Partial Cross Ownership and Innovation*

Sandro Shelegia and Yossi Spiegel†

May 24, 2023

Abstract

We study the effects of partial cross ownership (PCO) among rival firms on their incentives to innovate. PCO in our model gives rise to a *price effect* which encourages investment by softening price competition, but also to a *cannibalization effect* which discourages investment because each firm internalizes part of the negative externality of its investment on the rival's profit. We show that overall, PCO may benefit or harm consumers depending on the size of the PCO stakes and their degree of symmetry, the relative cost of the innovation and its size, and whether it is drastic or not.

JEL Classification: D43, L41

Keywords: Bertrand oligopoly, innovation, partial cross ownership, drastic and non-drastic innovation

*For helpful discussions and comments we thank John Asker, Patrick Rey, and seminar participants at the 2014 International Workshop on Evolution of Standards and Technology in Hitotsubashi University, the 2015 MaCCI conference in Mannaheim, the 2017 IIOC conference in Boston, and the 2022 Berlin IO day.

†Shelegia: Department of Economics and Business, Universitat Pompeu Fabra, BSE and CEPR, Sandro.Shelegia@upf.edu. Spiegel: Coller School of Management, Tel Aviv University, CEPR, and ZEW, spiegel@post.tau.ac.il

1 Introduction

Many industries feature a complex web of partial cross ownership (PCO) among rival firms. Examples include the Japanese and the U.S. automobile industries (Alley, 1997 and Ono et al. 2004), the Dutch Financial Sector (Dietzenbacher, Smid, and Volkerink, 2000), the Nordic power market (Amundsen and Bergman, 2002), and the global steel industry (Gilo, Moshe, and Spiegel, 2006). Nitta (2008) reports that cross-shareholding, i.e., situations in which two public companies mutually own each other's shares, accounted for $13\% - 15\%$ of the shares of public firms listed in the Tokyo, Osaka, and Nagoya stock exchanges during the 1990's, and remained above 8.5% by 2006.[1]

While horizontal mergers are subject to substantial antitrust scrutiny, passive investments in rival firms were either granted a de facto exemption from antitrust liability, or have gone unchallenged by antitrust agencies in recent cases (Gilo, 2000). Rock and Rubinfeld (2018) argue that the DOJ and the FTC generally have not challenged partial equity acquisitions of less than 20% with no evidence of control. This lenient approach is due to the courts' interpretation of the exemption for stock acquisitions "solely for investment" included in Section 7 of the Clayton Act and the fact that acquisitions of PCO stakes do not involve a conspiracy in restraint of trade and hence cannot be condemned under Section 1 of the Sherman Act (Rock and Rubinfeld, 2018).

Recently though, the European Commission has began to question the lenient approach towards passive investments (e.g., European Commission, 2013) and stated that "significant harm to competition and consumers can occur not only from acquisitions of control, but also from structural links."[2] Indeed, the literature has shown that PCO among rival firms can soften competition in the Cournot model (Reynolds and Snapp, 1986; Bolle and Güth, 1992; Flath 1991 and 1992;

[1] A related, but distinct, phenomenon is common ownership: cases where the same set of shareholders own several competing firms. Common ownership has attracted a lot of attention recently and there is a lively debate about its competitive implications. See for instance, Azar, Schmalz, and Tecu (2018), Antón et al (2022), Backus, Conlon, and Sinkinson (2021a, 2021b), Banal-Estañol, Seldeslachts, and Vives, (2020). While common and cross ownership are in general not the same (see the discussion below), they are isomorphic in our duopoly setting. However, for the sake of concretness, we will refer to links bewteen firms as cross ownership.

[2] The commission also mentioned common shareholding theory of harm in two recent merger reviews (Dow/DuPont in 2017 and Bayer/Monsanto in 2018), albeit it did not formally rely on this theory of harm in its final decsions. See Burnside and Kidane (2020). Interestingly, the commission approved the two mergers subject to divestitures of major businsses and assets, including R&D organisations, and argued in its Dow/DuPont decision that "the presence of significant common shareholding is likely to negatively affect the benefits of innovation competition for firms subject to this common shareholding." See European Commission case M.7932 – Dow/DuPont, Paragraph 2352 and case M. 8084 - Bayer/Monsanto.

Reitman, 1994; and Dietzenbacher, Smid, and Volkerink, 2000) or the Bertrand model (Shelegia and Spiegel, 2012), and can also facilitate collusion in a repeated Cournot model (Malueg, 1992) or the Bertrand model (Gilo, Moshe, and Spiegel, 2006).[3] Moreover, partial cross ownership (with and without control) among upstream and downstream firms can also lead to upstream and downstream foreclosure (e.g., Baumol and Ordover, 1994; Reiffen, 1998; Greenlee and Raskovich, 2006; Spiegel, 2013;, Hunold and Stahl, 2016; and Levy, Spiegel, and Gilo, 2018). The picture that emerges from existing literature is that PCO has adverse unilateral and coordinated competitive effects. This picture has received some empirical support (e.g., Dietzenbacher, Smid, and Volkerink, 2000; Brito, Ribeiro, and Vasconcelos, 2014; Nain and Wang, 2018; and Heim et al., 2022).

In this paper we show that PCO can also have a bright side: under certain conditions, PCO may promote innovation and benefit consumers. Specifically, we consider a Bertrand duopoly in which firms hold PCO stakes in each other and choose how much to invest before setting prices. Investments in our model can be in either process or product innovation and they either succeed or fail. A firm earns a positive profit only if its investment succeeds and the rival's investment fails. Otherwise, the firm earns a profit of 0 due to Bertrand competition. Importantly, the PCO stakes that firms hold in each other need not be identical. We are interested in finding how an increase in the stake that one firm holds in the rival affects investments and prices and ultimately consumer surplus, which is the most common welfare standard used by antitrust agencies.[4]

We show that the PCO stakes soften price competition when one firm innovates successfully and the other fails.[5] The resulting effect, which we term *the price effect* of PCO, encourages firms to invest in an attempt to be the sole innovator. At the same time, when a firm fails to innovate but the rival does, the firm still makes a profit due to its stake in the rival. Since investment cannibalizes the rival's profit, PCO weakens the incentive to invest. We term this the *cannibalization effect* of PCO. We then explore how the two conflicting effects play out in equilibrium and explore the implications for consumers' welfare. In particular, we assume that firm i's stake in firm j, α_i, is possibly larger than firm j's stake in firm i, α_j, and establish sufficient conditions for an increase in

[3] Malueg (1992) shows that PCO can also hinder collusion in a repeated Cournot model, but when it does, firms should have no incentives to acquire ownership stakes in one another in the first place.

[4] For instance, it is used in the U.S., the EU, and the UK (see OECD, 2012, p. 26-27).

[5] The reason is that in our Bertrand setting, the equilibrium price equals marginal cost, which is 0 if both firms succeed to innovate or c if both firms fail. Absent PCO, if only firm i succeeds, it continues to charge c (firm j's marginal cost) and (by a tie-breaking rule) serves the entire market. However, when firm j holds a stake in firm i, it is better off allowing firm i to serve the entire market (at 0 cost) and sharing firm i's profits than undercutting firm i and serving the market itself (at a cost of c). As a result, firm i can raise its price in equilibrium above c.

α_i (which expands the gap between α_i and α_j and makes the PCO structure more asymmetric) to harm consumers, and a sufficient condition for an increase in α_j (which makes the PCO structure more symmetric) to benefit consumers.

We then consider the case where firms face a unit demand function and explore the effect of PCO on consumers in more detail. Absent PCO, the innovation in our model is non-drastic: a sole innovator cannot act as an unconstrained monopolist. However, due to the price effect of PCO, large enough PCO stakes make the innovation drastic. That is, whether the innovation is drastic or not depends directly on the size of the PCO stakes.

When the PCO stakes are relatively small so the innovation is non-drastic, an increase in α_i unambiguously harms consumers. By contrast, an increase in α_j benefits consumers when the relative cost of innovation (the marginal cost of innovation divided by the size of the innovation) is low, but harms consumers otherwise. We also show that an increase in α_j benefits consumers for a larger set of parameters as the PCO structure becomes more asymmetric (the gap between α_i and α_j expands).

When the PCO stakes are sufficiently large to make the innovation drastic, a sole innovator already charges the monopoly price, so a further increase in the PCO stakes does not lead to a price effect. An increase in α_i or α_j can benefit consumers in this case if the relative cost of innovation (the marginal cost of innovation divided by the willingness of consumers to pay) is sufficiently high or the size of the innovation (relative to the willingness to pay) is sufficiently small. When the size of the innovation is large, an increase in α_j can still benefit consumers if the relative cost of innovation is small. By contrast, when the size of the innovation is large, an increase in α_i harms consumers and an increase in α_j also harms consumers if in addition the relative cost of innovation is high (but not too high). We also show that in the neighborhood of a symmetric PCO structure, an increase in α_i or α_j harms consumers when the relative cost of innovation is small.

Our analysis highlights the fact that PCO may benefit consumers by promoting innovation. The reason for this is that PCO softens price competition when only one firm innovates, which in turn boosts the incentive to invest and become the sole innovator. One may then wonder how PCO performs relative to other arrangements which are intended to boost investments by softening competition, like outright collusion in the product market (semicollusion), a research joint venture (RJV), or a full merger. We show that compared to a semicollusion, symmetric PCO with a sufficiently high stakes boosts investment and benefits consumers. Symmetric PCO also boosts investments more than a full merger, provided that the PCO stakes are not too large, and benefits consumers. Compared with an RJV, symmetric PCO leads to a higher investment, and at least in the unit demand case, it benefits consumers more than RJV if the relative cost of innovation is sufficiently small.

The rest of the paper is organized as follows. In Section 2 we review related literature and in Section 3 we present the model and characterize the equilibrium absent PCO. In Section 4 we characterize the equilibrium with PCO and study the welfare implications of PCO in Section 5. In Section 6 we consider the unit demand case in order to shed more light on the welfare implications of PCO. In Section 7 we compare PCO with semicollusion, RJV's, and full mergers. In Section 8 we conclude. All proofs are in the Appendix.

2 Related literature

Our paper is related to the small literature that studies the effects of overlapping ownership (both cross and common ownership) on innovation. Ideally, papers in this literature should allow for a general ownership structure, a general R&D process, and a general model of product market competition. Given the difficulty of deriving results with a very general model, papers in this literature have made progress by simplifying some of these aspects.

López and Vives (2019) consider a fairly general n-firm Cournot oligopoly model, but assume a deterministic cost-reducing R&D process and a symmetric overlapping ownership structure. They show that if demand is not too convex, an increase in the symmetric level of overlapping ownership increases investments and output when R&D spillovers are sufficiently high, increases investments and decreases output when R&D spillovers are intermediate, and decreases investments and output when R&D spillovers are low.[6]

Stenbacka and Van Moer (2022) consider a duopoly model with a stochastic product R&D process, but like López and Vives (2019) they also assume that the overlapping ownership structure is symmetric. They show that an increase in the symmetric levels of overlapping ownership can improve welfare even without R&D spillovers because it softens competition and therefore boosts the marginal benefit from investment. By contrast, the marginal benefit from investment in process innovation is proportional to output, so when competition is softer, firms have a weaker incentive to invest.

[6] These results are robust to a Bertrand model with differentiated products in which R&D levels are chosen before output levels. Also see Vives (2020) for an overview of the results.

Bayona and López (2018) consider a Hotelling duopoly model with possibly asymmetric common ownership, but consider a deterministic quality-enhancing R&D process.[7] They show that if the controlling shareholder of firm i holds a larger stake in firm j than the controlling shareholder of firm j holds in firm i, then firm i invests less and may also set a higher price than firm j.[8] Moreover, they show that common ownership may lead to higher or lower consumer and total surplus when only one controlling shareholder holds a stake in the rival, but symmetric common ownership always decreases consumer and total surplus.[9]

Antón et al. (2021) consider an n-firm Cournot oligopoly model with differentiated products and linear demand functions with possibly asymmetric common ownership, but consider a deterministic cost-reducing R&D process. They show that an increase in the weight that firm i assigns to firm j's profit increases firm i's R&D investment if and only if technological spillovers are sufficiently large relative to the degree of product differentiation. They provide empirical support for this result using data on publicly listed U.S. corporations.[10]

Our paper differs from the above papers in that we consider a Bertrand duopoly, but allow the ownership stakes to be asymmetric and consider a stochastic R&D process which could be viewed as either process or product innovation. In other words, we consider a fairly simple model of product market competition (under PCO, the Bertrand model is less simple than one may think), but consider a general ownership structure and an R&D process. Our modeling choice is motivated by the following consideration. First, models with symmetric ownership structure can, by design, only examine the competitive implications of an increase in the weights that all firms assign to the

[7] Although they allow common ownership to be asymmetric, they only study the welfare implications of either symmetric common ownership or common ownership in only one of the two firms.

[8] Specifically, they assume that investments deterministically increase the base utility that consumers receive and show that firm i sets a higher price than firm j if and only if the ratio of the marginal effect of investment on quality and the transportation cost is sufficiently low.

[9] Li and Zhang (2021) study a related model with Hotelling duopoly, where firms first choose locations (possibly outside the Hoetlling line) and then compete by setting prices. They show that an increase in the symmetric overlapping ownership level harms consumers because it induces firms to move further apart (outside the Hotelling line) and set higher prices. While there are no investments in quality in their model, the choice of locations outside the Hotelling line increases transportation costs and is akin to a decrease in quality.

[10] Specifically, they find that an increase in common ownership is associated with a decrease in citation-weighted patents when products are sufficiently close substitutes, but an increase in citation-weighted patents when technology spillovers are relatively large. Lewellen and Lowry (2021) use mergers of financial institutions outside the 2008-2009 period and find that although these mergers cause substantial and lasting increases in common ownership, there is no evidence that they affect firms' R&D spendings.

profits of all other firms by the exact same amount. By contrast, we can study the competitive implications of an increase in the stake that one firm holds in a rival, holding fixed the rival's PCO stake. This comparative statics exercise is policy relevant because, in practice, antitrust agencies evaluate acquisitions of ownership stakes one at a time.[11] Moreover, we show that the welfare effects of PCO depend, among other things, on how symmetric or asymmetric the PCO structure is. For example, when consumers have a unit demand function and the PCO stakes are symmetric and sufficiently low to ensure that the innovation is non drastic, an increase in a symmetric PCO stake always harms consumers, whereas an increase in the stake of only one firm can enhance welfare if the relative cost of innovation is sufficiently small.

Second, we consider a stochastic R&D process, which can either succeed or fail, rather than a deterministic R&D process as in López and Vives (2019), Bayona and López (2018), and Antón et al. (2021). This distinction is important because a deterministic R&D process leads to lower costs (or higher quality) and unambiguously benefits consumers. By contrast, an increase in R&D investments in our model is a double-edge sword from consumers' point of view: although it increases the likelihood that both firms innovate, which benefits consumers, it may also increase the likelihood that only one firm innovates, which harms consumers due to the price effect of PCO.

Our paper is also related to the literature that studies the effects of horizontal mergers on investments in innovation, (e.g., Federico, Langus, and Valletti, 2018; Jullien and Lefouili, 2018; and Motta and Tarantino, 2021). PCO can be viewed as a "partial merger," in which firms remain independent entities but still internalize part of their externality on rivals.

3 Model

Two firms produce a homogeneous good at a constant marginal cost, $c > 0$, and face a downward sloping demand, $Q(p)$. The strategic interaction between the two firms evolves in two stages. In stage 1, each firm i decides how much to invest in an innovation which either succeeds with probability λ_i or fails with probability $1 - \lambda_i$. If the innovation succeeds, marginal cost drops to 0, and if it fails, marginal cost remains $c > 0$. The parameter c then reflects the size of the innovation.[12] We assume that λ_i is a choice variable for the firm and for the sake of concreteness,

[11] Rock and Rubinfeld (2018) provide examples from the U.S. Interestingly, all of these examples involve cross ownership rather than common ownership.

[12] Alternatively, we can normalize marginal cost to 0 and assume that if innovation succeeds, the willingness of consumers to pay shifts up by a constant c. That is, the innovation in our framework can be viewed as either process

refer to it as "firm i's investment level."[13] The cost of investment is $\frac{k\lambda_i^2}{2}$, where $k > 0$ is the slope of the marginal cost of investment.

In stage 2, the firms observe each other's marginal costs and simultaneously choose prices. Consumers buy from the lowest price firm; if both firms charge the same price, consumers buy from the more efficient firm.[14] If both firms are equally efficient, consumers randomize between them.

We now make a few assumptions about the $Q(p)$, c, and k.

A1 $\pi(p) = pQ(p)$ (the profit of a monopoly when marginal cost is 0 and price is p) is concave in p

A2 $\varepsilon'(p) \geq 0$, where $\varepsilon(p) \equiv -\frac{pQ'(p)}{Q(p)}$ is the elasticity of demand

A3 $c < p^m < 2c$, where $p^m \equiv \arg\max_p \pi(p)$ is the monopoly price when marginal cost is 0

A4 $k > \pi^m \equiv p^m Q(p^m)$

Assumptions A1 and A2 ensure that the demand function behaves "nicely." Assumption A3 implies that absent PCO, the innovation is non-drastic in the sense that a firm cannot act as an unconstrained monopolist when it innovates and the rival fails, but with PCO, the innovation becomes drastic for sufficiently large PCO stakes.[15] Assumption A4 ensures that the equilibrium choices of λ_i and λ_j are below 1 (recall that λ_i and λ_j are probabilities).

Although we use a Bertrand setting to model competition in stage 2, our qualitative results should continue to hold under a more general setting. For instance, Aoki and Spiegel (2009) consider a similar model with stochastic R&D process, but assume that in stage 2, the profit of each firm is π_{yy} if both firms innovate, π_{nn} if both firms fail, π_{yn} if the firm innovates and the rival fails, and π_{ny} if the firm fails but the rival innovates.[16] They show that qualitative results in the Bertrand model, in which $\pi_{yn} > 0 = \pi_{ny} = \pi_{yy} = \pi_{nn}$, continue to hold in the more general model, so long

or product innovation. While the two formulations are isomorphic, we will use the process innovation interpretation for the sake of concreteness.

[13] Strictly speaking though, λ_i is the probability that firm i innovates successfully.

[14] The latter assumption is standard (see e.g., Deneckere and Kovenock, 1996). If consumers are also strategic players, this is actually a result rather than an assumption, because if consumers buy from the less efficient firm when prices are the same, the more efficient firm can undercut the less efficient firm slightly. Hence, a Nash equilibrium exists only if consumers buy from the most efficient firm when both firms charge the same price.

[15] For example, when demand is linear and given by $Q = A - p$, the monopoly price is $A/2$, so Assumption A3 implies that $2c < A < 4c$.

[16] Jullien and Lefouili (2018) and Stenbacka and Van Moer (2022) consider a similar setting, where $\pi_{ny} = \pi_{nn} = 0$.

as $\pi_{yn} + \pi_{ny} > \pi_{yy} + \pi_{nn}$. We chose to work with the Bertrand setting because in general, PCO affects all stage 2 profits, π_{yn}, π_{ny}, π_{yy}, and π_{nn}, so the model becomes too complex to analyze, especially since we focus on asymmetric PCO structure and cannot invoke symmetry to simplify the analysis. The Bertrand setting has the advantage that only π_{yn} and π_{ny} are affected by cross ownership, while π_{yy} and π_{nn} remain equal to 0.

We end this section with a characterization of the equilibrium in the no PCO benchmark. When both firms innovate, their marginal cost is 0 and they charge a price of 0 in stage 2. When both firms fail to innovate, their marginal cost is c, and in equilibrium they charge c in stage 2. In both cases, the two firms earn 0 in stage 2. Given that the innovation is non-drastic, when firm i innovates and firm j fails, firm i serves the entire market at a price c.[17] The resulting equilibrium profit of firm i in stage 2 is $\pi(c) = cQ(c)$, while firm j's profit in stage 2 is 0. Hence, the expected profit of firm i in stage 1 is

$$\lambda_i(1-\lambda_j)\pi(c) - \frac{k\lambda_i^2}{2}. \tag{1}$$

In equilibrium, both firms choose

$$\lambda^* = \frac{\pi(c)}{k+\pi(c)}.$$

Given Assumption A4, $k > \pi^m > \pi(c)$, so the equilibrium is unique and stable.[18]

4 Equilibrium with PCO

Now suppose that firm i holds a partial cross ownership (PCO) stake, α_i in firm j and firm j holds a stake α_j in firm i, where $\alpha_j \leq \alpha_i < \frac{1}{2}$. These stakes are passive and give each firm a share in its rival's profit, but no control over the rival's decisions. Using Π_i and Π_j to denote the standalone profits of the two firms, their overall values, including their stakes in their rival, are defined by the following system:

$$V_i = \Pi_i + \alpha_i V_j, \qquad V_j = \Pi_j + \alpha_j V_i.$$

Solving the system, yields

$$V_i = \frac{\Pi_i + \alpha_i \Pi_j}{1-\alpha_i\alpha_j}, \qquad V_j = \frac{\Pi_j + \alpha_j \Pi_i}{1-\alpha_i\alpha_j}.$$

[17] If the innovation were drastic, firm i would choose the monopoly price, p^m.

[18] To see why, note that firm i's best-response function is $\lambda_i = \frac{(1-\lambda_j)\pi(c)}{k}$ and its slope in the (λ_i, λ_j) space is above 1 in absolute value, while firm j's best-response function is $\lambda_j = \frac{(1-\lambda_{ij})\pi(c)}{k}$ and its slope in the (λ_i, λ_j) space is below 1 in absolute value. Consequently, the best-response function of firm i crosses the best-response function of firm j once and from above in the interior of the (λ_i, λ_j) space.

Note that each firm assigns a larger weight to its own standalone profit than to the rival's standalone profit. Also note that although V_i and V_j sum up to more than $\Pi_i + \Pi_j$, the share of "real" shareholders (not firms) in these values is $(1-\alpha_j)V_i + (1-\alpha_i)V_j = \Pi_i + \Pi_j$.

The decisions of each firm i are made by its controlling shareholder, whose ownership stake is β_i, where $\beta_i + \alpha_j \leq 1$; the remaining stake, $1-\beta_i-\alpha_j$ is held by dispersed shareholders. We assume that the controlling shareholder of each firm does not hold a stake in the rival firm, so his objective is to maximize $\beta_i V_i$.[19] Since β_i is a constant, there is no loss of generality in assuming that the controller's objective is to simply maximize V_i.

As in the standard Bertrand model, when both firms innovate in stage 1 or both fail, competition drives their standalone profits in stage 2 to 0. To see why, suppose that in stage 2 firm i charges a price p. If firm j undercuts p, its value is $V_j = \frac{pQ(p)}{1-\alpha_i\alpha_j}$ if both firms innovate and $V_j = \frac{(p-c)Q(p)}{1-\alpha_i\alpha_j}$ if both firms fail to innovate. If firm j sets a price above p, firm i serves the entire market, so $V_j = \frac{\alpha_j pQ(p)}{1-\alpha_i\alpha_j}$ if both firms innovate and $V_j = \frac{\alpha_j(p-c)Q(p)}{1-\alpha_i\alpha_j}$ if both firms fail. Since $\alpha_j < \frac{1}{2}$, undercutting p is more profitable for firm j, so the usual Bertrand equilibrium prevails.

Things are more involved when firm i innovates in stage 1 and its marginal cost drops to 0, while firm j fails and its marginal cost remains c. Then, when firm i charges a price p, firm j can either undercut p slightly, in which case $V_j = \frac{(p-c)Q(p)}{1-\alpha_i\alpha_j}$, or can let firm i serve the entire market at p, in which case $V_j = \frac{\alpha_j pQ(p)}{1-\alpha_i\alpha_j}$. Firm j will not undercut firm i if

$$\frac{(p-c)\,Q(p)}{1-\alpha_i\alpha_j} \leq \frac{\alpha_j pQ(p)}{1-\alpha_i\alpha_j}, \qquad \Longrightarrow \qquad p \leq \frac{c}{1-\alpha_j}.$$

By assumption, when both firm charge the same price, consumers buy from the more efficient firm; hence firm i can charge $\frac{c}{1-\alpha_j}$ and serve the entire market. However, if $\frac{c}{1-\alpha_j} > p^m$, where p^m is the monopoly price when marginal cost is 0, the innovation becomes drastic and firm i is better off charging p^m. By Assumption A3, this can happen however only when $\alpha_j > 0$. The threshold of α_j

[19] When firm i's controlling shareholder holds a stake β_{ij} in firm j, we also have common ownership: the shareholder's objective then is to maximize $\beta_i V_i + \beta_{ij} V_j = (\beta_i + \alpha_j\beta_{ij})\left(\Pi_i + \frac{\alpha_i\beta_i+\beta_{ij}}{\beta_i+\alpha_j\beta_{ij}}\Pi_j\right)$. Although the weight assigned to firm j's profit now exceeds α_i, concenptually nothing changes, which is why we set $\beta_{ij} = \beta_{ji} = 0$. Moreover, under pure common ownership (when $\alpha_i = 0$), the objective of firm i's controlling shareholder becomes $\beta_i V_i + \beta_{ij} V_j = \beta_i\left(\Pi_i + \frac{\beta_{ij}}{\beta_i}\Pi_j\right)$, which is again equivalent to the objective function that we consider, with $\frac{\beta_{ij}}{\beta_i}$ replacing α_i. While common and cross ownership are essentially equivalent in our duopoly setting, the equivalence may no longer hold if there are more than two firms and the ownership stakes are asymmetric. Then, as Gilo, Moshe, and Spiegel (2006) show, an increase in firm 1's stake in firm 2, say, may affect the effective weight that firm 3 assigns to the profits of firms 1 or 2. This cannot arise under common ownership.

above which the innovation becomes drastic is then $\alpha_j = \frac{p^m - c}{p^m}$. In what follows we will denote this threshold by σ; that is, the innovation is drastic for firm i if $\alpha_j \geq \sigma \equiv \frac{p^m - c}{p^m}$.[20]

Proposition 1 in Shelegia and Spiegel (2012) implies that when firm i has a lower cost than firm j, there are multiple Nash equilibria in stage 2 of the game. In these equilibria, firm i serves the entire market and the two firms charge the same price $p \in [0, p(\alpha_j)]$, where

$$p(\alpha_j) \equiv \begin{cases} \frac{c}{1-\alpha_j}, & \alpha_j < \sigma, \\ p^m, & \alpha_j \geq \sigma. \end{cases} \tag{2}$$

The associated stage 2 profit of firm i as a function of firm j's stake, α_j, is

$$\pi(\alpha_j) \equiv \pi(p(\alpha_j)) = \begin{cases} \frac{c}{1-\alpha_j} Q\left(\frac{c}{1-\alpha_j}\right), & \alpha_j < \sigma, \\ \pi^m, & \alpha_j \geq \sigma. \end{cases} \tag{3}$$

Of the above Nash equilibria, the only equilibrium in which firm j does not play a weakly dominated strategy is the one where both firms charge $p(\alpha_j)$ and firm i serves the entire market.[21] In what follows, we will restrict attention to this equilibrium.

A few comments are now in order. First, $p'(\alpha_j) \geq 0$ and $\pi'(\alpha_j) \geq 0$: when firm i is the sole innovator, it charges the same or a higher price as α_j increases, and its associated profit in stage 2 is weakly higher. Intuitively, as α_j increases, firm j is more willing to let firm i serve the entire market and share its profit than undercut firm i and serve the market itself at a higher cost. This allows firm i to raise its price without being undercut by firm j. When firm j does not hold a stake in firm i, i.e., $\alpha_j = 0$, then $p(0) = c$ and $\pi(0) = 0$, exactly as in the traditional Bertrand model.

Second, the threshold PCO level above which the innovation becomes drastic, $\sigma \equiv \frac{p^m - c}{p^m}$, reflects the size of the innovation, c, relative to the monopoly price, p^m, with higher values of c being associated with lower values of σ. That is, σ is inversely related to the size of the innovation; by Assumption A3, $\sigma \to 0$ as $c \to p^m$, in which case the innovation is largest, and $\sigma \to 1/2$ as $p^m \to 2c$, in which case the innovation is smallest. Since $\sigma < 1/2$, PCO levels that are sufficiently close to $1/2$ make the innovation drastic.

[20] Although $\frac{p^m - c}{p^m}$ looks like a price-cost margin, in fact it is not because p^m is the monopoly price when marginal cost is 0.

[21] To see why, consider an equilibrium such that $p_i = p_j = p^* \in [0, p(\alpha_j))$. While firm j makes no sales in equilibrium (recall that by assumption, when firms charge the same price, consumers buy from the more efficient firm, which here is i as its cost is 0 while firm j's cost is c), its strategy is weakly dominated by charging $p(\alpha_j)$ since any upward deviation by firm i from p^* leaves firm j with a stage 2 profit of $(p^* - c)\, Q(p^*)$; this profit is below $\alpha_j p^* Q(p^*)$, which is firm j's profit when it charges $p(\alpha_j)$ since $p^* < p(\alpha_j)$. Hence, a trembling hand argument will elliminate all such equilibria.

Third, $\alpha_j \leq \alpha_i$ implies that $p(\alpha_j) \leq p(\alpha_i)$; since $p(\alpha_i) \leq p^m$, it follows from Assumption A1 that $\pi(\alpha_i) \geq \pi(\alpha_j)$. Moreover, $\pi'(\alpha_i) \geq 0$ and $\pi'(\alpha_j) \geq 0$: firm i's profit increases with the stake that firm j holds in firm i and conversely for firm j.

We summarize these observations in the next lemma.

Lemma 1: *The equilibrium in stage 2 is as follows:*

(i) *When both firms innovate in stage 1 or both fail, the equilibrium price in stage 2 is equal to their marginal cost and their equilibrium profits in stage 2 are 0.*

(ii) *When firm i innovates in stage 1, while firm j fails, the unique equilibrium in stage 2 in which firms do not play weakly dominated strategies is such that both firms charge $p(\alpha_j)$ and firm i serves the entire market and earns $\pi(\alpha_j)$. Both $p(\alpha_j)$ and $\pi(\alpha_j)$ are (weakly) increasing with α_j. Since $\alpha_i \geq \alpha_j$, $p(\alpha_i) \geq p(\alpha_j)$ and $\pi(\alpha_i) \geq \pi(\alpha_j)$.*

Moving to stage 1 in which firms make investment decisions, note that with probability $\lambda_i(1-\lambda_j)$, firm i innovates and firm j fails, so firm i's stage 2 profit is $\pi(\alpha_j)$; with probability $\lambda_j(1-\lambda_i)$, firm j innovates and firm i fails, so firm j's stage 2 profit is $\pi(\alpha_j)$. The expected value of firm i when it chooses λ_i in stage 1 is therefore

$$V_i = \frac{\overbrace{\lambda_i(1-\lambda_j)\pi(\alpha_j) - \frac{k\lambda_i^2}{2}}^{\Pi_i} + \alpha_i \overbrace{\left(\lambda_j(1-\lambda_i)\pi(\alpha_i) - \frac{k\lambda_j^2}{2}\right)}^{\Pi_j}}{1-\alpha_i\alpha_j}.$$

The resulting best-response function of firm i against firm j is given by

$$BR_i(\lambda_j) = \begin{cases} 0 & \lambda_j > \frac{\pi(\alpha_j)}{\pi(\alpha_j)+\alpha_i\pi(\alpha_i)}, \\ (1-\lambda_j)\frac{\pi(\alpha_j)}{k} - \lambda_j\frac{\alpha_i\pi(\alpha_i)}{k} & \lambda_j \leq \frac{\pi(\alpha_j)}{\pi(\alpha_j)+\alpha_i\pi(\alpha_i)}. \end{cases} \tag{4}$$

The best-response function of firm j against firm i is analogous.

Notice that $BR_i'(\lambda_j) \leq 0$ and $BR_j'(\lambda_i) \leq 0$, implying that the choices of λ_i and λ_j are strategic substitutes: firm i invests less when firm j invests more. Intuitively, firm j's investment lowers firm i's chance to be the sole innovator, which is the only situation in which firm i makes money in period 2. Hence, a larger λ_j weakens firm i's incentive to invest. When $\lambda_j > \frac{\pi(\alpha_j)}{\pi(\alpha_j)+\alpha_i\pi(\alpha_i)}$, the marginal benefit of firm i from investing is below the associated cost, so firm i does not invest.

Also notice that by Lemma 1, $\pi'(\alpha_j) \geq 0$ and $\pi_j'(\alpha_i) \geq 0$, so $BR_i(\lambda_j)$ is increasing with α_j and decreasing with α_i. Consequently, PCO has two opposing effects on the incentive to invest, which we will refer to as the "price effect" and the "cannibalization effect." The price effect of PCO, given by $(1-\lambda_j)\frac{\pi(\alpha_j)}{k}$, reflects the extra profit that firm i makes in the event that it innovates and firm j fails. An increase in α_j boosts this term because a higher α_j allows firm i to set a higher price in this event; hence firm i's marginal benefit of investment is higher. Interestingly, the price effect is independent of α_i because, as (2) shows, the price that firm i charges when it innovates and firm j fails depends on α_j but not on α_i. The cannibalization effect of PCO, given by $\lambda_j\frac{\alpha_i\pi(\alpha_i)}{k}$, arises because firm i gets a share α_i of firm j's profit, and therefore internalizes the negative effect that an increase in λ_i has on firm j's chance to be a sole innovator and earn a profit of $\pi(\alpha_i)$.

A (subgame perfect) Nash equilibrium in stage 1 is a pair $(\lambda_i^*, \lambda_j^*)$, defined by the intersection of $BR_i(\lambda_j)$ and $BR_j(\lambda_i)$ in the (λ_i, λ_j) space. The following assumption ensures that the equilibrium in stage 1 is unique, interior, and stable (see the Appendix for a proof):

A5 k is sufficiently large: $k > \underline{k} \equiv \pi(\alpha_i)\left(1+\alpha_i\frac{\pi(\alpha_i)}{\pi(\alpha_j)}\right)$ for all $0 \leq \alpha_j \leq \alpha_i < 1/2$

Assumption A5 is stronger than Assumption A4. To see why, note that since $\pi(\alpha_i)$ is (weakly) increasing with α_i and $\pi(\alpha_j)$ is (weakly) increasing with α_j, $\underline{k}$ increases with α_i and decreases with α_j and hence is maximized at $\alpha_i = 1/2$ and $\alpha_j = 0$, where its value is $\underline{k} = \pi(1/2)\left(1+\frac{\pi(1/2)}{2\pi(0)}\right) > \pi(1/2) = \pi^m$, where the last equality follows from equation (3) and because $\sigma < 1/2$ by Assumption A3.

The equilibrium in stage 1 is illustrated in Figure 1. Assumption A5 ensure that $BR_i(\lambda_j)$ crosses $BR_j(\lambda_i)$ in the interior of (λ_i, λ_j) space once and from above.[22] In the Appendix we also show that Assumption A5 ensures that the slope of $BR_i(\lambda_j)$ in the (λ_i, λ_j) space exceeds 1 in absolute value, whereas the slope of $BR_j(\lambda_i)$ is below 1.

[22] When Assumption A5 fails, there are potentially two more equilibria: in one of them only firm i invests and in the other only firm j invests. Assumption A5 eliminates these equilibria and allows us to focus on the interior equilibrium in which both firms invest.

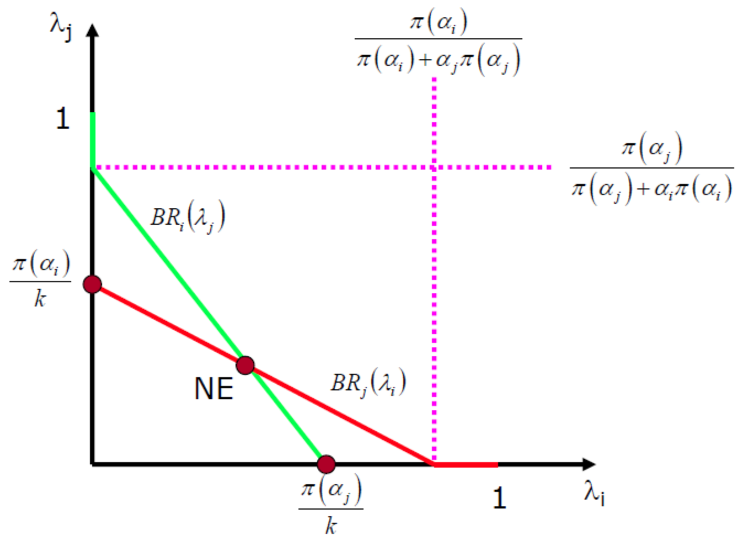

Figure 1: the best response functions in stage 1 and the Nash equilibrium

Given (4), the equilibrium investment levels are

$$\lambda_i^* = \frac{\pi(\alpha_j)\,k - \pi(\alpha_i)\left(\pi(\alpha_j) + \alpha_i \pi(\alpha_i)\right)}{k^2 - \left(\pi(\alpha_j) + \alpha_i \pi(\alpha_i)\right)\left(\pi(\alpha_i) + \alpha_j \pi(\alpha_j)\right)}, \tag{5}$$

and

$$\lambda_j^* = \frac{\pi(\alpha_i)\,k - \pi(\alpha_j)\left(\pi(\alpha_i) + \alpha_j \pi(\alpha_j)\right)}{k^2 - \left(\pi(\alpha_j) + \alpha_i \pi(\alpha_i)\right)\left(\pi(\alpha_i) + \alpha_j \pi(\alpha_j)\right)}. \tag{6}$$

Lemma 2: *The equilibrium investment levels chosen in stage 1 are given by (5) and (6), and have the following properties:*

(i) $0 < \lambda_i^* \leq \lambda_j^*$, *where* $\lambda_i^* < 1/2$ *and* $\lambda_j^* < 1$*;*

(ii) *as* $k \to \underline{k}$, $\lambda_i^* \to 0$ *and* $\lambda_j^* \to \frac{\pi(\alpha_j)}{\pi(\alpha_j)+\alpha_i\pi(\alpha_i)} > 0$ *if* $\alpha_j < \alpha_i$ *and* $\lambda_i^* = \lambda_j^* \to \frac{1}{2(1+\alpha)}$ *if* $\alpha_j = \alpha_i = \alpha$, *and as* $k \to \infty$, $\lambda_i^* \to 0$ *and* $\lambda_j^* \to 0$.

Proof: See the Appendix.

Lemma 2 states that given Assumption A5, the investment levels of both firms are strictly positive, and firm i, which holds the larger PCO stake, invests less than firm j. Firm i's investment level, λ_i^*, is bounded from above by $1/2$ (when $k \to \underline{k}$ and $\alpha_j = \alpha_i = 0$), while firm j's investment level, λ_j^*, is bounded from above by 1 (when $k \to \underline{k}$, $\alpha_j = 0$ and $\alpha_i \to 0$). Both λ_i^* and λ_j^* tend to 0 as the slope of the marginal cost of investment, k, tends to ∞. At the other extreme, as $k \to \underline{k}$,

$\lambda_i^* = \lambda_j^* \to 1/2$ if $\alpha_j = \alpha_i = 0$, and $\lambda_i^* \to 0$ and $\lambda_j^* \to 1$ if $\alpha_j = 0$, and $\alpha_i \to 0$. This latter result highlights the stark difference between symmetric and asymmetric PCO structures. Starting from no PCO's, even a small PCO by firm i in firm j has a large effect on the equilibrium investment levels. To see the logic for this, note from Figure 1 that as $k \to \underline{k}$, the vertical intercept of $BR_j(\lambda_i)$, $\frac{\pi(\alpha_i)}{k}$, tends to $\frac{\pi(\alpha_j)}{\pi(\alpha_j)+\alpha_i\pi(\alpha_i)}$, which is also the vertical intercept of $BR_i(\lambda_j)$; hence $\lambda_i^* \to 0$. When $\alpha_i \to 0$, $\frac{\pi(\alpha_j)}{\pi(\alpha_j)+\alpha_i\pi(\alpha_i)} \to 1$, so $\lambda_j^* \to 1$. However, when $\alpha_j = \alpha_i = \alpha$, the equilibrium is symmetric and $\lambda_i^* = \lambda_j^* = \frac{\pi(\alpha)}{k+\pi(\alpha)(1+\alpha)}$; when $\alpha = 0$ and $k \to \underline{k}$, this value tends to 1/2.

We now study the comparative statics of λ_i^* and λ_j^* with respect to the PCO stakes.

Proposition 1: *The PCO stake affect the equilibrium investment levels as follows:*

(i) *an increase in α_i lowers λ_i^* and increases λ_j^*, and an increase in α_j lowers λ_j^* and increases λ_i^*: $\frac{\partial \lambda_i^*}{\partial \alpha_i} < 0 < \frac{\partial \lambda_j^*}{\partial \alpha_i}$ and $\frac{\partial \lambda_j^*}{\partial \alpha_j} < 0 < \frac{\partial \lambda_i^*}{\partial \alpha_j}$; since $0 \leq \alpha_j \leq \alpha_i < \frac{1}{2}$, λ_i^* is largest under a symmetric PCO structure where $\alpha_i = \alpha_j$ and lowest under a maximally asymmetric PCO structure where $\alpha_i \to \frac{1}{2}$ and $\alpha_j = 0$, whereas λ_j^* is largest under a maximally asymmetric PCO structure and lowest under a symmetric PCO structure;*

(ii) *in the neighborhood of a symmetric PCO structure, where $\alpha_i = \alpha_j = \alpha < \sigma$, $\lambda_i^* + \lambda_j^*$ is increasing with α_i;*

(iii) *when $\alpha_i \geq \alpha_j \geq \sigma$, $\lambda_i^* + \lambda_j^*$ is decreasing with α_i and with α_j.*

The effect of changes in α_i and α_j on the equilibrium investment levels is illustrated in Figure 2. An increase in α_i induces firm i to cut λ_i^* due to the cannibalization effect; hence $BR_i(\lambda_j)$ rotates counterclockwise around its horizontal intercept, $\frac{\pi(\alpha_j)}{k}$. At the same time, an increase in α_i induces firm j to raise λ_j^* due to the price effect, so $BR_j(\lambda_i)$ shifts outward. The new equilibrium, NE_1 then lies northwest of the original equilibrium NE_0. Hence, at the new equilibrium, λ_i^* is lower and λ_j^* is higher than in the original equilibrium. In particular, starting from a symmetric PCO structure where $\alpha_i = \alpha_j$, an increase in α_i lowers λ_i^* and raises λ_j^*, so eventually, $\lambda_i^* < \lambda_j^*$.[23] In other words, firm i, which holds a bigger PCO stake in the rival, invests less than firm j. This result is consistent with Proposition 1 in Bayona and López (2018), albeit in their model, investments are deterministic rather than stochastic as in our model.

[23]Holding $BR_j(\lambda_i)$ fixed, the counterclockwise rotation of $BR_i(\lambda_j)$ around $\frac{\pi(\alpha_i)}{k}$ leads to a lower λ_i^* and a higher λ_j^*. The upward shift in $BR_j(\lambda_i)$ reinforces this effect.

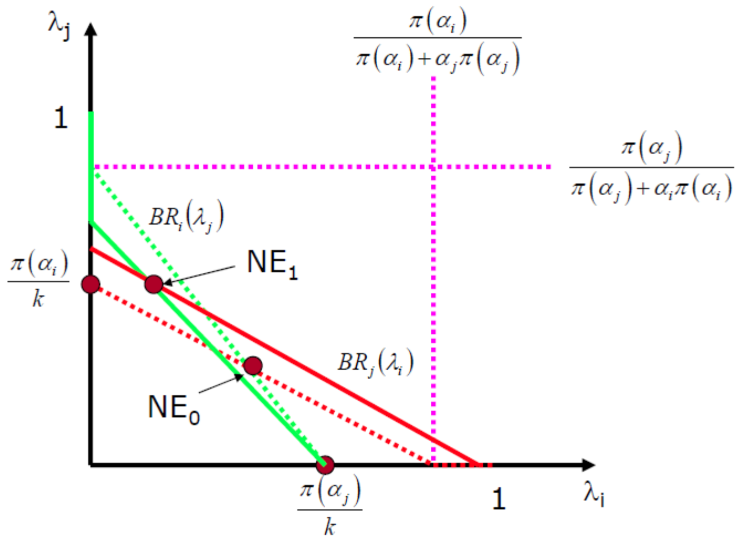

Figure 2: The effect of an increase in α on the Nash equilibrium

While Proposition 1(i) shows that an increase in α_i lowers λ_i^* and increases λ_j^*, the fact that both $BR_i(\lambda_j)$ and $BR_j(\lambda_i)$ are affected makes it hard to tell whether the change in λ_i^* is bigger than the change in λ_j^* or vice versa. Proposition 1(ii) shows that, starting from a symmetric PCO structure where $\alpha_i = \alpha_j = \alpha$, a small increase in α_i increases λ_j^* more than it decreases λ_i^*. By contrast, Proposition 1(iii) shows that when α_i are α_j sufficiently high to ensure that $\pi(\alpha_i) = \pi(\alpha_j) = \pi^m$, a small increase in α_i decreases λ_i^* more than it increases λ_j^*. The reason for this is that when $\pi(\alpha_i) = \pi(\alpha_j) = \pi^m$, an increase α_i does not lead to a price effect, so only $BR_i(\lambda_j)$ rotates counterclockwise around its horizontal intercept, while $BR_j(\lambda_i)$ stays intact. The new equilibrium then, lies on $BR_j(\lambda_i)$; since the slope of $BR_j(\lambda_i)$ is less than 1 in absolute value, λ_i^* decreases by more than λ_j^* increases.

In the next proposition, we examine how λ_i^* and λ_j^* are affected by changes in k, which is the slope of the marginal cost of investment; c, which reflects the innovation size; and π^m, which is the monopoly profit that a sole innovator earns when the innovation drastic (i.e., when $\alpha_i \geq \alpha_j \geq \sigma$).

Proposition 2: *The equilibrium investment levels are affected by k, c, and π^m, as follows,*

(i) *λ_i^* is first increasing and then decreasing with k if $\alpha_j < \alpha_i$ and is decreasing with k for all $k > \underline{k}$ if $\alpha_i = \alpha_j$, while λ_j^* is decreasing with k for all $k > \underline{k}$;*

(ii) *when* $\alpha_j \leq \alpha_i < \sigma$, $\lambda_i^* + \lambda_j^*$ *is increasing with* c*, i.e., either* λ_i^**, or* λ_j^**, or both, increase with* c*;*

(iii) *when* $\alpha_i \geq \alpha_j \geq \sigma$, λ_i^* *and* λ_j^* *are independent of* c *and depend on* π^m *only through* k/π^m*, so the effect of* π^m *is the opposite of the effect of* k*.*

Proof: See the Appendix.

Proposition 2(i) shows that, as one might expect, an increase in the slope of the marginal cost of investment, k, induces firm j to cut λ_j^*. Surprisingly, however, this is not necessarily true for λ_i^*: when k is low, an increase in k actually induces firm i to invest more. This counterintuitive result arises because of the strategic interaction between the two firms. An increase in k raises the marginal cost of investment and induces firm j to invest less. Since λ_j^* decreases, firm i is more likely to become a sole innovator, so its marginal benefit of investment increases. Although the marginal cost of firm i increases as well, when $\alpha_j < \alpha_i$ and $k \to \underline{k}$, $\lambda_i^* \to 0$, so the increase in firm i's marginal cost, $k\lambda_i^*$, is lower than the increase in the marginal benefit, so firm i invests more. By continuity, this is also true when k is not too far from $\underline{k}$. But as k increases further, the increase in $k\lambda_i^*$ eventually outweighs the associated increase in firm i's marginal benefit, so λ_i^* begins to decrease with k.[24] As $k \to \infty$, λ_i^*, as well as λ_j^*, drop to 0.

The effect of k on the equilibrium levels of investment can also be seen from Figure 1. So long as $\alpha_j < \alpha_i$, the best-response functions, $BR_i(\lambda_j)$ and $BR_j(\lambda_i)$, intersect (almost) on the vertical axis when k tends to its lower bound $\underline{k}$, so $\lambda_i^* \to 0$ and $\lambda_j^* \geq 0$. As k increases, $BR_i(\lambda_j)$ rotates clockwise around its vertical intercept, while $BR_j(\lambda_i)$ rotates counterclockwise around its horizontal intercept, so now $BR_i(\lambda_j)$ and $BR_j(\lambda_i)$ intersect at the interior of the (λ_i, λ_j) space, implying that λ_i^* becomes positive, whereas λ_j^* falls. As $k \to \infty$, $BR_i(\lambda_j)$ and $BR_j(\lambda_i)$ intersect at the origin, so $\lambda_i^* = \lambda_j^* = 0$. Overall then, λ_i^* is first increasing with k and then decreases with k, whereas λ_i^* is decreasing with k throughout. When $\alpha_j = \alpha_i$, $BR_i(\lambda_j)$ and $BR_j(\lambda_i)$ intersect on a 45^0 line, but as k increases, they shift inward, so their intersection moves closer to the origin.

Proposition 2(ii)-(iii) show that the comparative statics of λ_i^* and λ_j^* with respect to c depend on the size of the PCO stakes. When $\alpha_j \leq \alpha_i < \sigma$ (the innovation is non drastic), an increase in c (though it cannot increase by too much because by Assumption A3, $\frac{p^m}{2} < c < p^m$) implies that

[24] For firm j, the increase in marginal cost, $k\lambda_j^*$, when k increases, always outweighs the increase in marginal benefit because, unlike λ_i^*, λ_j^* is never very small.

the innovation confers a larger advantage on the innovating firm when the rival fails. This has two implications. First, an increase in c magnifies the price effect of PCO because firm i can charge a higher price as a sole innovator and therefore earn a higher profit. This effect encourages investment and is stronger when α_j is higher (the price effect is then stronger) and when λ_j is lower (firm i is more likely to be a sole innovator). Second, an increase in c also magnifies the cannibalization effect of PCO, because then firm j also earns a higher profit as a sole innovator, so firm i's innovation imposes a larger negative externality on firm j's profit. Firm i internalizes part of this negative externality due to its stake in firm j and hence it invests less when α_i is larger and when λ_j^* is larger (firm j is more likely to innovate). While in general, we cannot tell the net effect on λ_i^* and λ_j^* separately, Proposition 2(ii) shows that an increase in c shifts both $BR_i(\lambda_j)$ and $BR_j(\lambda_i)$ outward, so they intersect further away from the origin, implying that $\lambda_i^* + \lambda_j^*$ increases. Proposition 2(iii) shows by contrast that when the innovation drastic (i.e., $\alpha_i \geq \alpha_j \geq \sigma$), an increase in c does not affect λ_i^* and λ_j^* because in this range, the equilibrium price of a sole innovator is independent of c and equals p^m.

Proposition 2(iii) also shows that when the innovation is drastic (i.e., $\alpha_i \geq \alpha_j \geq \sigma$), π^m affects λ_i^* and λ_j^* only through k/π^m, and hence has the opposite effect of k.[25] In particular, λ_i^* is first increasing with π^m and then decreasing with π^m. That is, when π^m is small, an increase in π^m induces firm 1 to lower its investment. As before, this counterintuitive result occurs because the increase in π^m encourages firm j to invest more, thus lowering the marginal benefit of firm i from investment. As π^m becomes larger, both firms raise their investment levels when π^m increases.

5 Welfare analysis

In this section we examine the effect of PCO on consumer surplus, which as mentioned earlier, is the most common welfare standard in antitrust enforcement. To this end, recall that in equilibrium, consumers pay 0 if both firms innovate, c if both firms fail to innovate, $p(\alpha_j)$ if only firm i innovates, and $p(\alpha_i)$ if only firm j innovates. Therefore, expected consumer surplus, as a function of the PCO stakes, α_i and α_j, is given by

$$\begin{aligned} CS(\alpha_i, \alpha_j) &= \lambda_i^* \lambda_j^* S(0) + \lambda_i^* \left(1 - \lambda_j^*\right) S(p(\alpha_j)) + \lambda_j^* (1 - \lambda_i^*) S(p(\alpha_i)) \\ &\quad + (1 - \lambda_i^*) \left(1 - \lambda_j^*\right) S(c), \end{aligned} \tag{7}$$

[25] The only caveat is that by Assumption A5, $k > \underline{k} \equiv \pi^m + \frac{(\pi^m)^2}{2cQ(c)}$, so holding k constant, π^m cannot increase by too much. Also, naturally, $\pi^m > cQ(c)$, so π^m cannot be too low.

where λ_i^* and λ_j^* are given by (5) and (6) and $S(p) = \int_p^\infty Q(x)dx$. Since $\alpha_i \geq \alpha_j \geq 0$, (2) implies that $S(0) > S(c) \geq S(p(\alpha_j)) \geq S(p(\alpha_i))$, with strict inequalities when $\alpha_i > \alpha_j > 0$.

PCO affects expected consumer surplus both directly through the equilibrium prices, $p(\alpha_i)$ and $p(\alpha_j)$, when only one firm innovates, and indirectly through the equilibrium investment levels λ_i^* and λ_j^*. Straightforward differentiation reveals that

$$\frac{\partial CS(\alpha_i, \alpha_j)}{\partial \alpha_i} = \frac{\partial CS(\alpha_i, \alpha_j)}{\partial \lambda_i} \underbrace{\frac{\partial \lambda_i^*}{\partial \alpha_i}}_{(-)} + \frac{\partial CS(\alpha_i, \alpha_j)}{\partial \lambda_j} \underbrace{\frac{\partial \lambda_j^*}{\partial \alpha_i}}_{(+)} + \lambda_j^* (1 - \lambda_i^*) S'(p(\alpha_i)) p'(\alpha_i), \tag{8}$$

where, evaluated at λ_i^* and λ_j^*,

$$\frac{\partial CS(\alpha_i, \alpha_j)}{\partial \lambda_i} = \lambda_j^* \underbrace{(S(0) - S(p(\alpha_i)))}_{(+)} - \left(1 - \lambda_j^*\right) \underbrace{(S(c) - S(p(\alpha_j)))}_{(+)}, \tag{9}$$

and

$$\frac{\partial CS(\alpha_i, \alpha_j)}{\partial \lambda_j} = \lambda_i^* \underbrace{(S(0) - S(p(\alpha_j)))}_{(+)} - (1 - \lambda_i^*) \underbrace{(S(c) - S(p(\alpha_i)))}_{(+)}. \tag{10}$$

The first term in (8) is the effect of α_i on the probability that firm i innovates. Since $\frac{\partial \lambda_i^*}{\partial \alpha_i} < 0$ by Proposition 1, the sign of this term is equal to the sign of $-\frac{\partial CS(\alpha_i, \alpha_j)}{\partial \lambda_i}$. The second term in (8) is the effect of α_i on the probability that firm j innovates. By Proposition 1, $\frac{\partial \lambda_j^*}{\partial \alpha_i} > 0$, so the sign of this term is equal to the sign of $\frac{\partial CS(\alpha_i, \alpha_j)}{\partial \lambda_j}$. The third term in (8) reflects the price effect of PCO. Note that $S'(p(\alpha_i)) = -Q(p(\alpha_i)) < 0$, and note from (2) that $p'(\alpha_i) < 0$ when $\alpha_i < \sigma$ and $p'(\alpha_i) = 0$ when $\alpha_i \geq \sigma$. Hence the price effect is negative and harms consumers when $\alpha_i < \sigma$, because then an increase in α_i raises $p(\alpha_i)$, which is the equilibrium price that firm j charges when it serves the entire market. When $\alpha_i \geq \sigma$, firm j already charges p^m when it serves the entire market, so there is no price effect when α_i increases further.

Equations (9) and (10) show that an increase in λ_i^* and λ_j^* is a double-edge sword from the perspective of consumers: although it increases the likelihood that both firms innovate, in which case the equilibrium price drops to 0, it may also raise the likelihood that only one firm innovates, which is the worst situation from consumers' point of view because then the price is $p(\alpha_j)$ or $p(\alpha_i)$ instead of 0 or c. Hence, from the perspective of consumers, innovation may either be insufficient or excessive. In particular, an increase in λ_i^* boosts consumer surplus when λ_j^* is sufficiently large because then both firms are more likely to innovate, and likewise, an increase in λ_j^* boosts consumer surplus when λ_i^* is sufficiently large. Recalling from Lemma 2 that λ_j^* is particularly large under maximal asymmetry in the PCO structure, whereas λ_i^* is particularly large when the PCO structure is symmetric, $\frac{\partial CS(\alpha_i,\alpha_j)}{\partial \lambda_i}$ is more likely to be positive when the PCO structure is asymmetric and negative when the PCO structure is symmetric, and conversely for $\frac{\partial CS(\alpha_i,\alpha_j)}{\partial \lambda_j}$.

Noting that the price effect of PCO, captured by the third term in (2), is nonpositive, we can now establish sufficient conditions for an increase in α_i to harm consumers:

Proposition 3: *Given the equilibrium investment levels, $\frac{\partial CS(\alpha_i,\alpha_j)}{\partial \lambda_i} \geq \frac{\partial CS(\alpha_i,\alpha_j)}{\partial \lambda_j}$. The following conditions are sufficient for $\frac{\partial CS(\alpha_i,\alpha_j)}{\partial \alpha_i} \leq 0$:*

(i) $\frac{\partial CS(\alpha_i,\alpha_j)}{\partial \lambda_i} \geq 0 \geq \frac{\partial CS(\alpha_i,\alpha_j)}{\partial \lambda_j}$;

(ii) $\frac{\partial CS(\alpha_i,\alpha_j)}{\partial \lambda_i} \leq 0$ *and* $\lambda_i^* + \lambda_j^*$ *is increasing with* α_i*;*

(iii) $\frac{\partial CS(\alpha_i,\alpha_j)}{\partial \lambda_j} \geq 0$ *and* $\lambda_i^* + \lambda_j^*$ *is decreasing with* α_i*;*

Moreover, $\frac{\partial CS(\alpha_i,\alpha_j)}{\partial \alpha_i} < 0$ if $p'(\alpha_i) < 0$, or at least one inequality in (i)-(iii) is strict.

Proof: See the Appendix.

Proposition 3 shows that an increase in λ_i^* (the smaller investment) benefits consumers more, or harms them less, than an increase in λ_j^* (the larger investment). Perhaps more importantly, Proposition 3 provides three sufficient conditions for an increase in α_i to harm consumers. By implication then, an increase in α_i can benefit consumers only if the three conditions fail. To see the logic behind the three conditions, recall that following an increase in α_i, λ_i^* falls and λ_j^* increases. Condition (i) requires that both changes harm consumers. Condition (ii) requires that the increase in λ_j^* outweighs the decrease in λ_i^* and $\frac{\partial CS(\alpha_i,\alpha_j)}{\partial \lambda_i} \leq 0$. By (9), the latter condition is more likely to hold when λ_j^* is relatively small. Intuitively, by part (i), $\frac{\partial CS(\alpha_i,\alpha_j)}{\partial \lambda_i} \leq 0$ implies that $\frac{\partial CS(\alpha_i,\alpha_j)}{\partial \lambda_j} \leq 0$, so the increase in λ_j^* harms consumers. Although the decrease in λ_i^* benefits consumers, the harm exceeds the benefit, so overall consumers are worse off. Conversely, condition (iii) requires that the decrease in λ_i^* outweighs the increase in λ_j^* and $\frac{\partial CS(\alpha_i,\alpha_j)}{\partial \lambda_j} \geq 0$, which by (10) holds when λ_i^* is relatively high. By part (i), $\frac{\partial CS(\alpha_i,\alpha_j)}{\partial \lambda_j} \geq 0$ implies that $\frac{\partial CS(\alpha_i,\alpha_j)}{\partial \lambda_i} \geq 0$, so the decrease in λ_i^* harms consumers and outweighs the associated benefit due to the increase in λ_j^*.

The next corollary reports two special cases where the sufficient conditions in Proposition 3 become tighter.

Corollary 1: *Given the equilibrium investment levels,*

(i) *in the neighborhood of a symmetric PCO structure,* $\frac{\partial CS(\alpha_i,\alpha_j)}{\partial \lambda_i} \leq 0$ *is sufficient for* $\frac{\partial CS(\alpha_i,\alpha_j)}{\partial \alpha_i} \leq 0$*, with strict inequality when* $\frac{\partial CS(\alpha_i,\alpha_j)}{\partial \lambda_i} < 0$*;*

(ii) *when* $\alpha_i \geq \alpha_j \geq \sigma$*,* $\frac{\partial CS(\alpha_i,\alpha_j)}{\partial \lambda_j} \geq 0$ *is sufficient for* $\frac{\partial CS(\alpha_i,\alpha_j)}{\partial \alpha_i} \leq 0$*, with strict inequality when* $\frac{\partial CS(\alpha_i,\alpha_j)}{\partial \lambda_j} > 0$.

Proof: See the Appendix.

Corollary 1 shows two cases where an increase in α_i surely harms consumers. In the first case, $\frac{\partial CS(\alpha_i,\alpha_j)}{\partial \lambda_i} \leq 0$ implies $\frac{\partial CS(\alpha_i,\alpha_j)}{\partial \lambda_j} \leq 0$, and in the neighborhood of symmetric PCO structure, the negative effect due to the increase of λ_j^* outweighs the beneficial effect of the decrease in λ_i^*. In the second case, $\frac{\partial CS(\alpha_i,\alpha_j)}{\partial \lambda_j} \geq 0$ implies that $\frac{\partial CS(\alpha_i,\alpha_j)}{\partial \lambda_i} \geq 0$, and the when the innovation is drastic, the negative effect due to the decrease in λ_i^* outweighs the benefit due to the increase in λ_j^*.

So far we provided sufficient conditions for an increase in α_i to harm consumers. In the next proposition we can provide a sufficient condition for an increase in α_j to benefit consumers.

Proposition 4: *Suppose that the PCO stakes are sufficiently large to make the innovation drastic, i.e.,* $\alpha_i \geq \alpha_j \geq \sigma$*. Given the equilibrium investment levels,* $\frac{\partial CS(\alpha_i,\alpha_j)}{\partial \lambda_i} \leq 0$ *is sufficient for* $\frac{\partial CS(\alpha_i,\alpha_j)}{\partial \alpha_j} \geq 0$*, with strict inequality if* $\frac{\partial CS(\alpha_i,\alpha_j)}{\partial \lambda_i} < 0$.

Proof: See the Appendix.

6 The unit demand case

To shed more light on the welfare effects of PCO in our model, we will now consider the case where firms face a unit demand function with willingness to pay B.[26] Since the monopoly price in this case is $p^m = B$, Assumption A3 requires that $c < B < 2c$. The equilibrium price of firm i, $p(\alpha_j)$, is still given by (2), with $p^m = B$. The equilibrium price of firm j, $p(\alpha_i)$, is analogous. With a unit demand function, $p(\alpha_j)$ and $p(\alpha_i)$ are also the standalone profits of firms i and j.

In what follows we denote by $z \equiv k/c$ the ratio of the slope of marginal cost of investment, k, and the innovation size, c, and will refer to z as the "relative cost of innovation." Note that now,

[26] With a unit demand function, the market size is fixed. This property also holds in spatial models of competition (e.g., Hotelling or the circular city model) under the common assumption that the market is covered. Indeed this assumption is made in Bayona and López (2018) and Li and Zhang (2021).

the threshold of the PCO levels above which the innovation becomes drastic is $\sigma \equiv \frac{p^m - c}{p^m} = \frac{B-c}{B}$. The assumption that $c < B < 2c$ implies that $0 < \sigma < 1/2$. We will now consider two cases: the case where $\alpha_j \leq \alpha_i < \sigma$, in which the PCO stakes are relatively small so the innovation is non-drastic, and the case where $\alpha_i \geq \alpha_j \geq \sigma$, in which the PCO stakes are sufficiently large to make the innovation drastic.[27]

6.1 Non-drastic innovation: $\alpha_j \leq \alpha_i < \sigma$

The equilibrium prices in this case are $p(\alpha_j) = \frac{c}{1-\alpha_j}$ and $p(\alpha_i) = \frac{c}{1-\alpha_i}$. Hence, an increase in PCO affects consumers both directly through the equilibrium prices, as well as indirectly through the equilibrium investment levels. With a unit demand function, the standalone profits in stage 2 are $\pi(\alpha_j) = \frac{c}{1-\alpha_j}$ and $\pi(\alpha_i) = \frac{c}{1-\alpha_i}$. Assumption A5 then requires that $k > \underline{k} = \frac{c(1-\alpha_i\alpha_j)}{(1-\alpha_i)^2}$; using $z \equiv k/c$, the inequality can be written as $z > \underline{z} \equiv \frac{1-\alpha_i\alpha_j}{(1-\alpha_i)^2}$, where $\underline{z}$ is the lower bound on the relative cost of innovation. Note that $\underline{z}$ increases with α_i and decreases with α_j and hence is highest when $\alpha_i \to 1/2$ and $\alpha_j = 0$, where its value is 4 and is lowest when $\alpha_i = \alpha_j = 0$, where its value is 1.

Substituting $\pi(\alpha_j) = \frac{c}{1-\alpha_j}$ and $\pi(\alpha_i) = \frac{c}{1-\alpha_i}$ in (5) and (6), and using the definitions of z and $\underline{z}$, the equilibrium investment levels become

$$\lambda_i^* = \frac{(1-\alpha_j)(z-\underline{z})}{z^2(1-\alpha_j)^2 - \underline{z}^2(1-\alpha_i)^2}, \qquad \lambda_j^* = \frac{(1-\alpha_i)\left(\frac{z(1-\alpha_j)^2}{(1-\alpha_i)^2} - \underline{z}\right)}{z^2(1-\alpha_j)^2 - \underline{z}^2(1-\alpha_i)^2}. \tag{11}$$

The equilibrium investment levels depend only on the PCO stakes, α_i and α_j, and on the relative cost of innovation, z. In Lemma 2 we already established that $0 < \lambda_i^* \leq \lambda_j^*$, $\lambda_i^* < 1/2$, $\lambda_j^* < 1$, $\frac{\partial \lambda_i^*}{\partial \alpha_i} < 0 < \frac{\partial \lambda_j^*}{\partial \alpha_i}$ and $\frac{\partial \lambda_i^*}{\partial \alpha_j} > 0 > \frac{\partial \lambda_j^*}{\partial \alpha_j}$. In the next lemma we establish additional properties of λ_i^* and λ_j^* in the unit demand case when $\alpha_j \leq \alpha_i < \sigma$.

Lemma 3: *Suppose $\alpha_j \leq \alpha_i < \sigma$. Then,*

(i) *as $z \to \underline{z}$, $\lambda_i^* \to 0$ and $\lambda_j^* \to \frac{1-\alpha_i}{1-\alpha_i\alpha_j}$ if $\alpha_j < \alpha_i$ and $\lambda_i^* = \lambda_j^* \to \frac{1}{2(1+\alpha)}$ if $\alpha_j = \alpha_i = \alpha$, and as $z \to \infty$, $\lambda_i^* \to 0$ and $\lambda_j^* \to 0$;*

(ii) *$\lambda_i^* + \lambda_j^*$ is increasing with α_i and with α_j;*

[27] There is also an intermediate case where $\alpha_j < \sigma \leq \alpha_i$. This case is a hybrid of the two cases that we consider and we therefore do not study it.

(iii) *λ_i^* is first increasing and then decreasing with z if $\alpha_j < \alpha_i$ and is decreasing with z for all $z > \underline{z}$ if $\alpha_i = \alpha_j$, while λ_j^* is decreasing with z for all $z > \underline{z}$.*

Proof: See the Appendix.

There are two notable differences between Lemma 3 and Propositions 1 and 2. First, in Proposition 1 we can compare the effects of α_i and α_j on $\lambda_i^* + \lambda_j^*$ only in the neighborhood of a symmetric PCO structure or when $\alpha_i \geq \alpha_j \geq \sigma$. In Lemma 3 by contrast, we can make the comparison for all $0 \leq \alpha_j \leq \alpha_i < \sigma$. Second, noting that c is inversely related to z, part (iii) of Lemma 3 shows how λ_i^* and λ_j^* respond to changes in c. Hence, while in Proposition 2 we are only able to show that $\lambda_i^* + \lambda_j^*$ is increasing with c, here we can also show how c affects λ_i^* and λ_j^* separately. In particular, λ_j^* always increases with c, while λ_i^* is U-shaped in c. As in the case of an increase in k, the latter is driven by two conflicting effects. Holding λ_j^* fixed, an increase in c raises firm i's profit from being a sole innovator and hence encourages investment. But since λ_j^* increases, there is a countervailing effect as firm i has a lower chance to be the sole innovator. The second negative effect dominates when c is low, while the first positive effect dominates when c is high. Firm j also faces the same two effects, but the first positive effect dominates the second negative effect for all c.[28]

As in Lemma 2, we get a stark difference between symmetric and asymmetric PCO structures. In particular, when $z \to \underline{z}$, $\lambda_i^* = \lambda_j^* \to 1/2$ if $\alpha_j = \alpha_i = 0$, whereas $\lambda_i^* \to 0$ and $\lambda_j^* \to 1$ if $\alpha_j = 0$ and $\alpha_i \to 0$. That is, even a small asymmetry in the PCO structure can have a large effect on the equilibrium investment levels.

Consumer surplus in the unit demand case is given by $B - p$; recalling that $p = 0$ when both firms innovate, $p = c$ when both firms fail, $p = \frac{c}{1-\alpha_j}$ when only firm i innovates, and $p = \frac{c}{1-\alpha_i}$ when only firm j innovates, expected consumer surplus is therefore

$$CS(\alpha_i, \alpha_j) = B - (1 - \lambda_i^*)\left(1 - \lambda_j^*\right) c - \lambda_i^*\left(1 - \lambda_j^*\right) \frac{c}{1 - \alpha_j} - \lambda_j^*\left(1 - \lambda_i^*\right) \frac{c}{1 - \alpha_i}. \tag{12}$$

We now prove the following result:

[28] Obviously, the second effect cannot be negative and dominate the first positive effect for both firms. To see why suppose by way of negation that is is. Then both λ_i^* and λ_j^* are decreasing with c. But in order for the second effect to be negative for both firms, λ_i^* and λ_j^* must be increasing with c, a contradiction. The reason the second effect is negative and dominates the first positive effect for firm i and not for firm j is that in equilibrium $\lambda_i^* \leq \lambda_j^*$, which implies that the second negative effect is stronger for firm i than it is for firm j.

Proposition 5: *Suppose* $\alpha_j \leq \alpha_i < \sigma$. *Then,*

(i) $\frac{\partial CS(\alpha_i,\alpha_j)}{\partial \alpha_i} < 0$ *for all* $z > \underline{z}$;

(ii) $\frac{\partial CS(\alpha_i,\alpha_j)}{\partial \alpha_j} \geq 0$ *for* z *sufficiently close to* $\underline{z}$, *with strict inequality for* $\alpha_j < \alpha_i$, *and* $\frac{\partial CS(\alpha_i,\alpha_j)}{\partial \alpha_j} < 0$ *for* z *sufficiently large*;

(iii) *in the neighborhood of a symmetric PCO structure,* $\frac{\partial CS(\alpha_i,\alpha_j)}{\partial \alpha_j} < 0$ *for all* $z > \underline{z}$ *with* $\lim_{z \to \underline{z}} \frac{\partial CS(\alpha_i,\alpha_j)}{\partial \alpha_j} = 0$.

Proof: See the Appendix.

Part (i) of Proposition 5 shows that when the PCO stakes are below σ so the innovation is non-drastic, an increase in α_i, which increases the asymmetry in the PCO structure (the gap between α_i and α_j expands), unambiguously harms consumers. Part (ii) shows that an increase in α_j, which leads to a greater symmetry in the PCO structure, benefits consumers when the relative cost of innovation, z, is small, but harms consumers when z is large. How small z should be such that an increase in α_j still benefits consumers, depends on the PCO structure. In particular, part (iii) of the proposition shows that in the neighborhood of a symmetric PCO structure, an increase in α_j never benefits consumers. But when the PCO structure becomes more asymmetric, an increase in α_j benefits consumers for a larger set of values of z.[29]

Proposition 5 implies that when innovations are non drastic, antitrust agencies that pursue a consumer welfare standard, should not allow the firm with the larger PCO stake to increase its stake, but may allow the rival to increase its stake, provided that the relative cost of innovation, z, is low. In particular, for high levels of z, consumer surplus decreases as α_j increases towards α_i (which is by assumption higher than α_j), so antitrust authorities that pursue a consumer welfare standard should oppose acquisitions of PCO by both firms. For low levels of z, consumer surplus first increases as α_j increases (by part ii), but then decreases as α_j gets closer to α_i (by part iii).[30]

[29] For instance, if $\alpha_j = 0$, the largest z for which $\frac{\partial CS(\alpha_i,\alpha_j)}{\partial \alpha_j} \geq 0$ is 1.96 if $\alpha_i = 0.1$, 2.99 if $\alpha_i = 0.2$, 5.06 if $\alpha_i = 0.33$, and 10.24 as $\alpha_i \to 1/2$.

[30] To illustrate, let $B \to 2$ and $c = 1$ (in which case $\sigma \to 1/2$). If $z = 1.7$ and $\alpha_i = 0.1$, then consumer surplus increases with α_j, so long as $\alpha_j < 0.05$, and it then decreases as α_j approaches 0.1.

Hence, antitrust authorities that pursue a consumer welfare standard should encourage the firm with the low PCO stake to increase its stake, but not all the way up to α_i.

To see the relationship between Proposition 5 and the sufficient conditions in Proposition 3, recall from part (ii) of Lemma 3 that $\lambda_i^* + \lambda_j^*$ increases with α_i. Hence, by Proposition 3(ii), $\frac{\partial CS(\alpha_i,\alpha_j)}{\partial \lambda_i^*} \leq 0$ is sufficient for $\frac{\partial CS(\alpha_i,\alpha_j)}{\partial \alpha_i} \leq 0$. Differentiating (12) yields

$$\frac{\partial CS\left(\alpha_i,\alpha_j\right)}{\partial \lambda_i^*} = \frac{\left(1-\alpha_i\alpha_j\right)c}{\left(1-\alpha_i\right)\left(1-\alpha_j\right)}\left(\lambda_j^* - \frac{\alpha_j\left(1-\alpha_i\right)}{1-\alpha_i\alpha_j}\right) \leq 0, \qquad \Leftrightarrow \qquad \lambda_j^* \leq \frac{\alpha_j\left(1-\alpha_i\right)}{1-\alpha_i\alpha_j}.$$

Recalling from Lemma 3 that λ_j^* is increasing with α_i and decreasing with α_j and with z, and noting that $\frac{\alpha_j(1-\alpha_i)}{1-\alpha_i\alpha_j}$ is decreasing with α_i and increasing with α_j, the above condition is more likely to hold when α_j is large and α_i and z are small. In particular, when $\alpha_i = \alpha_j = \alpha$, $\lambda_j^* = \frac{1}{1+\alpha+z(1-\alpha)}$ and $\frac{\alpha_j(1-\alpha_i)}{1-\alpha_i\alpha_j} = \frac{\alpha}{1+\alpha}$, so the condition holds when $\alpha \geq \frac{1}{z-1}$. When $\alpha < \frac{1}{z-1}$, the sufficient condition fails, despite the fact that by Proposition 5, $\frac{\partial CS(\alpha_i,\alpha_j)}{\partial \alpha_i} < 0$. The reason for this is that the sufficient condition does not take into account the price effect of PCO: the fact that an increase in α_i leads to a higher price which harms consumers when firm j is the sole innovator.

Finally, as mentioned in Section 2, several existing papers have studied symmetric models and examined what happens when all firms assign a higher weight to the profits of all other firms by the exact same amount. For instance López and Vives (2019) conclude that an increase in PCO harms consumers unless there are large enough R&D spillovers. Indeed, in our model (which does not have R&D spillovers), if we evaluate (12) at $\alpha_j = \alpha_i = \alpha < \sigma$ and differentiate $CS\left(\alpha,\alpha\right)$ with respect to α, we get

$$\frac{\partial CS\left(\alpha,\alpha\right)}{\partial \alpha} = \frac{-2c\left(1+\left(z\left(1-\alpha\right)+\alpha\right)\left(z\left(1-\alpha^2\right)+\alpha^2\right)\right)}{\left(1-\alpha\right)^2\left(z\left(1-\alpha\right)+1+\alpha\right)^3} < 0.$$

That is, an increase in the stake that both firms hold in each other, α, unambiguously harms consumers. Proposition 5 shows however that when the PCO structure is not symmetric, an increase in α_j, holding α_i fixed, benefits consumers if z is sufficiently small. In other words, not every increase in PCO necessarily harms consumers absent R&D spillovers.

6.2 Drastic innovation: $\alpha_i \geq \alpha_j \geq \sigma$

Here the PCO stake are sufficiently large to make the innovation drastic. Note in particular that this case can arise only when both firms hold stakes in one another. When $\alpha_i \geq \alpha_j \geq \sigma$, the equilibrium prices are $p\left(\alpha_j\right) = p\left(\alpha_j\right) = p^m = B$, implying that an increase in PCO only affects the equilibrium investment levels, but does not have a price effect. With a unit demand function,

$\pi(\alpha_i) = \pi(\alpha_i) = B$. Assumption A5 then requires that $k > \underline{k} \equiv (1+\alpha_i) B$, or using $m \equiv k/B$, it requires that $m > 1 + \alpha_i$. Notice that m is the analog of z in the monopoly case and reflects the relative cost of innovation when it is drastic.

Substituting $\pi(\alpha_i) = \pi(\alpha_i) = B$ in (5) and (6), the equilibrium investment levels become

$$\lambda_i^* = \frac{m - (1+\alpha_i)}{m^2 - (1+\alpha_i)(1+\alpha_j)}, \qquad \lambda_j^* = \frac{m - (1+\alpha_j)}{m^2 - (1+\alpha_i)(1+\alpha_j)}. \tag{13}$$

λ_i^* and λ_j^* depend only on the PCO stakes, α_i and α_j, and on relative cost of innovation, m.[31] It is easy to verify that, as in Lemma 2, $0 < \lambda_i^* \leq \lambda_j^*$, $\lambda_i^* < 1/2$, and $\lambda_j^* < 1$ and $\frac{\partial \lambda_i^*}{\partial \alpha_i} < 0 < \frac{\partial \lambda_j^*}{\partial \alpha_i}$, $\frac{\partial \lambda_i^*}{\partial \alpha_j} > 0 > \frac{\partial \lambda_j^*}{\partial \alpha_j}$. In the next lemma we establish additional properties of λ_i^* and λ_j^*.

Lemma 4: *Suppose $\alpha_i \geq \alpha_j \geq \sigma$. Then,*

(i) *as $m \to 1+\alpha_i$, $\lambda_i^* \to 0$ and $\lambda_j^* \to \frac{1}{1+\alpha_i}$ if $\alpha_j < \alpha_i$ and $\lambda_i^* = \lambda_j^* \to \frac{1}{2(1+\alpha)}$ if $\alpha_j = \alpha_i = \alpha$, and as $m \to \infty$, $\lambda_i^* \to 0$ and $\lambda_j^* \to 0$;*

(ii) *$\lambda_i^* + \lambda_j^*$ is decreasing with α_i and with α_j;*

(iii) *λ_i^* is first increasing with m for $m < \widehat{m}$ and then decreasing with m for $m > \widehat{m}$, where $\widehat{m} \equiv 1 + \alpha_i + \sqrt{(\alpha_i - \alpha_j)(1+\alpha_i)}$, if $\alpha_j < \alpha_i$ and is decreasing with m for all $m > 1+\alpha_i$ if $\alpha_i = \alpha_j$, while λ_j^* is decreasing with m for all $m > 1 + \alpha_i$.*

Proof: See the Appendix.

Lemma 4(i) is consistent with Lemmas 2 and 3. Lemma 4(ii) is the opposite Lemma 3(ii): when the PCO stakes are sufficiently large to make the innovation drastic, an increase in α_i has a larger effect on λ_i^* than on λ_j^* and vice versa for an increase in α_j. That is, an increase in the PCO of a firm has a bigger effect on its own investment level than on the rival's investment level.

Lemma 4(iii) shows that λ_j^* is decreasing with the relative cost of innovation, m, and hence is bounded from above by $\frac{1}{1+\alpha_i}$, which is the value of λ_j^* as $m \to 1+\alpha_i$; as $\alpha_i \to 0$, $\lambda_j^* \to 1$. Note again the stark difference between symmetric and asymmetric PCO structures. In particular, if $m \to 1 + \alpha_i$, then $\lambda_i^* = \lambda_j^* \to 1/2$ if $\alpha_j = \alpha_i = 0$, but $\lambda_i^* \to 0$ and $\lambda_j^* \to 1$ if $\alpha_j = 0$ and $\alpha_i \to 0$. Lemma 4(iii) also shows that so long as $\alpha_j < \alpha_i$, λ_i^* is an inverse U-shaped function of m. The cutoff $\widehat{m}$ above which λ_i^* is decreasing with m is increasing with α_i and decreasing with α_j.

[31] That is, the parameters k and B affect the equilibrium only through the ratio $m \equiv k/B$.

That is, the set of values of m for which firm i increases its equilibrium investment is expands as the PCO structure become more asymmetric, and is particularly large under maximal asymmetry where $\alpha_j \to \sigma$ and $\alpha_i \to 1/2$. By contrast, when the PCO structure is symmetric, i.e., $\alpha_i = \alpha_j = \alpha$, $\widehat{m} = 1 + \alpha < m$, where the last inequality follows from Assumption A5, so λ_i^* is decreasing with m for all feasible values of m.

Turning to consumer surplus, expected consumer surplus is given by

$$CS(\alpha_i, \alpha_j) = \lambda_i^* \lambda_j^* B + (1 - \lambda_i^*)\left(1 - \lambda_j^*\right)(B - c) = B\left[\lambda_i^* \lambda_j^* + (1 - \lambda_i^*)\left(1 - \lambda_j^*\right)\sigma\right]. \tag{14}$$

Equation (14) reflects the fact that with probability $\lambda_i^* \lambda_j^*$, both firms innovate, so $p = 0$ and consumer surplus is B; with probability $\lambda_i^*\left(1 - \lambda_j^*\right) + \lambda_j^*(1 - \lambda_i^*)$, only one firm innovates, so $p = B$ and consumer surplus is 0; and with probability $(1 - \lambda_i^*)\left(1 - \lambda_j^*\right)$, both firms do not innovate, so $p = c$ and consumer surplus is $B - c$. In the next proposition we examine how $CS(\alpha_i, \alpha_j)$ is affected by α_i and α_j.

Proposition 6: *Suppose $\alpha_i \geq \alpha_j \geq \sigma$. Then,*

(i) $\frac{\partial CS(\alpha_i,\alpha_j)}{\partial \alpha_i} > 0$ *and* $\frac{\partial CS(\alpha_i,\alpha_j)}{\partial \alpha_j} > 0$ *when m is sufficiently large, or when $\sigma \to 1/2$;*

(ii) $\frac{\partial CS(\alpha_i,\alpha_j)}{\partial \alpha_i} < 0$ *and* $\frac{\partial CS(\alpha_i,\alpha_j)}{\partial \alpha_j} > 0$ *when m is not too much above its lower bound, $1 + \alpha_i$;*

(iii) *when $\sigma \to 0$ and m is such that $\sigma m \to 0$,* $\frac{\partial CS(\alpha_i,\alpha_j)}{\partial \alpha_i} < 0$ *and* $\frac{\partial CS(\alpha_i,\alpha_j)}{\partial \alpha_j} > 0$ *as $m < \widehat{m}$ and* $\frac{\partial CS(\alpha_i,\alpha_j)}{\partial \alpha_j} < 0$ *as $m > \widehat{m}$; as σ increases, it is more likely that* $\frac{\partial CS(\alpha_i,\alpha_j)}{\partial \alpha_j} > 0$*;*

(iv) *in the neighborhood of a symmetric PCO structure, where $\alpha_i = \alpha_j = \alpha \geq \sigma$,* $\frac{\partial CS(\alpha_i,\alpha_j)}{\partial \alpha_i} = \frac{\partial CS(\alpha_i,\alpha_j)}{\partial \alpha_j} > 0$ *if $m > \frac{1}{\sigma} - \alpha$ and* $\frac{\partial CS(\alpha_i,\alpha_j)}{\partial \alpha_i} = \frac{\partial CS(\alpha_i,\alpha_j)}{\partial \alpha_j} < 0$ *if $m < \frac{1}{\sigma} - \alpha$.*

Parts (i)-(iii) of Proposition 6 show that when the PCO stakes are sufficiently large to make the innovation drastic, a further increase in α_i benefits consumers when the relative cost of innovation, m, is sufficiently large, or the inverse size of innovation, σ, is large, and harms consumers when m or σ are small. Parts (i)-(ii) of the proposition show in addition that an increase in α_j benefits consumers when m or σ are sufficiently large, or when m is close to its lower bound, $1 + \alpha_i$. Part (iii) suggests that in fact, an increase in α_j may benefit consumers for all $m > 1 + \alpha_i$, provided that σ is sufficiently large. The reason for this is that $\frac{\partial CS(\alpha_i,\alpha_j)}{\partial \alpha_j}$ is a positive cubic function of m (the coefficient on m^3 is positive) and is increasing with σ. It turns out that at $m = 1 + \alpha_i$, $\frac{\partial CS(\alpha_i,\alpha_j)}{\partial \alpha_j} > 0$ and is increasing with m at a decreasing rate. Hence, $\frac{\partial CS(\alpha_i,\alpha_j)}{\partial \alpha_j}$ reaches a local maximum at some m above $1+\alpha_i$, but then it decreases and reaches a local minimum before increasing again with m.[32] If the value at the local minimum is still positive (which is more likely when σ is large), then $\frac{\partial CS(\alpha_i,\alpha_j)}{\partial \alpha_j} > 0$ for all $m > 1+\alpha_i$. If the value at the local minimum is negative, then $\frac{\partial CS(\alpha_i,\alpha_j)}{\partial \alpha_j} > 0$ for low or high values of m, but $\frac{\partial CS(\alpha_i,\alpha_j)}{\partial \alpha_j} < 0$ for intermediate values of m.[33] Proposition 6(iii) also shows that when $\sigma \to 0$, an increase in α_j may either benefit consumers when m is small or harm them when m is large.[34] This is because $\frac{\partial CS(\alpha_i,\alpha_j)}{\partial \alpha_j}$ is no longer cubic in m when $\sigma \to 0$ and $\sigma m \to 0$, but rather an inverse U-shaped function of m and hence $\frac{\partial CS(\alpha_i,\alpha_j)}{\partial \alpha_j} < 0$ when m is large. Finally, part (iv) of Proposition 6 shows that an increase in the PCO stakes when they are symmetric benefits consumers if m is large, but harms them if m is small.[35]

Proposition 6 implies that when the PCO stakes are sufficiently large to make innovation drastic and the relative cost of innovation, m, or the inverse size of innovation, σ, are large, an increase in both PCO stakes is beneficial for consumers. Consumer surplus is then maximized when $\alpha_j = \alpha_i \to 1/2$. By contrast, when m is not too much above its lower bound, $1+\alpha_i$, consumer surplus is largest when α_i is small and α_j is large; since $\alpha_j \leq \alpha_i$, this occurs when $\alpha_j = \alpha_i = \sigma$. Note that in both case, a symmetric PCO structure benefits consumers.

To see how Proposition 6 is related to the sufficient conditions in Proposition 3, notice that by Lemma 4, $\lambda_i^* + \lambda_j^*$ is decreasing with α_i, and recall from Proposition 3(iii) that in this case, $\frac{\partial CS(\alpha_i,\alpha_j)}{\partial \lambda_j^*} \geq 0$ is sufficient for $\frac{\partial CS(\alpha_i,\alpha_j)}{\partial \alpha_i} \leq 0$. Differentiating (14),

$$\frac{\partial CS\left(\alpha_i,\alpha_j\right)}{\partial \lambda_j^*} = (1+\sigma)\, B\left(\lambda_i^* - \frac{\sigma}{1+\sigma}\right) \geq 0, \qquad \Leftrightarrow \qquad \lambda_i^* \geq \frac{\sigma}{1+\sigma}.$$

This condition can hold only when $\sigma < 1/3$, otherwise $\frac{\sigma}{1+\sigma} > 1/2$ which is the upper bound on λ_i^* by Lemma 2. Indeed, Proposition 6 shows that $\frac{\partial CS(\alpha_i,\alpha_j)}{\partial \alpha_i} < 0$ when $\sigma \to 0$ and $\frac{\partial CS(\alpha_i,\alpha_j)}{\partial \alpha_i} > 0$

[32] $\frac{\partial CS(\alpha_i,\alpha_j)}{\partial \alpha_i}$ is also a positive cubic function of m, but at $m = 1+\alpha_i$ it is negative and decreasing with m, implying that it continues to decrease until it reaches a local minimum and then it is increasing with m. As a result, $\frac{\partial CS(\alpha_i,\alpha_j)}{\partial \alpha_i} < 0$ for m not too far from $1+\alpha_i$ and $\frac{\partial CS(\alpha_i,\alpha_j)}{\partial \alpha_i} > 0$ for m sufficiently large.

[33] For instance, if $\alpha_i = 0.4$, $\alpha_j = 0.3$, and $\sigma = 0.25$, then $\frac{\partial CS(\alpha_i,\alpha_j)}{\partial \alpha_j} > 0$ if $1.4 < m < 2.14$, $\frac{\partial CS(\alpha_i,\alpha_j)}{\partial \alpha_j} < 0$ if $2.1 < m < 3.34$, and $\frac{\partial CS(\alpha_i,\alpha_j)}{\partial \alpha_j} > 0$ if $m > 3.34$. However, if $\alpha_i = 0.4$, $\alpha_j = 0.3$, and $\sigma = 0.3$, then $\frac{\partial CS(\alpha_i,\alpha_j)}{\partial \alpha_j} > 0$ for all $m > 1.4$.

[34] When $\sigma \to 0$ and $\sigma m \to 0$, $\frac{\partial CS(\alpha_i,\alpha_j)}{\partial \alpha_j}$ is no longer cubic in m but rather negative quadratic function of m, so $\frac{\partial CS(\alpha_i,\alpha_j)}{\partial \alpha_j} < 0$ when m is large.

[35] Note that as $\sigma \to 1/2$, $m > \frac{1}{\sigma} - \alpha$, because by Assumption A5, $m > 1+\alpha_i$, where $1+\alpha_i$ is bounded from above by 3/2, whereas $\frac{1}{\sigma} - \alpha \to 3/2$ as $\alpha \geq \sigma \to 1/2$. Hence, part (vi) of the proposition implies that $\frac{\partial CS(\alpha_i,\alpha_j)}{\partial \alpha_i} = \frac{\partial CS(\alpha_i,\alpha_j)}{\partial \alpha_j} > 0$, which is consistent with part (i).

when $\sigma \to 1/2$.

Finally, we contrast Proposition 6 with the case where both firms hold the exact same stakes in each other and these stakes increase (rather than a unilateral increase in one of the two stakes). To this end, we evaluate (14) at $\alpha_j = \alpha_i = \alpha$ and differentiate with respect to α:

$$\frac{\partial CS(\alpha, \alpha)}{\partial \alpha} = \frac{2\sigma B \left(m - \left(\frac{1}{\sigma} - \alpha\right)\right)}{(m + \alpha + 1)^3}.$$

The derivative is positive if $m > \frac{1}{\sigma} - \alpha$ and negative if $m < \frac{1}{\sigma} - \alpha$, similarly to Proposition 6(iii). Now, an increase in α has the same effect as the increase in the stake of only one firm when the stakes are (nearly) symmetric. The reason is that as (13) shows, λ_i^* and λ_j^* are symmetric when the innovation is drastic (the price is B regardless of which firm is the sole innovator).

7 PCO vs. semicollusion, RJV, and mergers

The reason why PCO may boost investment is that it softens price competition when only one firm innovates and hence boosts expected profits. One may wonder then how PCO performs relative to other arrangements that also soften competition, like outright collusion in the product market, a research joint venture (RJV), or a full merger. The difference between the three arrangements is that under collusion, firms coordinate their pricing strategies in stage 2, but compete in stage 1. This situation is often referred to in the literature as semicollusion. Under an RJV the reverse is true: here firms fully coordinate their investments in stage 1 but then compete in the product market. Under a full merger, firms fully coordinate their strategies in both stages.

It should be noted that policymakers recognize the importance of RJVs and mergers for promoting innovation. For instance, in the U.S., the 2000 "Antitrust Guidelines for Collaborations among Competitors" of the DOJ and the FTC state that "Such collaborations often are not only benign but procompetitive."[36] Likewise, the 2010 "Horizontal Merger Guidelines" of the DOJ and FTC state that "the Agencies consider the ability of the merged firm to conduct research or development more effectively. Such efficiencies may spur innovation but not affect short-term pricing." As for semicollusion, Fershtman and Gandal (1994) and Brod and Shivakumar (1999) show that semicollusion can promote investments in R&D and benefit consumers.[37]

[36] Indeed the U.S. Congress has protected certain collaborations from full antitrust liability by passing the National Cooperative Research Act of 1984 and the National Cooperative Research and Production Act of 1993 (codified together at 15 U.S.C. § § 4301-06).

[37] Similarly, Schinkel and Spiegel (2017) show that when firms invest in the sustainability of their respective products

To simplify the comparisons, we will consider in what follows a symmetric PCO structure such that $\alpha_i = \alpha_j = \alpha \in \left[\sigma, \frac{1}{2}\right)$; that is, we will assume that the PCO stakes are sufficiently large to make the innovation drastic. Then, $p(\alpha) = p^m$ and $\pi(\alpha) = \pi^m$. Since $p(\alpha) = p^m$, PCO does not have a price effect, so the only difference between PCO and the other three arrangements is due to the cannibalization effect. Substituting $\alpha_i = \alpha_j = \alpha$ and $\pi(\alpha) = \pi^m$ in (5) and (6), the equilibrium investment levels are:

$$\lambda_i^* = \lambda_j^* = \lambda(\alpha) \equiv \frac{\pi^m}{k + (1+\alpha)\pi^m}. \tag{15}$$

Expected consumer surplus under PCO is

$$CS(\alpha) = \lambda(\alpha)^2 S(0) + 2\lambda(\alpha)(1-\lambda(\alpha)) S(p^m) + (1-\lambda(\alpha))^2 S(c). \tag{16}$$

This expression reflects the idea that with probability $\lambda(\alpha)^2$ both firms succeed and the price is 0, with probability $2\lambda(\alpha)(1-\lambda(\alpha))$ only one firm succeeds so the price is p^m, and with probability $(1-\lambda(\alpha))^2$, both firms fail and the price is c.

7.1 PCO vs. semicollusion

We begin by comparing PCO with semicollusion: firms compete in stage 1 when choosing their investment levels, but then collude in stage 2 when they set prices.[38] To simplify matters, we will focus on a pure price fixing scheme whereby firms charge the same price and split the market equally. The collusive price is p^m when both firms have a marginal cost 0, $p_c^m \equiv \arg\max (p-c) Q(p)$ when both firms have a marginal cost c, and $\hat{p} \in [p^m, p_c^m]$ when one firm innovates and its marginal cost is 0, while the other firm fails and its marginal cost is c. Recall that $\pi^m \equiv p^m Q(p^m)$ is the monopoly profit when marginal cost is 0 and let $\pi_c^m \equiv (p_c^m - c) Q(p_c^m)$ be the monopoly profit when marginal cost is c and firms charge p_c^m, $\hat{\pi} \equiv \hat{p} Q(\hat{p})$ be the monopoly profit when marginal cost is 0 and firms charge a price $\hat{p}$, and $\hat{\pi}_c \equiv (\hat{p} - c) Q(\hat{p})$ be the corresponding profit when marginal cost is c.

(which boosts the willingness of consumers to pay) before competing in the product market, semicollusion promotes investements and may benefit consumers if the slope of the marginal cost of investment is sufficiently low relative to the degree of product differentiation.

[38] The collusive scheme can be supported by repeated interaction in stage 2 (for details, see the Appendix in Schinkel and Spiegel, 2017).

The expected value of firm i in stage 1 of the game is,[39]

$$\lambda_i \lambda_j \frac{\pi^m}{2} + \lambda_i (1 - \lambda_j) \frac{\hat{\pi}}{2} + \lambda_j (1 - \lambda_i) \frac{\hat{\pi}_c}{2} + (1 - \lambda_i)(1 - \lambda_j) \frac{\pi_c^m}{2} - \frac{k\lambda_i^2}{2}.$$

In a symmetric equilibrium,

$$\lambda^{**} = \frac{\hat{\pi} - \pi_c^m}{2k + \hat{\pi} + \hat{\pi}_c - \pi_c^m - \pi^m}.$$

We now prove the following result:

Proposition 7: *Suppose that the two firms hold the same PCO stake $\alpha \geq \sigma$ in each other, so under PCO, $p(\alpha) = p^m$. Then, PCO leads to more investment than semicollusion and also yields a higher expected consumer surplus.*

Proof: See the Appendix.

Proposition 7 says that a symmetric PCO which is large enough to make the innovation drastic boosts investment more than semicollusion and benefits consumers. To see the intuition, note first that our assumption that $p(\alpha) = p^m$ implies an increase in α does not have a price effect, so the only difference between PCO and semicollusion is due to the cannibalization effect. Under semicollusion, firms charge a price above cost and earn a profit even when they both innovate. By contrast, under PCO, a firm earns a profit only if it innovates and its rival fails. Conditional on firm j innovating, innovation by firm i raises its profit from $\frac{\hat{\pi}_c}{2}$ to $\frac{\pi^m}{2}$ under semicollusion, whereas under PCO it entails a loss of $\alpha\pi^m$, because competition eliminates firm j's profit, which firm i shares through PCO. And, conditional on firm j failing to innovate, innovation raises firm i's profit under semicollusion from $\frac{\pi_c^m}{2}$ to $\frac{\hat{\pi}}{2}$, whereas under PCO it raises it by π^m. Hence, the marginal benefit of investment is greater under semicollusion when firm j innovates, but it is smaller when firm j fails. But since Assumption A4 implies that k is large, the probability that firm j innovates is relatively small, so the advantage of PCO over semicollusion is greater than the disadvantage.

As for expected consumer surplus, note that under PCO, the price is either 0 if both firms innovate, p^m when only one firm innovates, and c if neither firm innovates. Under semicollusion by contrast, the corresponding prices are p^m, $\widehat{p}$, and p_c^m and are all higher. Hence, consumers are better off under PCO, because innovation is more likely, and prices are lower in each regime.

[39] When the interaction in the product market is repeated, the per-period profits must be divided by the discount factor. To make the model comparable to our two-stage game, we can assume that although the innovation takes place once and for all in stage 1, firms must incur the cost of the innovation in every period (e.g., each firm i pays a "royalty" or a maintenance cost $\frac{k\lambda_i^2}{2}$ in every period).

7.2 PCO vs. RJV

We now compare PCO with an RJV. Under an RJV, firms cooperate in stage 1 when they choose investments, but then compete in the product market in stage 2. We follow Choi (1993) by assuming that by forming an RJV in stage 1, firms perfectly coordinate their R&D investments, but their respective probabilities of success, as well as their stage 2 prices, are independent across firms. In other words, under an RJV, firms coordinate their investments, but implement the innovation independently, so the realization of each firm's cost is independent of the rival's cost.[40]

For now, we will assume that under an RJV, firms also have symmetric PCO stakes in each other such that $\alpha_i = \alpha_j = \alpha \in \left[\sigma, \frac{1}{2}\right)$. The difference is that under (pure) PCO, the two firms invest independently in stage 1, whereas with an RJV, they choose λ_i and λ_j in stage 1, to maximize the sum of their values, given by,

$$\begin{aligned} V_i + V_j &= \frac{(1+\alpha)\left(\lambda_i\left(1-\lambda_j\right)\pi^m - \frac{k\lambda_i^2}{2}\right) + (1+\alpha)\left(\lambda_j(1-\lambda_i)\pi^m - \frac{k\lambda_j^2}{2}\right)}{1-\alpha^2} \\ &= \frac{\left(\lambda_i\left(1-\lambda_j\right)+\lambda_j(1-\lambda_i)\right)\pi^m - \frac{k\lambda_i^2}{2} - \frac{k\lambda_j^2}{2}}{1-\alpha}. \end{aligned}$$

The investment levels which maximize this expression are

$$\lambda_i = \lambda_j = \lambda^{RJV} \equiv \frac{\pi^m}{k+2\pi^m}. \tag{17}$$

Notice from (15) that $\lambda^{RJV} < \lambda(\alpha)$ as $\alpha < 1/2$: firms invest less under RJV than under (pure) PCO. In fact, λ^{RJV} is even lower if firms do not have PCO stakes in each other under an RJV, because then the profit under RJV is $\pi(c) < \pi^m$. Intuitively, under RJV, firms fully internalize the cannibalization effect, while under PCO they only partially internalize it; hence they invest more under PCO. Moreover, if the innovation is non-drastic, PCO leads to a price effect which RJV does not, so the incentive to innovate under PCO is even larger.

The expected consumer surplus under RJV is also given by (16) with λ^{RJV} replacing $\lambda(\alpha)$. To compare consumer surplus under (pure) PCO and under RJV, it is useful to rewrite (16) as

$$CS(\alpha) = S(c) + \lambda(\alpha)^2 \underbrace{\left(S(0) - S(c)\right)}_{(+)} + 2\lambda(\alpha)\left(1-\lambda(\alpha)\right)\underbrace{\left(S(p^m) - S(c)\right)}_{(-)}.$$

[40] In Choi (1993), the RJV also generates an information spillovers, which boost the profits of the two firms. We abstract from such spillovers. Kamien, Muller, and Zang (1992) refer to the case where firms jointly choose investments, but the marginal cost of each firm depends on its own investment, as "R&D cartelization." Under RJV, firms choose investments independently, but the marginal cost of each firm depends on the sum of the investments.

That is, the baseline consumer surplus is $S(c)$. However with probability $\lambda(\alpha)^2$, both firms innovate so consumer surplus increases from $S(c)$ to $S(0)$, and with probability $2\lambda(\alpha)(1-\lambda(\alpha))$, only one firm innovates, in which case consumer surplus drops from $S(c)$ to $S(p^m)$. Note that by Assumption A3, $p^m < 2c$; hence $p^m - c < c - 0$. Since $S(\alpha)$ is decreasing and convex, $S(0) - S(c) > S(p^m) - S(c)$.[41] But then by Assumption A4, $k > \pi^m$, so from (15) it follows that $\lambda(\alpha) < 1/2$. Hence, $\lambda(\alpha)^2 < 2\lambda(\alpha)(1-\lambda(\alpha))$, implying that the sum of the second and third terms in $CS(\alpha)$ may be either positive or negative. Recalling that $\lambda^{RJV} < \lambda(\alpha)$, it follows that the second positive term is smaller under RJV, but the third negative term is also smaller in absolute value. In general then, consumers can be better or worse off under PCO relative to RJV. In the next proposition we can compare consumer surplus under PCO and under RJV.

Proposition 8: *Suppose that the two firms hold the same PCO stake α in each other. Then $\lambda(\alpha) > \lambda^{RJV}$, regardless of whether firms also hold symmetric PCO stakes in each other under an RJV. In the unit demand case, consumer surplus is larger under PCO if $m \equiv k/B$ is sufficiently small and larger under RJV if m is sufficiently large.*

Proof: See the Appendix.

7.3 PCO vs. full merger

Finally, we consider the possibility that the two firms fully merge. Then, the two firms can fully coordinate both their investment levels in stage 1, and their pricing strategies in stage 2. This is unlike semicollusion where there is coordination in stage 2 but competition in stage 1, or RJV where the opposite is true. A full merger differs from PCO in that PCO can be viewed as a "partial merger," in which firms internalize only part of their externality on rivals, while under a full merger they fully internalize it.

When firms fully merge, they charge the monopoly price p^m when at least one of them has a marginal cost 0 (only the efficient firm serves the market in this case) and charge $p_c^m \equiv \arg\max(p-c)Q(p)$ when they both have a marginal cost c. Hence, in stage 1, the merged firm chooses λ_i and λ_j to maximize the sum of the firms' values, given by

$$(1-(1-\lambda_i)(1-\lambda_j))\pi^m + (1-\lambda_i)(1-\lambda_j)\pi_c^m - \frac{k\lambda_i^2}{2} - \frac{k\lambda_j^2}{2},$$

[41] Note that $S'(p) = -Q(p) < 0$ and $S''(p) = -Q'(p) > 0$.

where $\pi_c^m \equiv (p_c^m - c) Q(p_c^m)$. The investment levels which maximize this expression are,

$$\lambda_i = \lambda_j = \lambda^m = \frac{\pi^m - \pi_c^m}{k + \pi^m - \pi_c^m}.$$

Comparing with (15), yields,

Proposition 9: *Suppose that the two firms hold the same PCO stake $\alpha \geq \sigma$ in each other, so under PCO, $p(\alpha) = p^m$. Then, $\lambda(\alpha) > \lambda^m$ if and only if*

$$\alpha < \frac{\pi_c^m k}{\pi^m (\pi^m - \pi_c^m)}. \tag{18}$$

Expected consumer surplus is higher under PCO.

Proof: See the Appendix.

Proposition 9 says that a sufficiently large symmetric PCO which induces firms to charge the monopoly price, p^m, when only one of them innovates, boosts investments more than a full merger, provided that it is not too large. In particular, PCO boosts investment more than a full merger provided that $\sigma \leq \alpha < \frac{\pi_c^m k}{\pi^m(\pi^m - \pi_c^m)}$. This condition surely holds when $\alpha \to 1/2$ because then $\sigma \leq \alpha$ and when k is sufficiently large.[42] Intuitively, there are two differences between a full merger and a PCO. Conditional on firm j innovating, firm i's innovation has no value under a full merger (one innovation is enough for the merged entity to lower its cost), whereas under PCO it implies a loss of $\alpha\pi^m$, due to the fact that firm j's profit (which firm i shares due to its PCO stake α) drops from π^m to 0. And conditional on firm j failing, innovation by firm i raises the merged entity's profit from π_c^m to π^m, whereas under PCO it implies a gain of π^m. Whether the marginal benefit of investment is greater under a full merger or under PCO therefore depends on the value of $\alpha\pi^m$ relative to the value of $\pi^m - \pi_c^m$.

As for prices, under a full merger they are always higher than under PCO since there is always a monopoly in the product market, whereas under PCO there is competition when both firms innovate or both firms fail.

[42] To see that the set of parameters for which the condition holds is non empty, note that in the unit demand case, $p^m = p_c^m = B$, $\pi^m = B$, and $\pi_c^m = B - c$. Then, $\frac{p^m - c}{p^m} = \frac{B-c}{B} \equiv \sigma$ and $\frac{\pi_c^m k}{\pi^m(\pi^m - \pi_c^m)} = \frac{(B-c)k}{Bc} \equiv \sigma z > \sigma$, because $z > \underline{z} \geq 1$.

8 Conclusion

We have explored the competitive effects of partial cross ownership (PCO) in rival firms in the context of a duopoly model in which firms first invest in innovation and then compete in prices. Innovation in our model is stochastic and can either succeed or fail. When both firms succeed or both fail, they engage in Bertrand competition and make 0 profit. But when only one firm succeeds, it captures the entire market and earns a positive profit. This profit is even higher when the failing firm holds a stake in the innovating firm, as the former is reluctant to undercut the latter because it shares its profit. PCO then creates a price effect, which boosts the incentive to invest and become the innovating firm. Although the innovation in our model is non-drastic absent PCO, it becomes drastic when the PCO stakes are large; then a sole innovator charges the monopoly price, so an increase in PCO does not lead to a further price increase. Apart from a price effect, PCO also creates a cannibalization effect: when a firm innovates, it cannibalizes the rival's profit if the rival also innovates. This effect weakens the incentive to innovate and more so the larger is the stake that the firm holds in the rival.

Interestingly, more investment is not always good for consumers: although more investment increases the likelihood that both firms innovate, which is the best outcome for consumers, it also raises the likelihood that only one firm innovates, which is the worst outcome for consumers. We provide sufficient conditions for PCO to harm or benefit consumers and then explore the welfare implications of PCO in greater detail under the assumption that consumers have a unit demand function. Importantly, we allow firm i to possibly have a bigger PCO stake in firm j than firm j has in firm i. Hence, an increase in firm i's stake in firm j makes the PCO structure more asymmetric, whereas an increase in firm j's stake in firm i makes it more symmetric.

The welfare implications of PCO in our model depend on whether the PCO stakes are or are not large enough to make the innovation drastic. When they are relatively small so the innovation is non-drastic, an increase in firm i's stake in firm j always harms consumers. An increase in firm j's stake in firm i can benefit consumers provided that the relative cost of innovation is not too high; otherwise an increase in firm j's stake in firm i also harms consumers.

When the PCO stakes are sufficiently large to make the innovation drastic, an increase in the PCO stakes only affects the equilibrium investment levels, but does not create a price effect. The PCO stakes can now benefit consumers if the relative cost of innovation is sufficiently large, or its size is large relative to the willingness of consumers to pay. Moreover, an increase in firm

j's stake in firm i can also benefit consumers if the relative cost of innovation or its size are small. We also show that increases in the PCO stakes can harm consumers; for instance, when the size of the innovation is large, an increase in firm i's stake in firm j harms consumers and an increase in firm j's stake in firm i harms consumers when the relative cost of innovation is large, or in the neighborhood of a symmetric PCO structure when the relative cost of innovation is small.

The upshot is that PCO stakes may harm or benefit consumers depending on various factors, including the size of the PCO stakes and how symmetric they are, whether the innovation is drastic or non-drastic, and the size of the innovation and its relative cost. In particular, our analysis importance of asymmetries in the PCO structure for evaluating the welfare effects of PCO.

9 Appendix

Following are technical proofs.

Existence and uniqueness of a stable interior Nash equilibrium in stage 1: Note that Assumption A5 is equivalent to $\frac{\pi(\alpha_j)}{\pi(\alpha_j)+\alpha_i\pi(\alpha_i)} > \frac{\pi(\alpha_i)}{k}$, where $\frac{\pi(\alpha_j)}{\pi(\alpha_j)+\alpha_i\pi(\alpha_i)}$ is the vertical intercept of $BR_i(\lambda_j)$ in the (λ_i, λ_j) space and $\frac{\pi(\alpha_i)}{k}$ is the vertical intercept of $BR_j(\lambda_i)$.[43] Moreover, recall that the assumption that $\alpha_i \geq \alpha_j$ implies that $\pi(\alpha_i) \geq \pi(\alpha_j)$, which implies in turn that

$$\pi(\alpha_i)\left(1 + \frac{\alpha_i \pi(\alpha_i)}{\pi(\alpha_j)}\right) \geq \pi(\alpha_j)\left(1 + \frac{\alpha_j \pi(\alpha_j)}{\pi(\alpha_i)}\right).$$

Hence, Assumption A5 also implies that $\frac{\pi(\alpha_i)}{\pi(\alpha_i)+\alpha_j\pi(\alpha_j)} > \frac{\pi(\alpha_j)}{k}$, where $\frac{\pi(\alpha_i)}{\pi(\alpha_i)+\alpha_j\pi(\alpha_j)}$ is the horizontal intercept of $BR_i(\lambda_j)$ in the (λ_i, λ_j) space and $\frac{\pi(\alpha_j)}{k}$ is the horizontal intercept of $BR_j(\lambda_i)$.[44] Together with the fact that by (4), the slope of $BR_i(\lambda_j)$ is constant, and likewise the slope of $BR_j(\lambda_i)$ is constant, Assumption A5 ensures that $BR_i(\lambda_j)$ crosses $BR_j(\lambda_i)$ in the interior of the (λ_i, λ_j) space once and from above, which ensures in turn the existence of a unique and stable Nash equilibrium in stage 1.

In fact, Assumption A5 ensures that the slope of $BR_i(\lambda_j)$ in the (λ_i, λ_j) space exceeds 1 in absolute value, whereas the slope of $BR_j(\lambda_i)$ is below 1 in absolute value. To see why, note that

[43] That is, $\frac{\pi(\alpha_j)}{\pi(\alpha_j)+\alpha_i\pi(\alpha_i)}$ is the value of λ_j for which $\lambda_i = 0$, i.e., $BR_i\left(\frac{\pi(\alpha_j)}{\pi(\alpha_j)+\alpha_i\pi(\alpha_i)}\right) = 0$, and $\frac{\pi(\alpha_i)}{k}$ is firm j's best response against $\lambda_i = 0$, i.e., $BR_j(0) = \frac{\pi(\alpha_i)}{k}$.

[44] That is, $\frac{\pi(\alpha_i)}{\pi(\alpha_i)+\alpha_j\pi(\alpha_j)}$ is the value of λ_i for which $\lambda_j = 0$ and $\frac{\pi(\alpha_j)}{k}$ is firm i's best response against $\lambda_j = 0$.

fully differentiating (4), evaluating at $\lambda_i = BR_i(\lambda_j)$, and using Assumption A5,

$$\left|\frac{\partial \lambda_j}{\partial \lambda_i}\right| = \frac{k}{\pi(\alpha_j) + \alpha_i \pi(\alpha_i)} > \frac{\pi(\alpha_i)\left(1 + \alpha_i \frac{\pi(\alpha_i)}{\pi(\alpha_j)}\right)}{\pi(\alpha_j) + \alpha_i \pi(\alpha_i)} = \frac{\pi(\alpha_i)}{\pi(\alpha_j)} \geq 1,$$

where the last inequality follows because $\alpha_i \geq \alpha_j$ implies that $\pi(\alpha_i) \geq \pi(\alpha_j)$. Similarly, evaluated at $\lambda_j = BR_j(\lambda_i)$,

$$\left|\frac{\partial \lambda_j}{\partial \lambda_i}\right| = \frac{\pi(\alpha_i) + \alpha_j \pi(\alpha_j)}{k} < \frac{\pi(\alpha_i) + \alpha_j \pi(\alpha_j)}{\pi(\alpha_i)\left(1 + \alpha_i \frac{\pi(\alpha_i)}{\pi(\alpha_j)}\right)} = \frac{1 + \alpha_j \frac{\pi(\alpha_j)}{\pi(\alpha_i)}}{1 + \alpha_i \frac{\pi(\alpha_i)}{\pi(\alpha_j)}} \leq 1,$$

where the last inequality follows because $\alpha_i \geq \alpha_j$ and $\pi(\alpha_i) \geq \pi(\alpha_j)$.

Finally, we assume that $k > \underline{k}$ for the following reason. Suppose that $k = \underline{k}$ and $\alpha_i = \alpha_j = \alpha$. Then $\underline{k} \equiv \pi(\alpha_i)\left(1 + \alpha_i \frac{\pi(\alpha_i)}{\pi(\alpha_j)}\right) = \pi(\alpha)(1+\alpha)$, so the vertical intercept of $BR_j(\lambda_i)$ in the (λ_i, λ_j) space is $\frac{\pi(\alpha)}{\underline{k}} = \frac{\pi(\alpha)}{\pi(\alpha)(1+\alpha)} = \frac{1}{1+\alpha}$. The vertical intercept of $BR_i(\lambda_j)$ in turn is $\frac{\pi(\alpha)}{\pi(\alpha)+\alpha\pi(\alpha)} = \frac{1}{1+\alpha}$. Since by symmetry the same holds for horizontal intercepts, $BR_i(\lambda_j)$ and $BR_j(\lambda_i)$ coincide. ■

Proof of Lemma 2: (i) First, note that $\lambda_i^*, \lambda_j^* > 0$ because the equilibrium in stage 1 is interior as we have proved above. Moreover, using (5) and (6) and rearranging terms,

$$\lambda_j^* - \lambda_i^* = \frac{\pi(\alpha_i)(k + \alpha_i \pi(\alpha_i)) - \pi(\alpha_j)(k + \alpha_j \pi(\alpha_j))}{k^2 - (\pi(\alpha_j) + \alpha_i \pi(\alpha_i))(\pi(\alpha_i) + \alpha_j \pi(\alpha_j))} \geq 0,$$

where the inequality follows because Assumption A5 ensures that the denominator is positive, and $\alpha_i \geq \alpha_j$ implies $\pi(\alpha_i) \geq \pi(\alpha_j)$ by Lemma 1.[45] Therefore, $0 < \lambda_i^* \leq \lambda_j^*$.

Second, in Proposition 1(i) below we prove that λ_i^* is maximized when $\alpha_j = \alpha_i = \alpha$. Evaluating (5) at $\alpha_j = \alpha_i = \alpha$,

$$\lambda_i^* = \frac{\pi(\alpha)}{k + (1+\alpha)\pi(\alpha)} < \frac{1}{2(1+\alpha)} \leq \frac{1}{2},$$

where the first inequality follows because Assumption A5 implies that when $\alpha_j = \alpha_i = \alpha$, $k > \pi(\alpha)(1+\alpha)$, and the second inequality follows because $\frac{1}{2(1+\alpha)}$ is maximized when $\alpha = 0$.

[45] In particular, note that since by Assumption A5, $k > \underline{k}$,

$$\begin{aligned} & k^2 - (\pi(\alpha_j) + \alpha_i \pi(\alpha_i))(\pi(\alpha_i) + \alpha_j \pi(\alpha_j)) > \underline{k}^2 - (\pi(\alpha_j) + \alpha_i \pi(\alpha_i))(\pi(\alpha_i) + \alpha_j \pi(\alpha_j)) \\ = \; & (\pi(\alpha_j) + \alpha_i \pi(\alpha_i))\left[\frac{\pi(\alpha_i)^2}{\pi(\alpha_j)^2}(\pi(\alpha_j) + \alpha_i \pi(\alpha_i)) - (\pi(\alpha_i) + \alpha_j \pi(\alpha_j))\right] \\ \geq \; & (\pi(\alpha_j) + \alpha_i \pi(\alpha_i))\left[(1-\alpha_j)\pi(\alpha_j) - (1-\alpha_i)\pi(\alpha_i)\right] > 0. \end{aligned}$$

Third, in Proposition 2(i) below we prove that λ_j^* decreases with k. Then using (6),

$$\lambda_j^* < \frac{\pi(\alpha_j)}{\pi(\alpha_j) + \alpha_i \pi(\alpha_i)} < 1,$$

where the upper bound on λ_j^* is its value as $k \to \underline{k} \equiv \pi(\alpha_i)\left(1 + \alpha_i \frac{\pi(\alpha_i)}{\pi(\alpha_j)}\right)$, and the last inequality follows because $\frac{\pi(\alpha_j)}{\pi(\alpha_j)+\alpha_i\pi(\alpha_i)}$ is maximized when $\alpha_j \to \alpha_i \to 0$.

(ii) Suppose that $k \to \underline{k}$. Then (5) and (6) imply that if $\alpha_j < \alpha_i$, $\lambda_i^* \to 0$ and $\lambda_j^* \to \frac{\pi(\alpha_j)}{\pi(\alpha_j)+\alpha_i\pi(\alpha_i)}$, which tends to 1 as $\alpha_i \to 0$. If $\alpha_j = \alpha_i = \alpha$, then by Assumption A5, $\underline{k} = \pi(\alpha)(1+\alpha)$, so as $k \to \underline{k}$,

$$\lambda_i^* = \lambda_j^* \to \frac{\pi(\alpha)}{\underline{k} + \pi(\alpha)(1+\alpha)} = \frac{1}{2(1+\alpha)},$$

which equals 1/2 when $\alpha = 0$. As $k \to \infty$, (5) and (6) imply that $\lambda_i^* \to 0$ and $\lambda_j^* \to 0$. ■

Proof of Proposition 1: (i) Recalling from Lemma 1 that $\pi'(\alpha_j) \geq 0$,

$$\frac{\partial BR_i(\lambda_j)}{\partial \alpha_i} = -\frac{\lambda_j(\pi(\alpha_i) + \alpha_i \pi'(\alpha_i))}{k} < 0, \qquad \frac{\partial BR_j(\lambda_i)}{\partial \alpha_i} = \frac{(1-\lambda_i)\pi'(\alpha_i)}{k} \geq 0.$$

Since the best-response functions are downward sloping, an increase in α_i shifts their intersection point northwest in the (λ_i, λ_j) space. Hence, λ_i^* decreases and λ_j^* increases. The comparative statics with respect to α_j are analogous.

(ii) Using (5) and (6) again,

$$\lambda_i^* + \lambda_j^* = \frac{(\pi(\alpha_i) + \pi(\alpha_j))k - 2\pi(\alpha_i)\pi(\alpha_j) - \alpha_i(\pi(\alpha_i))^2 - \alpha_j(\pi(\alpha_j))^2}{k^2 - (\pi(\alpha_j) + \alpha_i\pi(\alpha_i))(\pi(\alpha_i) + \alpha_j\pi(\alpha_j))}.$$

Differentiating with respect to α_i, evaluating at $\alpha_i = \alpha_j = \alpha$, and rearranging yields

$$\left.\frac{\partial(\lambda_i^* + \lambda_j^*)}{\partial \alpha_i}\right|_{\alpha_i=\alpha_j=\alpha} = \frac{\pi'(\alpha)k - (\pi(\alpha))^2}{(k + (1+\alpha)\pi(\alpha))^2}.$$

Note that when $\alpha_i = \alpha_j = \alpha$, Assumption A5 implies

$$k > \underline{k} \equiv \pi(\alpha)\left(1 + \frac{\alpha\pi(\alpha)}{\pi(\alpha)}\right) = (1+\alpha)\pi(\alpha).$$

Moreover, when $\alpha_i = \alpha_j = \alpha < \sigma$, $\pi(\alpha) = \frac{c}{1-\alpha}Q\left(\frac{c}{1-\alpha}\right)$, so

$$\pi'(\alpha) = \frac{c}{(1-\alpha)^2}\left[Q\left(\frac{c}{1-\alpha}\right) + \frac{c}{1-\alpha}Q'\left(\frac{c}{1-\alpha}\right)\right] > \frac{c}{(1-\alpha)^2}Q\left(\frac{c}{1-\alpha}\right) = \frac{\pi(\alpha)}{1-\alpha}.$$

Hence,

$$\left.\frac{\partial(\lambda_i^* + \lambda_j^*)}{\partial \alpha_i}\right|_{\alpha_i=\alpha_j=\alpha} > \frac{\frac{\pi(\alpha)}{1-\alpha} \times (1+\alpha)\pi(\alpha) - (\pi(\alpha))^2}{(k + (1+\alpha)\pi(\alpha))^2} = \frac{(\pi(\alpha))^2 \times \frac{2\alpha}{1-\alpha}}{(k + (1+\alpha)\pi(\alpha))^2} > 0.$$

(iii) When $\alpha_i \geq \alpha_j \geq \sigma$, $p(\alpha_i) = p(\alpha_j) = p^m$ and $\pi(\alpha_i) = \pi(\alpha_j) = p^m Q(p^m) \equiv \pi^m$. Then,

$$\frac{\partial\left(\lambda_i^* + \lambda_j^*\right)}{\partial \alpha_i} = -\left(\frac{\pi^m\left(k - (1+\alpha_j)\pi^m\right)}{k^2 - (1+\alpha_i)(1+\alpha_j)(\pi^m)^2}\right)^2 < 0.$$

Noting that $\lambda_i^* + \lambda_j^*$ is symmetric with respect to α_i and α_j, we also have $\frac{\partial\left(\lambda_i^* + \lambda_j^*\right)}{\partial \alpha_j} < 0$. ■

Proof of Proposition 2: (i) Starting with λ_i^*, notice from (5) that

$$\frac{\partial \lambda_i^*}{\partial k} = \frac{\pi(\alpha_j) - 2k\lambda_i^*}{k^2 - \left(\pi(\alpha_j) + \alpha_i \pi(\alpha_i)\right)\left(\pi(\alpha_i) + \alpha_j \pi(\alpha_j)\right)},$$

where the denominator is positive by Assumption A5. By Lemma 2, if $\alpha_j < \alpha_i$, then $\lambda_i^* \to 0$ as $k \to \underline{k}$, implying that $\frac{\partial \lambda_i^*}{\partial k} > 0$. Next, note that

$$\begin{aligned}
\frac{\partial^2 \lambda_i^*}{\partial k^2} &= \frac{-2\left(\lambda_i^* + k\frac{\partial \lambda_i^*}{\partial k}\right)}{k^2 - \left(\pi(\alpha_j) + \alpha_i \pi(\alpha_i)\right)\left(\pi(\alpha_i) + \alpha_j \pi(\alpha_j)\right)} - 2k \times \underbrace{\frac{\pi(\alpha_j) - 2k\lambda_i^*}{\left(k^2 - \left(\pi(\alpha_j) + \alpha_i \pi(\alpha_i)\right)\left(\pi(\alpha_i) + \alpha_j \pi(\alpha_j)\right)\right)^2}}_{\frac{\partial \lambda_i^*}{\partial k}} \\
&= \frac{-2\left(\lambda_i^* + 2k\frac{\partial \lambda_i^*}{\partial k}\right)}{k^2 - \left(\pi(\alpha_j) + \alpha_i \pi(\alpha_i)\right)\left(\pi(\alpha_i) + \alpha_j \pi(\alpha_j)\right)}.
\end{aligned}$$

Hence, $\frac{\partial^2 \lambda_i^*}{\partial k^2} \leq 0$ whenever $\frac{\partial \lambda_i^*}{\partial k} = 0$, so any extremum point must be a maximum. Since λ_i^* is a continuous function of k, it is first increasing with k (from 0) and then decreasing with k (to 0) and attains a unique maximum when $\frac{\partial \lambda_i^*}{\partial k} = 0$. If $\alpha_j = \alpha_i = \alpha$, then the numerator of $\frac{\partial \lambda_i^*}{\partial k}$ is such that

$$\pi(\alpha_j) - 2k\lambda_i^* < \pi(\alpha_j) - \frac{2\underline{k}}{2(1+\alpha)} = -\pi(\alpha_j) < 0,$$

where the first inequality follows because Lemma 2 implies that as $k \to \underline{k}$, $\lambda_i^* = \frac{1}{2(1+\alpha)}$, and the equality follows because $\underline{k} = 1 + \alpha$ when $\alpha_j = \alpha_i = \alpha$. Hence, $\frac{\partial \lambda_i^*}{\partial k} < 0$ for all $k > \underline{k}$.

Turning to λ_j^*, recalling that $\alpha_i \geq \alpha_j$ implies $\pi(\alpha_i) \geq \pi(\alpha_j)$ by Lemma 1, it follows that evaluated at $k \to \underline{k}$,

$$\lambda_j^* = \frac{\pi(\alpha_i)\underline{k} - \pi(\alpha_j)\left(\pi(\alpha_i) + \alpha_j \pi(\alpha_j)\right)}{\underline{k}^2 - \left(\pi(\alpha_j) + \alpha_i \pi(\alpha_i)\right)\left(\pi(\alpha_i) + \alpha_j \pi(\alpha_j)\right)} = \frac{\pi(\alpha_j)}{\pi(\alpha_j) + \alpha_i \pi(\alpha_i)} > 0.$$

At the other extreme where $k \to \infty$, $\lambda_j^* \to 0$. Moreover,

$$\frac{\partial^2 \lambda_j^*}{\partial k^2} = \frac{-2\left(\lambda_j^* + 2k\frac{\partial \lambda_j^*}{\partial k}\right)}{k^2 - \left(\pi(\alpha_j) + \alpha_i \pi(\alpha_i)\right)\left(\pi(\alpha_i) + \alpha_j \pi(\alpha_j)\right)},$$

so $\frac{\partial^2 \lambda_j^*}{\partial k^2} < 0$ (recall that $\lambda_j^* > 0$ in the relevant range) whenever $\frac{\partial \lambda_j^*}{\partial k} = 0$, implying that as a function of k, λ_j^* attains a unique maximum. We now show that this maximum is below $\underline{k}$, implying that

λ_j^* decreases with k for all $k > \underline{k}$. To this end, note that

$$\frac{\partial \lambda_j^*}{\partial k} = \frac{\pi(\alpha_i) - 2k\lambda_j^*}{k^2 - (\pi(\alpha_j) + \alpha_i \pi(\alpha_i))(\pi(\alpha_i) + \alpha_j \pi(\alpha_j))}.$$

The sign of the derivative depends on the sign of the numerator. Evaluating it at $k \to \underline{k}$ and recalling that $\lambda_j^* = \frac{\pi(\alpha_j)}{\pi(\alpha_j)+\alpha_i\pi(\alpha_i)}$ as $k \to \underline{k}$, yields

$$\pi(\alpha_i) - 2 \times \frac{\pi(\alpha_i)(\pi(\alpha_j) + \alpha_i \pi(\alpha_i))}{\pi(\alpha_j)} \times \frac{\pi(\alpha_j)}{\pi(\alpha_j) + \alpha_i \pi(\alpha_i)} = -\pi(\alpha_i) < 0.$$

Hence, λ_j^* is decreasing with k for all $k > \underline{k}$.

(ii) When $\alpha_j \leq \alpha_i < \sigma$, the horizontal intercept of $BR_i(\lambda_j)$, $\frac{\pi(\alpha_j)}{k}$, shifts to the right as c increases, while the vertical intercept of $BR_j(\lambda_i)$, $\frac{\pi(\alpha_i)}{k}$, shifts up. As for the vertical intercept of $BR_i(\lambda_j)$, $\frac{\pi(\alpha_j)}{\pi(\alpha_j)+\alpha_i\pi(\alpha_i)}$, substituting $\pi(\alpha_i) = \frac{c}{1-\alpha_i}Q\left(\frac{c}{1-\alpha_i}\right)$ and $\pi(\alpha_j) = \frac{c}{1-\alpha_j}Q\left(\frac{c}{1-\alpha_j}\right)$ and rearranging, yields

$$\frac{\pi(\alpha_j)}{\pi(\alpha_j) + \alpha_i \pi(\alpha_i)} = \frac{1}{1 + \frac{\alpha_i(1-\alpha_j)}{1-\alpha_i} \frac{Q\left(\frac{c}{1-\alpha_i}\right)}{Q\left(\frac{c}{1-\alpha_j}\right)}}.$$

Recalling that $\varepsilon(p) \equiv -\frac{pQ'(p)}{Q(p)}$ is the elasticity of demand

$$\begin{aligned}
\frac{\partial}{\partial c}\left(\frac{Q\left(\frac{c}{1-\alpha_i}\right)}{Q\left(\frac{c}{1-\alpha_j}\right)}\right) &= \frac{\frac{1}{1-\alpha_i}Q'\left(\frac{c}{1-\alpha_i}\right)Q\left(\frac{c}{1-\alpha_j}\right) - \frac{1}{1-\alpha_j}Q'\left(\frac{c}{1-\alpha_j}\right)Q\left(\frac{c}{1-\alpha_i}\right)}{\left(Q\left(\frac{c}{1-\alpha_j}\right)\right)^2} \\
&= \frac{Q\left(\frac{c}{1-\alpha_i}\right)}{cQ\left(\frac{c}{1-\alpha_j}\right)}\left[\frac{\frac{c}{1-\alpha_i}Q'\left(\frac{c}{1-\alpha_i}\right)}{Q\left(\frac{c}{1-\alpha_i}\right)} - \frac{\frac{c}{1-\alpha_j}Q'\left(\frac{c}{1-\alpha_j}\right)}{Q\left(\frac{c}{1-\alpha_j}\right)}\right] \\
&= -\frac{Q\left(\frac{c}{1-\alpha_i}\right)}{cQ\left(\frac{c}{1-\alpha_j}\right)}\left[\varepsilon\left(\frac{c}{1-\alpha_i}\right) - \varepsilon\left(\frac{c}{1-\alpha_j}\right)\right] \leq 0,
\end{aligned}$$

where the inequality follows because by Assumption A2, $\varepsilon'(p) \geq 0$, so $\alpha_i \geq \alpha_j$ implies that $\varepsilon\left(\frac{c}{1-\alpha_i}\right) \geq \varepsilon\left(\frac{c}{1-\alpha_j}\right)$. Hence, the vertical intercept of $BR_i(\lambda_j)$, $\frac{\pi(\alpha_j)}{\pi(\alpha_j)+\alpha_i\pi(\alpha_i)}$, shifts up when c increases. Likewise, the horizontal intercept of $BR_j(\lambda_i)$ shifts to the right. Given that both $BR_i(\lambda_j)$ and $BR_j(\lambda_i)$ shift outward, the sum of λ_i^* and λ_j^* increases.

(iii) When $\alpha_i \geq \alpha_j \geq \sigma$, $p(\alpha_i) = p(\alpha_j) = p^m$ and $\pi(\alpha_i) = \pi(\alpha_j) = \pi^m$. Substituting in (5) and (6), the equilibrium investments become

$$\lambda_i^* = \frac{\frac{k}{\pi^m} - (1+\alpha_i)}{\left(\frac{k}{\pi^m}\right)^2 - (1+\alpha_i)(1+\alpha_j)}, \qquad \lambda_j^* = \frac{\frac{k}{\pi^m} - (1+\alpha_j)}{\left(\frac{k}{\pi^m}\right)^2 - (1+\alpha_i)(1+\alpha_j)}.$$

λ_i^* and λ_j^* are independent of c and depend on π^m only through $\frac{k}{\pi^m}$. Hence, the comparative static result with respect to π^m are the opposite of those with respect to k. ■

Proof of Proposition 3: Recalling from part (i) of Proposition 1 that $\lambda_i^* \leq \lambda_j^*$,

$$\begin{aligned}\frac{\partial CS(\alpha_i,\alpha_j)}{\partial \lambda_i} - \frac{\partial CS(\alpha_i,\alpha_j)}{\partial \lambda_j} &\geq \lambda_i^*\left(S(0) - S(p(\alpha_i))\right) - (1-\lambda_i^*)\left(S(c) - S(p(\alpha_j))\right) \\ &\quad - \left[\lambda_i^*\left(S(0) - S(p(\alpha_j))\right) - (1-\lambda_i^*)\left(S(c) - S(p(\alpha_i))\right)\right] \\ &= S(p(\alpha_j)) - S(p(\alpha_i)) \geq 0,\end{aligned}$$

where the last inequality follows because by Lemma 1, $\alpha_i \geq \alpha_j$ implies $p(\alpha_i) \geq p(\alpha_j)$.

Noting that $S'(p) < 0$ and $p'(\alpha_i) \geq 0$, it follows that $\frac{\partial CS(\alpha_i,\alpha_j)}{\partial \alpha_i} \leq 0$ if the sum of the first two terms in (8) is negative, i.e.,

$$\frac{\partial CS(\alpha_i,\alpha_j)}{\partial \alpha_i} = \frac{\partial CS(\alpha_i,\alpha_j)}{\partial \lambda_i}\underbrace{\frac{\partial \lambda_i^*}{\partial \alpha_i}}_{(-)} + \frac{\partial CS(\alpha_i,\alpha_j)}{\partial \lambda_j}\underbrace{\frac{\partial \lambda_j^*}{\partial \alpha_i}}_{(+)} \leq 0, \tag{19}$$

where the signs of $\frac{\partial \lambda_i^*}{\partial \alpha_i}$ and $\frac{\partial \lambda_j^*}{\partial \alpha_i}$ are due to Proposition 1(i). The following conditions ensure that (19) holds and are therefore sufficient for $\frac{\partial CS(\alpha_i,\alpha_j)}{\partial \alpha_i} \leq 0$:

(i) If $\frac{\partial CS(\alpha_i,\alpha_j)}{\partial \lambda_i} \geq 0 \geq \frac{\partial CS(\alpha_i,\alpha_j)}{\partial \lambda_j}$, both terms of (19) are negative.

(ii) If $\frac{\partial CS(\alpha_i,\alpha_j)}{\partial \lambda_i} \leq 0$ and $\lambda_i^* + \lambda_j^*$ is increasing with α_i; using (19) and recalling from part (i) that $\frac{\partial CS(\alpha_i,\alpha_j)}{\partial \lambda_i} \geq \frac{\partial CS(\alpha_i,\alpha_j)}{\partial \lambda_j}$,

$$\frac{\partial CS(\alpha_i,\alpha_j)}{\partial \alpha_i} = \underbrace{\frac{\partial CS(\alpha_i,\alpha_j)}{\partial \lambda_i}}_{(-)}\underbrace{\frac{\partial \lambda_i^*}{\partial \alpha_i}}_{(-)} + \underbrace{\frac{\partial CS(\alpha_i,\alpha_j)}{\partial \lambda_j}}_{(-)}\underbrace{\frac{\partial \lambda_j^*}{\partial \alpha_i}}_{(+)} \leq \underbrace{\frac{\partial CS(\alpha_i,\alpha_j)}{\partial \lambda_i}}_{(-)}\underbrace{\frac{\partial\left(\lambda_i^* + \lambda_j^*\right)}{\partial \alpha_i}}_{(+)} \leq 0.$$

(iii) Likewise, if $\frac{\partial CS(\alpha_i,\alpha_j)}{\partial \lambda_j} \geq 0$ and $\lambda_i^* + \lambda_j^*$ is decreasing with α_i, then,

$$\frac{\partial CS(\alpha_i,\alpha_j)}{\partial \alpha_i} = \underbrace{\frac{\partial CS(\alpha_i,\alpha_j)}{\partial \lambda_i}}_{(+)}\underbrace{\frac{\partial \lambda_i^*}{\partial \alpha_i}}_{(-)} + \underbrace{\frac{\partial CS(\alpha_i,\alpha_j)}{\partial \lambda_j}}_{(+)}\underbrace{\frac{\partial \lambda_j^*}{\partial \alpha_i}}_{(+)} \leq \underbrace{\frac{\partial CS(\alpha_i,\alpha_j)}{\partial \lambda_i}}_{(+)}\underbrace{\frac{\partial\left(\lambda_i^* + \lambda_j^*\right)}{\partial \alpha_i}}_{(-)} \leq 0.$$

If the above inequalities are strict, or $p'(\alpha_i) > 0$, then $\frac{\partial CS(\alpha_i,\alpha_j)}{\partial \alpha_i} < 0$. ■

Proof of Corollary 1: By Proposition 1(ii), $-\frac{\partial \lambda_i^*}{\partial \alpha_i} < \frac{\partial \lambda_i^*}{\partial \alpha_i}$ in the neighborhood of a symmetric PCO structure, whereas by Proposition 1(iii), $-\frac{\partial \lambda_i^*}{\partial \alpha_i} > \frac{\partial \lambda_j^*}{\partial \alpha_i}$ when $\alpha_i \geq \alpha_j \geq \sigma$. The result then follows immediately from Proposition 3(ii) and (iii). ■

Proof of Proposition 4: Differentiating $CS(\alpha_i, \alpha_j)$ with respect to α_j yields,

$$\frac{\partial CS(\alpha_i,\alpha_j)}{\partial \alpha_j} = \frac{\partial CS(\alpha_i,\alpha_j)}{\partial \lambda_i}\underbrace{\frac{\partial \lambda_i^*}{\partial \alpha_j}}_{(+)} + \frac{\partial CS(\alpha_i,\alpha_j)}{\partial \lambda_j}\underbrace{\frac{\partial \lambda_j^*}{\partial \alpha_j}}_{(-)} + \lambda_i^*\left(1-\lambda_j^*\right)S'\left(p\left(\alpha_j\right)\right)p'\left(\alpha_j\right). \tag{20}$$

If $\alpha_i \geq \alpha_j \geq \sigma$, then $p'(\alpha_j) = 0$ and Proposition 1(iii), $\frac{\partial \lambda_i^*}{\partial \alpha_j} < -\frac{\partial \lambda_j^*}{\partial \alpha_j}$. If in addition $0 \geq \frac{\partial CS(\alpha_i,\alpha_j)}{\partial \lambda_i} \geq \frac{\partial CS(\alpha_i,\alpha_j)}{\partial \lambda_j}$, then:

$$\underbrace{\frac{\partial CS(\alpha_i,\alpha_j)}{\partial \lambda_i}}_{(-)}\underbrace{\frac{\partial \lambda_i^*}{\partial \alpha_j}}_{(+)} + \underbrace{\frac{\partial CS(\alpha_i,\alpha_j)}{\partial \lambda_j}}_{(-)}\underbrace{\frac{\partial \lambda_j^*}{\partial \alpha_j}}_{(-)} \geq \underbrace{\frac{\partial \lambda_i^*}{\partial \alpha_j}}_{(+)}\left(\underbrace{\frac{\partial CS(\alpha_i,\alpha_j)}{\partial \lambda_i} - \frac{\partial CS(\alpha_i,\alpha_j)}{\partial \lambda_j}}_{(+)}\right) \geq 0.$$

If $\frac{\partial CS(\alpha_i,\alpha_j)}{\partial \lambda_i} < 0$ then $\frac{\partial CS(\alpha_i,\alpha_j)}{\partial \alpha_j} > 0$. ■

Proof of Lemma 3: (i) It is clear from (11) that as $z \to \underline{z}$, $\lambda_i^* \to 0$ and $\lambda_j^* \to \frac{1-\alpha_i}{1-\alpha_i\alpha_j}$, and as $z \to \infty$, $\lambda_i^* \to 0$ and $\lambda_j^* \to 0$.

(ii) Recalling that $\underline{z} \equiv \frac{1-\alpha_i\alpha_j}{(1-\alpha_i)^2}$, straightforward differentiation using (11) yields

$$\frac{\partial\left(\lambda_i^* + \lambda_j^*\right)}{\partial \alpha_i} = \frac{\left(1-\alpha_j\right)^2(z-1)}{\left(z\left(1-\alpha_i\right)\left(1-\alpha_j\right) + \left(1-\alpha_i\alpha_j\right)\right)^2} > 0,$$

and

$$\frac{\partial\left(\lambda_i^* + \lambda_j^*\right)}{\partial \alpha_j} = \frac{\left(1-\alpha_i\right)^2(z-1)}{\left(z\left(1-\alpha_i\right)\left(1-\alpha_j\right) + \left(1-\alpha_i\alpha_j\right)\right)^2} > 0.$$

(iii) Straightforward differentiation using (11), yields

$$\frac{\partial \lambda_i^*}{\partial z} = \frac{\left(1-\alpha_j\right)T_i(z)}{\left(1-\alpha_i\right)^2\left(z^2\left(1-\alpha_j\right)^2 - \left(1-\alpha_i\right)^2\underline{z}^2\right)^2},$$

where

$$T_i(z) = -\left(1-\alpha_j\right)^2 z^2 + 2\left(1-\alpha_j\right)^2 z\underline{z} - \left(1-\alpha_i\right)^2\underline{z}^2.$$

The sign of $\frac{\partial \lambda_i^*}{\partial z}$ depends on the sign of $T_i(z)$, which is concave in z and maximized at $z = \underline{z}$. Hence $T_i'(z) < 0$ for all $z > \underline{z}$, and recalling that $\alpha_j \leq \alpha_i < 1/2$,

$$\lim_{z \to \underline{z}} T_i(z) = \left(\alpha_i - \alpha_j\right)\left(2 - \alpha_i - \alpha_j\right)\underline{z}^2 \geq 0,$$

with strict inequality when $\alpha_j < \alpha_i$. Since $T_i(z) < 0$ for z sufficiently large as the coefficient of z^2 is negative, it follows that λ_i^* is first increasing (when $T_i(z) > 0$) and then decreasing with z (when $T_i(z) < 0$) if $\alpha_j < \alpha_i$ and is decreasing with z for all $z > \underline{z}$ if $\alpha_j = \alpha_i$.

Likewise,

$$\frac{\partial \lambda_j^*}{\partial z} = \frac{(1-\alpha_j)^2 T_j(z)}{(1-\alpha_i)\left(z^2(1-\alpha_j)^2 - (1-\alpha_i)^2 \underline{z}^2\right)^2},$$

where

$$T_j(z) = -(1-\alpha_j)^2 z^2 + 2(1-\alpha_i)^2 z\underline{z} - (1-\alpha_i)^2 \underline{z}^2.$$

The sign of $\frac{\partial \lambda_j^*}{\partial z}$ depends on the sign of $T_j(z)$, which is concave in z. Now, for all $z > \underline{z}$,

$$T_j'(z) = -2\left[(1-\alpha_j)^2 z - (1-\alpha_i)^2 \underline{z}\right] < 0.$$

Moreover, recalling that $\alpha_j \le \alpha_i$,

$$\lim_{z \to \underline{z}} T_j(z) = -(\alpha_i - \alpha_j)(2 - \alpha_i - \alpha_j)\underline{z}^2 \le 0.$$

Hence, $T_j(z) < 0$ for all $z > \underline{z}$, so λ_j^* is decreasing with z for all $z > \underline{z}$. ■

Proof of Proposition 5: (i) Differentiating (12) with respect to α_i, yields

$$\frac{\partial CS(\alpha_i, \alpha_j)}{\partial \alpha_i} = \frac{cH_i}{\left(z^2(1-\alpha_i)^2(1-\alpha_j)^2 - (1-\alpha_i\alpha_j)^2\right)^3},$$

where the sign of the derivative depends on the sign of

$$\begin{aligned} H_i \equiv\ & (1-\alpha_j(1-\alpha_j))(1-\alpha_i\alpha_j)^4 \\ & -(1-\alpha_j)^2(1-\alpha_i\alpha_j)^3(2(1-\alpha_i)(1-\alpha_j) + 1 - \alpha_i\alpha_j)z \\ & +2(1-\alpha_i)(1-\alpha_j)^2(1-\alpha_i\alpha_j)^2(2-\alpha_j(4-\alpha_i-\alpha_i\alpha_j))z^2 \\ & -2(1-\alpha_i)(1-\alpha_j)^3(1-\alpha_i\alpha_j)\left((1-\alpha_i)^2 + (1-\alpha_j)^2 + \alpha_j(1-\alpha_i)(1-\alpha_i(2-\alpha_j))\right)z^3 \\ & +(1-\alpha_i)^3(1-\alpha_j)^4(3+\alpha_j(1-\alpha_j)+\alpha_i(1-\alpha_j(3+\alpha_j)))z^4 \\ & -(1+\alpha_i)(1-\alpha_i)^3(1-\alpha_j)^6 z^5. \end{aligned}$$

Since the coefficient of z^5 is negative, $H_i < 0$ when z is sufficiently large. Using Mathematica, it turns out that $H_i < 0$ for all $z > \underline{z}$ and all $0 \le \alpha_j \le \alpha_i < 1/2$.[46]

(ii) Differentiating (12) with respect to α_j, yields

$$\frac{\partial CS(\alpha_i, \alpha_j)}{\partial \alpha_j} = \frac{cH_j}{\left(z^2(1-\alpha_i)^2(1-\alpha_j)^2 - (1-\alpha_i\alpha_j)^2\right)^3},$$

[46] The commad we use is Reduce[$H_i \ge 0$ && $z > \underline{z}$ && $0 \le \alpha_i < 1/2$&& $0 \le \alpha_j \le \alpha_i$,{$z, \alpha_i, \alpha_j$}], where $\underline{z} \equiv \frac{1-\alpha_i\alpha_j}{(1-\alpha_i)^2}$. The command returns the output "False", implying that, given the parameter restrictions, $H_i < 0$.

where H_j is similar to H_i, except that α_i and α_j switch roles. The sign of $\frac{\partial CS(\alpha_i,\alpha_j)}{\partial \alpha_j}$ is equal to the sign of H_j, where

$$\lim_{z\to \underline{z}} H_j = \frac{\alpha_i (\alpha_i - \alpha_j)^2 (1+\alpha_i-\alpha_j)(2-\alpha_i-\alpha_j)^2 (1-\alpha_i\alpha_j)^4}{(1-\alpha_i)^4} \geq 0,$$

with strict inequality for $\alpha_j < \alpha_i$. By contrast, $H_j < 0$ when z is sufficiently large because the coefficient of z^5 is negative.

(iii) Evaluated at $\alpha_j = \alpha_i = \alpha$,

$$H_j = -(1-\alpha)^7 \left(1 + (z(1-\alpha)+\alpha)\left(z\left(1-\alpha^2\right)+\alpha^2\right)\right)\left(z - \frac{1+\alpha}{1-\alpha}\right)^3 < 0, \tag{21}$$

where the inequality follows because by Assumption A5, $z > \underline{z} = \frac{1+\alpha}{1-\alpha}$ when $\alpha_j = \alpha_i = \alpha$. ■

Proof of Lemma 4: (i) Note from (13) that when $\alpha_j = \alpha_i$ and $m \to 1+\alpha_i$, $\lambda_i^* \to 0$, and

$$\lambda_j^* \to \frac{(1+\alpha_i)-(1+\alpha_j)}{(1+\alpha_i)^2 - (1+\alpha_i)(1+\alpha_j)} = \frac{1}{1+\alpha_i}.$$

If $\alpha_j = \alpha_i = \alpha$,

$$\lambda_i^* = \lambda_j^* = \frac{m-(1+\alpha)}{m^2-(1+\alpha)^2} = \frac{1}{m+1+\alpha},$$

which is approaching $\frac{1}{2(1+\alpha)}$ when $m \to 1+\alpha$. Moreover, note that $\lambda_i^* \to 0$ and $\lambda_j^* \to 0$ as $m \to \infty$.

(ii) By straightforward differentiation,

$$\frac{\partial\left(\lambda_i^*+\lambda_j^*\right)}{\partial \alpha_i} = -\frac{(m-(1+\alpha_j))^2}{(m^2-(1+\alpha_i)(1+\alpha_j))^2} < 0, \qquad \frac{\partial\left(\lambda_i^*+\lambda_j^*\right)}{\partial \alpha_j} = -\frac{(m-(1+\alpha_i))^2}{(m^2-(1+\alpha_i)(1+\alpha_j))^2} < 0.$$

(iii) By straightforward differentiation,

$$\frac{\partial \lambda_i^*}{\partial m} = \frac{-(m-(1+\alpha_i))^2 + (\alpha_i-\alpha_j)(1+\alpha_i)}{(m^2-(1+\alpha_i)(1+\alpha_j))^2},$$

and

$$\frac{\partial \lambda_j^*}{\partial m} = -\frac{(m-(1+\alpha_j))^2 + (\alpha_i-\alpha_j)(1+\alpha_j)}{(m^2-(1+\alpha_i)(1+\alpha_j))^2}.$$

Clearly, $\frac{\partial \lambda_j^*}{\partial m} < 0$ for all $m > 1+\alpha_i$ and all $\alpha_j \leq \alpha_i$. Moreover, $\frac{\partial \lambda_i^*}{\partial m} \lesseqgtr 0$ if $m \gtreqless \widehat{m} \equiv 1+\alpha_i+\sqrt{(\alpha_i-\alpha_j)(1+\alpha_i)}$. When $\alpha_j = \alpha_i$, $\frac{\partial \lambda_i^*}{\partial m} < 0$ for all $m > 1+\alpha_i$. ■

Proof of Proposition 6: Straightforward differentiation, using (13), yields

$$\frac{\partial CS(\alpha_i,\alpha_j)}{\partial \alpha_i} = \frac{B(m-(1+\alpha_j))M_i}{(m^2-(1+\alpha_i)(1+\alpha_j))^3}, \qquad \frac{\partial CS(\alpha_i,\alpha_j)}{\partial \alpha_j} = \frac{B(m-(1+\alpha_i))M_j}{(m^2-(1+\alpha_i)(1+\alpha_j))^3},$$

where

$$M_i \equiv -(1+\alpha_i)(1+\alpha_j)(1-\alpha_j\sigma)+(1+\alpha_j)(2+(1-\alpha_i)\sigma)m-(1+(2+\alpha_j)\sigma)m^2+\sigma m^3,$$

and M_j is similar to M_i, except that α_i and α_j switch roles. The signs of $\frac{\partial CS(\alpha_i,\alpha_j)}{\partial\alpha_i}$ and $\frac{\partial CS(\alpha_i,\alpha_j)}{\partial\alpha_j}$ depend on the signs of M_i and M_j, which in turn depend on 4 parameters: σ, m, α_i, and α_j, where $m>1+\alpha_i$, $0<\sigma\le\alpha_j\le\alpha_i<1/2$.

(i) M_i and M_j are positive cubic functions of m (the coefficient of m^3 is positive); hence $M_i>0$ and $M_j>0$ for m sufficiently large. If $\sigma\to 1/2$, then $\alpha_j,\alpha_i\to 1/2$; since $m>1+\alpha_i\to 3/2$ by Assumption A5, $M_i=M_j\to(m-3/2)^3/2>0$. Hence, in both cases, $\frac{\partial CS(\alpha_i,\alpha_j)}{\partial\alpha_i}>0$ and $\frac{\partial CS(\alpha_i,\alpha_j)}{\partial\alpha_j}>0$.

(ii) Note that,

$$\lim_{m\to 1+\alpha_i} M_i=-(1+\alpha_i)(\alpha_i-\alpha_j)(1-(\alpha_i-\alpha_j))\sigma\le 0.$$

Moreover, $\lim_{m\to 1+\alpha_i}\frac{\partial M_i}{\partial m}=-(\alpha_i-\alpha_j)(2-(1+3\alpha_i)\sigma)<0$, where the inequality follows because $\sigma\le\alpha_i<1/2$ implies that $2-(1+3\alpha_i)\sigma>2-(1+3/2)/2>0$. Since M_i is a positive cubic, it follows that $M_i<0$ for values of m not too much above $1+\alpha_i$, and $M_i>0$ otherwise.

Likewise,

$$\lim_{m\to 1+\alpha_i} M_j=(1+\alpha_i)(\alpha_i-\alpha_j)(1+\sigma)\ge 0,$$

and $\lim_{m\to 1+\alpha_i}\frac{\partial M_j}{\partial m}=(1+\alpha_i)(\alpha_i-\alpha_j)\sigma>0$, $\lim_{m\to 1+\alpha_i}\frac{\partial^2 M_j}{\partial m^2}=-2(1+(1-2\alpha_i)\sigma)<0$, where the inequality follows because $\sigma\le\alpha_i<1/2$. Since M_j is a positive cubic, $M_j>0$ for m not too far above $1+\alpha_i$ and it is possible that $M_j>0$ for all m if M_j has only one root (rather than 3).

(iii) If $\sigma\to 0$ and $\sigma m\to 0$,

$$M_i\to -(m-(1+\alpha_j))^2-(\alpha_i-\alpha_j)(1+\alpha_j)<0,$$

implying that $\frac{\partial CS(\alpha_i,\alpha_j)}{\partial\alpha_i}<0$, and

$$M_j\to -(m-(1+\alpha_i))^2+(\alpha_i-\alpha_j)(1+\alpha_i),$$

which is negative if $m>\widehat{m}\equiv 1+\alpha_i+\sqrt{(\alpha_i-\alpha_j)(1+\alpha_i)}$ and positive if $m<\widehat{m}$. Hence, $\frac{\partial CS(\alpha_i,\alpha_j)}{\partial\alpha_j}<0$ if $m>\widehat{m}$ and $\frac{\partial CS(\alpha_i,\alpha_j)}{\partial\alpha_j}>0$ if $m<\widehat{m}$.

Straightforward differentiation reveals that

$$\frac{\partial M_j}{\partial\sigma}=\alpha_i(1+\alpha_i)(1+\alpha_j)+(1+\alpha_i)(1-\alpha_j)m-(2+\alpha_i)m^2+m^3.$$

Using Mathematica, the derivative is strictly positive for $m > 1 + \alpha_i$ and all $0 \leq \alpha_j \leq \alpha_i < 1/2$.[47] Hence, as σ increases, M_j and hence $\frac{\partial CS(\alpha_i,\alpha_j)}{\partial \alpha_j}$ are more likely to be positive.

(iv) Evaluated at $\alpha_i = \alpha_j = \alpha \geq \sigma$, $M_i = M_j \equiv M$, where

$$M = \sigma \left(m - (1+\alpha)\right)^2 \left(m - \left(\frac{1}{\sigma} - \alpha\right)\right).$$

Hence, $\frac{\partial CS(\alpha_i,\alpha_j)}{\partial \alpha_i} = \frac{\partial CS(\alpha_i,\alpha_j)}{\partial \alpha_j} \gtreqless 0$ as $m \gtreqless \frac{1}{\sigma} - \alpha$. ■

Proof of Proposition 7: Recall that when the PCO structure is symmetric, $k > \pi(\alpha)(1+\alpha)$; since $\pi(\alpha) = \pi^m$ as $\alpha \geq \sigma$, (15) implies that $\lambda(\alpha) \equiv \frac{\pi^m}{k + (1+\alpha)\pi^m} > \frac{\pi^m}{2k}$. Hence,

$$\begin{aligned}
\lambda(\alpha) - \lambda^{**} &> \frac{\pi^m}{2k} - \frac{\hat{\pi} - \pi_c^m}{2k + \hat{\pi} + \hat{\pi}_c - \pi_c^m - \pi^m} \\
&= \frac{\pi^m \left(2k + \hat{\pi} + \hat{\pi}_c - \pi_c^m - \pi^m\right) - 2k\left(\hat{\pi} - \pi_c^m\right)}{2k\left(2k + \hat{\pi} + \hat{\pi}_c - \pi_c^m - \pi^m\right)} \\
&= \frac{2k\left(\pi^m - \hat{\pi} + \pi_c^m\right) - \pi^m\left(\pi^m - \hat{\pi} + \pi_c^m - \hat{\pi}_c\right)}{2k\left(2k + \hat{\pi} + \hat{\pi}_c - \pi_c^m - \pi^m\right)} \\
&= \frac{\left(2k - \pi^m\right)\left(\pi^m - \hat{\pi} + \pi_c^m\right) + \pi^m \hat{\pi}_c}{2k\left(2k + \hat{\pi} + \hat{\pi}_c - \pi_c^m - \pi^m\right)} > 0,
\end{aligned}$$

where the last inequality follows because by Assumption A4, $k > \pi^m \geq \hat{\pi} \equiv \pi(\hat{p})$.

Expected consumer surplus under semicollusion is given by

$$CS^{**} = \left(\lambda^{**}\right)^2 S\left(p^m\right) + 2\lambda^{**}\left(1 - \lambda^{**}\right) S\left(\widehat{p}\right) + \left(1 - \lambda^{**}\right)^2 S\left(p_c^m\right),$$

which is analogous to (16). Comparing CS^{**} with (16) and noting that $p_c^m > \widehat{p} > p^m > c > 0$, it is clear that $CS(\alpha) > CS^{**}$. ■

Proof of Proposition 8: In the text we show that $\lambda(\alpha) > \lambda^{RJV}$. Turning to the unit demand case, substituting $\pi^m = B$ and $m \equiv k/B$ in (15) and (17), the equilibrium levels of investments under PCO are under RJV are given by

$$\lambda(\alpha) = \frac{1}{m + 1 + \alpha}, \qquad \lambda^{RJV} = \frac{1}{m+2}.$$

Note that $\lambda(\alpha) > \lambda^{RJV}$ as $\alpha < 1/2$.

Using (14), expected consumer surplus under PCO and under RJV are given by

$$CS(\alpha) = \frac{B\left(1 + \sigma(\alpha + m)^2\right)}{(m + 1 + \alpha)^2}, \qquad CS^{RJV} = \frac{B\left(1 + \sigma(1+m)^2\right)}{(m+2)^2}.$$

[47] The commad we use is Reduce[$D[M_j, \sigma] \leq 0$ && $m > 1 + \alpha_i$ && $0 \leq \alpha_j \leq \alpha_i < 1/2, \{\alpha_i, \alpha_j.m\}$]. The command returns the output "False", implying that, given the parameter restrictions, $\frac{\partial M_j}{\partial \sigma} > 0$.

Notice that $CS(1) = CS^{RJV}$ and

$$CS'(\alpha) = \frac{2B\left(\sigma(\alpha+m)^2 - 1\right)}{(m+1+\alpha)^3}.$$

If m is sufficiently large, $CS'(\alpha) > 0$, so $CS(\alpha) < CS^{RJV}$ for all $0 \leq \alpha < 1/2$. By contrast, if $m \to 1+\alpha$ (the lower bound on m by Assumption A5), then

$$\lim_{m \to 1+\alpha} CS'(\alpha) \to \frac{B(\sigma(2\alpha+1) - 1)}{4(1+\alpha)^3} < \frac{B(2\sigma - 1)}{4(1+\alpha)^3} < 0,$$

where the first inequality follows because $\alpha < 1/2$ and the second follows because $0 < \sigma < 1/2$ by Assumption A3. Hence, if $m \to 1+\alpha$, $CS(\alpha) > CS^{RJV}$ for all $0 \leq \alpha < 1/2$. By continuity this is also the case when m is sufficiently small. ■

Proof of Proposition 9: Note that,

$$\begin{aligned} \lambda(\alpha) - \lambda^m &= \frac{\pi_c^m k - \alpha \pi^m (\pi^m - \pi_c^m)}{(k + (1+\alpha)\pi^m)(k + \pi^m - \pi_c^m)} \\ &= \frac{\pi^m (\pi^m - \pi_c^m)}{(k + (1+\alpha)\pi^m)(k + \pi^m - \pi_c^m)} \left[\frac{\pi_c^m k}{\pi^m (\pi^m - \pi_c^m)} - \alpha\right]. \end{aligned}$$

Since by definition $\pi^m > \pi_c^m$, $\lambda(\alpha) > \lambda^m$ if and only if (18) holds. Consumer surplus is higher under PCO since prices under full merger are either p^m or $p_c^m > p^m$, whereas under PCO they are lower and equal to p^m, $p(\alpha) \leq p^m$, or c. ■

10 References

Airline Business (1998), "Airlines with Ownership by Other Carriers," June 1998, News, **48**.

Alley W. (1997), "Partial Ownership Arrangements and Collusion in the Automobile Industry," *Journal of Industrial Economics*, **45**, 191-205.

Amundsen E. and L. Bergman (2002), "Will Cross-Ownership Re-Establish Market Power in the Nordic Power Market," *The Energy Journal*, **23**, 73-95.

Antón M., F. Ederer, M. Giné, and M. Schmalz (2021), "Innovation: The Bright Side of Common Ownership?" https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3099578.

Antón M., F. Ederer, M. Giné, and M. Schmalz (2022), "Common Ownership, Competition, and Top Management Incentives," mimeo, https://florianederer.github.io/common_ownership.pdf

Aoki R. and Y. Spiegel (2009), "Pre-Grant Patent Publication and Cumulative Innovation," *International Journal of Industrial Organization*, **27**, 333–345.

Azar J., M. Schmalz, and I. Tecu (2018), "Anti-competitive Effects of Common Ownership," *The Journal of Finance*, **73(4)**, 1513-1565.

Backus M., C. Conlon, and M. Sinkinson (2021), "Common Ownership and Competition in the Ready-to-Eat Cereal Industry," Discussion paper, National Bureau of Economic Research.

Backus M., C. Conlon, and M. Sinkinson (2021), "Common Ownership in America: 1980-2017," *American Economic Journal: Microeconomics*, 13(3), 273-308.

Banal-Estañol A., J. Seldeslachts, and X. Vives, (2020), "Diversification, Common Ownership, and Strategic Incentives," *AEA Papers and Proceedings*, **110**, 561–564.

Baumol W. and J. Ordover (1994), "On the Perils of Vertical Control by a Partial Owner of a Downstream Enterprise," *Revue d'économie Industrielle*, **69** (3e trimestre), 7-20.

Bayona A. and A. L. López, (2018), "Silent Financial Interests and Product Innovation," *Economics Letters*, **170**, 109–112.

Bolle F. and W. Güth (1992), "Competition Among Mutually Dependent Sellers," *Journal of Institutional and Theoretical Economics,* **148**, 209-239.

Bresnahan T. and S. Salop (1986), "Quantifying the Competitive Effects of Production Joint Ventures," *International Journal of Industrial Organization*, **4**, 155-175.

Brito D., R. Ribeiro, and H. Vasconcelos (2014), "Measuring Unilateral Effects in Partial Horizontal Acquisitions." *International Journal of Industrial Organization*, **33**, 22–36.

Brod A. and R. Shivakumar (1999), "Advantageous Semi-Collusion," *The Journal of Industrial Economics*, **47(2)**, 221-230.

Burnside A. and A.Kidane (2020), "Common Ownership: An EU Perspective," *Journal of Antitrust Enforcement*, **8**, 456–510.

Choi J. P. (1993), "Cooperative R&D with Product Market Competition," *International Journal of Industrial Organization,* **11**, 553–571.

Dietzenbacher E., B. Smid, and B. Volkerink (2000), "Horizontal Integration in the Dutch Financial Sector," *International Journal of Industrial Organization,* **18**, 1223-1242.

European Commission (2013), "Towards More Effective EU Merger Control," Commission Staff Working Document, available at http://ec.europa.eu/competition/consultations/2013_merger_control/

Federico, G., G. Langus, and T. Valletti (2018), "Horizontal Mergers and Product Innovation," *International Journal of Industrial Organization,* **59**, 1–23.

Fershtman C. and N. Gandal (1994), "Disadvantageous Semicollusion," *International Journal of Industrial Organization*, **12**, 141-154.

Flath D. (1991), "When is it Rational for Firms to Acquire Silent Interests in Rivals?," *International Journal of Industrial Organization,* **9**, 573-583.

Flath D. (1992), "Horizontal Shareholding Interlocks," *Managerial and Decision Economics,* **13**, 75-77.

Gilo D. (2000), "The Anticompetitive Effect of Passive Investment," *Michigan Law Review*, **99**, 1-47.

Gilo D., Y. Moshe, and Y. Spiegel (2006), "Partial Cross Ownership and Tacit Collusion," *RAND Journal of Economics*, **37**, 81-99.

Greenlee P., and A. Raskovich (2006), “Partial Vertical Ownership,” *European Economic Review,* **50(4)**, 1017–1041.

Heim S., K. Hüschelrath, U. Laitenberger, and Y. Spiegel (2022), “The Anticompetitive Effect of Minority Share Acquisitions: Evidence from the Introduction of National Leniency Programs,” *American Economic Journal: Microeconomics,* **14**, 366-410.

Hunold M. and K. Stahl (2016), “Passive Vertical Integration and Strategic Delegation,” *RAND Journal of Economics*, **42**, 891-913.

Jullien B. and Y. Lefouili (2018), “Mergers and Investments in New Products,” TSE Working Paper, n. 18-949.

Kamien K., E. Muller, and I. Zang (1992), “Research Joint Ventures and R&D Cartels,” *The American Economic Review*, **82(5)**, 1293-1306.

Levy N, Y. Spiegel, and D. Gilo (2018), “Partial Vertical Integration, Ownership Structure and Foreclosure,” *American Economics Journal: Microeconomics*, **10(1)**, 132-180.

Lewellen K. and M. Lowry (2021), “Does Common Ownership Really Increase Firm Coordination?” *Journal of Financial Economics*, **141(1)**, 322-344.

Li Y. and J. Zhang (2021), “Product Positioning with Overlapping Ownership,” *Economics Letters*, **208**, 110058.

López Á. and X. Vives (2019), “Overlapping Ownership, R&D Spillovers, and Antitrust Policy,” *Journal of Political Economy*, 127, 2394-2437.

Malueg D. (1992), “Collusive Behavior and Partial Ownership of Rivals,” *International Journal of Industrial Organization,* **10**, 27-34.

Motta M. and Tarantino E. (2021), “The Effect of Horizontal Mergers, when Firms Compete in Prices and Investments,” *International Journal of Industrial Organization*, **78**, 102774.

Nain A. and Y. Wang (2016), “The Product Market Impact of Minority Stake Acquisitions,” *Management Science*, **64(2)**, 825-844.

Nitta K. (2008), “Corporate Ownership Structure in Japan—Recent Trends and Their Impact,” Financial Research Group, Available at http://www.nli-research.co.jp/english/economics/2008/eco080331.pdf

OECD (2012), "The Role of Efficiency Claims in Antitrust Proceedings," Policy Roundtables, Available at http://www.oecd.org/competition/EfficiencyClaims2012.pdf.

Ono H., T. Nakazato T., C. Davis, and W. Alley (2004), "Partial Ownership Arrangements in the Japanese Automobile Industry; 1990-2000, *Journal of Applied Economics*, **12(2)**, 355-367.

Reiffen D. (1998), "Partial Ownership and Foreclosure: An Empirical Analysis," *Journal of Regulatory Economics*, **13(3)**, 227-44.

Reitman D. (1994), "Partial Ownership Arrangements and the Potential for Collusion," *Journal of Industrial Economics,* **42**, 313-322.

Reynolds R. and B. Snapp (1986), "The Competitive Effects of Partial Equity Interests and Joint Ventures," *International Journal of Industrial Organization,* **4**, 141-153.

Rock E. and D. Rubinfeld (2018), "Antitrust for Institutional Investors," *Antitrust Law Journal*, **82(1)**, 221-278.

Schinkel M.P. and Y. Spiegel (2017), "Can Collusion Promote Sustainable Production?," *International Journal of Industrial Organization,* **53**, 371-398.

Shelegia S. and Y. Spiegel (2012), "Bertrand Competition when Firms Hold Passive Ownership Stakes in One Another," *Economics Letters,* **114**, 136-138.

Spiegel (2013), "Backward Integration, Forward Integration, and Vertical Foreclosure," CEPR Discussion Paper No. DP9617

Stenbacka R. and G. Van Moer (2022), "Overlapping Ownership and Product Innovation," Mimeo

Vives X. (2020), "Common Ownership, Market Power, and Innovation," *International Journal of Industrial Organization*, **70**, 102528.