text{(D50)} \\
& +\frac{1}{s_2\Delta s}\left[c(s_2)\left(s_2^2-s_1^2+s_1s_2\right)+c(s_1)\left(2\frac{s_2^3}{s_1}-4s_2^2+s_1s_2\right)\right] \\
= \; & \overline{\theta}(s_2-2s_1)-2(s_2+\Delta s)c'(s_1) \qquad \text{(D51)} \\
& +\frac{1}{s_2\Delta s}\Bigg[s_2^2\underbrace{\left(c(s_2)+2c(s_1)\left(\frac{s_2}{s_1}-2\right)\right)}_{\geq c(s_1)\left[1+2\frac{s_2}{s_1}-4\right]=c(s_1)\left[\frac{2\Delta s-s_1}{s_1}\right]}+c(s_2)s_1\Delta s+c(s_1)s_1s_2\Bigg] \\
\geq \; & \overline{\theta}(s_2-2s_1)-2(s_2+\Delta s)c'(s_1) \qquad \text{(D52)} \\
& +\frac{1}{s_2\Delta s}\left[c(s_1)\frac{s_2^2}{s_1}(2\Delta s)+c(s_2)s_1\Delta s+c(s_1)(-s_2\Delta s)\right] \\
= \; & \overline{\theta}(s_2-2s_1)+2s_2\left(\frac{c(s_1)}{s_1}-c'(s_1)\right)-2c'(s_1)\Delta s+\frac{1}{s_2}\underbrace{(c(s_2)s_1-c(s_1)s_2)}_{=c(s_2)s_1-c(s_1)s_1+c(s_1)s_1-c(s_1)s_2} \qquad \text{(D53)} \\
= \; & \overline{\theta}(s_2-2s_1)+2s_2\left(\frac{c(s_1)}{s_1}-c'(s_1)\right)-2c'(s_1)\Delta s+\frac{s_1\Delta s}{s_2}\left(\frac{\Delta c}{\Delta s}-\frac{c(s_1)}{s_1}\right) \qquad \text{(D54)} \\
= \; & \overline{\theta}(s_2-2s_1)+\left(2s_2-\Delta s\frac{s_1}{s_2}\right)\left[\frac{c(s_1)}{s_1}-c'(s_1)\right]+\Delta s\frac{s_1}{s_2}\left[\frac{\Delta c}{\Delta s}-c'(s_1)\right]+2\Delta s\left(-c'(s_1)\right).
\end{aligned}
$$

□

D4 Proof of Proposition 24

It remains to derive the profitable deviation $s_3 = (s_2^2 + s_1 s_2 - s_1^2)/s_2$. To that end, let s_3 be some quality choice with $s_2 < s_3 \leq \overline{s}$. Then with $\Delta_{ij}s := (s_j - s_i)$ and $\Delta_{ij}c := c(s_j) - c(s_i)$ the following inequalities are equivalent

$$
\begin{aligned}
\Pi_2(s_2,s_3) \; &> \; \Pi_1(s_1,s_2) \qquad \text{(D55)} \\
\Leftrightarrow \Delta_{23}s\,\beta(s_2,s_3)^2 \; &> \; \Delta_{12}s\frac{s_2}{s_1}\alpha(s_1,s_2)^2 \\
\Leftrightarrow \Delta_{23}s\left[\frac{s_3}{3s_3+\Delta_{23}s}\left(2\overline{\theta}-\frac{c(s_3)}{s_3}-\frac{\Delta_{23}c}{\Delta_{23}s}\right)\right]^2 \; &> \; \Delta_{12}s\frac{s_2}{s_1}\left[\frac{s_1}{3s_2+\Delta_{12}s}\left(\frac{\Delta_{12}c}{\Delta_{12}s}-2\frac{c(s_1)}{s_1}+\overline{\theta}\right)\right]^2 \\
\Leftrightarrow \left(\frac{\Delta_{23}s}{\Delta_{12}s}\right)\left(\frac{s_1}{s_2}\right)\left(\frac{s_3^2}{s_1^2}\right)\frac{(3s_2+\Delta_{12}s)^2}{(3s_3+\Delta_{23}s)^2} \; &> \; \left(\frac{\frac{c(s_2)-c(s_1)}{\Delta_{12}s}-2\frac{c(s_1)}{s_1}+\overline{\theta}}{2\overline{\theta}-\frac{c(s_3)}{s_3}-\frac{c(s_3)-c(s_2)}{\Delta_{23}s}}\right)^2 \qquad \text{(D56)}
\end{aligned}
$$

with α and β for $s_i < s_j$ as defined in (D20) and (D21) at the beginning of Appendix D2. Suppose we can choose s_3 in the admissible interval such that

$$\frac{(s_3 - s_2)}{(s_2 - s_1)} = \frac{s_1}{s_2} \tag{D57}$$

$$\Leftrightarrow s_3 = \frac{s_2^2 + s_1 s_2 - s_1^2}{s_2}. \tag{D58}$$

With $s_1/s_2 \leq s_3/s_2$ and this particular choice of s_3 we have

$$(s_3 - s_2) \leq \frac{s_3}{s_2}(s_2 - s_1), \tag{D59}$$

which implies

$$\frac{3s_2 + (s_2 - s_1)}{3s_3 + (s_3 - s_2)} \geq \frac{s_2}{s_3}. \tag{D60}$$

Hence, for this specific choice of s_3, the LHS of (D56) reads

$$\left(\frac{s_1}{s_2}\right)\left(\frac{s_1}{s_2}\right)\left(\frac{s_3^2}{s_1^2}\right)\frac{(3s_2 + \Delta_{12}s)^2}{(3s_3 + \Delta_{23}s)^2} \geq \left(\frac{s_3^2}{s_2^2}\right)\left(\frac{s_2^2}{s_3^2}\right) = 1, \tag{D61}$$

while we know that the RHS of (D56) is smaller than 1 if and only if

$$2\overline{\theta} - \frac{c(s_3)}{s_3} - \frac{c(s_3) - c(s_2)}{(s_3 - s_2)} > \frac{c(s_2) - c(s_1)}{(s_2 - s_1)} - 2\frac{c(s_1)}{s_1} + \overline{\theta}. \tag{D62}$$

But (D62) holds, since from (B4') we know

$$2\overline{\theta} - \frac{c(s_3)}{s_3} - c'(s_3) > c'(s_3) - 2\frac{c(s_1)}{s_1} + \overline{\theta}$$

$$\Leftrightarrow 2\overline{\theta} - \frac{c(s_3)}{s_3} - \frac{\Delta_{23}c}{(s_3 - s_2)} + 2\frac{c(s_1)}{s_1} - \overline{\theta} - \frac{\Delta_{12}c}{(s_2 - s_1)} > \underbrace{c'(s_3) - \frac{\Delta_{23}c}{(s_3 - s_2)}}_{>0} + \underbrace{c'(s_3) - \frac{\Delta_{12}c}{(s_2 - s_1)}}_{>0}.$$

Hence, (D56) holds and this particular choice of s_3 is in fact a profitable deviation. When is this choice of s_3 infeasible? Suppose $s_2 < \overline{s}$. For $s_3 = s_2$ the LSH of (D57) is zero. As s_3 increases, the expression on the LHS increases. Hence, either (D57) holds for some s_3 - in which case we have found a profitable deviation - or $(\overline{s} - s_2) < s_1/s_2(s_2 - s_1)$. This deviation is infeasible if

$$(\overline{s} - s_2) < \frac{s_1}{s_2}(s_2 - s_1) = \frac{s_1}{s_2}\left(1 - \frac{s_1}{s_2}\right)s_2, \tag{D63}$$

which, since the RHS is smaller equal than $s_2/4$, holds if

$$\frac{4}{5}\overline{s} < s_2. \tag{D64}$$

□