THE JOURNAL OF INDUSTRIAL ECONOMICS 0022-1821
Volume LXXII December 2024 No. 4

HORIZONTAL PARTIAL CROSS OWNERSHIP AND INNOVATION*

SANDRO SHELEGIA[†,§]
YOSSI SPIEGEL[‡,§,¶]

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* due to its effect on price competition and hence on the marginal benefit from investment, as well as a *cannibalization effect* which arises 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, their degree of symmetry, the size of the innovation, its marginal cost, and whether it is drastic or not.

I. 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]; Ono *et al.* [2004]), the Dutch Financial Sector (Dietzenbacher *et al.* [2000]), the Nordic power market (Amundsen and Bergman [2002]), and the global steel industry (Gilo *et al.* [2006]). Nitta [2008] reports that cross-shareholding, that is, situations in which two firms 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 1990s, and remained above 8.5% by 2006.[1]

*For helpful discussions and comments we thank John Asker, Patrick Rey, Luca Sandrini, Maarten Pieter Schinkel, and seminar participants at the 2014 International Workshop on Evolution of Standards and Technology in Hitotsubashi University, the 2015 and 2023 MaCCI conferences in Mannheim, the 2017 IIOC conference in Boston, the 2022 Berlin IO day, the 2023 Cresse conference in Rhodes, the 2024 SweRIE conference in Stockholm, NYU, and University of Zurich. Yossi Spiegel thanks the Jeremy Coller Foundation and the Henry Crown Institute of Business Research in Israel for financial support.

[†]Authors' affiliations: Department of Economics and Business, Universitat Pompeu Fabra, BSE, Bacelona, Spain.
e-mail: Sandro.Shelegia@upf.edu
[‡]Coller School of Management, Tel Aviv University, Tel Aviv, Israel.
e-mail: spiegel@post.tau.ac.il
[§]CEPR, Paris, France.
[¶]ZEW, Mannheim, Germany.

[1] A related, but in general distinct, phenomenon is common ownership: cases where firms have common shareholders. Common ownership has attracted a lot of attention recently and there

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 VII 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 early literature on PCO has shown that horizontal PCO among rival firms can soften competition in the Cournot model (Reynolds and Snapp [1986]; Flath [1991, 1992]; Bolle and Güth [1992]; Reitman [1994]; Dietzenbacher *et al.* [2000]) or the Bertrand model (Shelegia and Spiegel [2012]), and can facilitate collusion in an infinitely repeated Cournot model (Malueg [1992]) or infinitely repeated Bertrand model (Gilo *et al.* [2006]).[3] Moreover, vertical PCO (with and without control) among upstream and downstream firms can lead to upstream and downstream foreclosure (e.g., Baumol and Ordover [1994]; Reiffen [1998]; Greenlee and Raskovich [2006]; Spiegel [2013]; Hunold and Stahl [2016]; Levy *et al.* [2018]) and can also raise prices and harm consumers without foreclosure (Flath [1989]; Fiocco [2016];

is a lively debate about its competitive implications. See for instance, Azar *et al.* [2018], Antón *et al.* [2022], Backus *et al.* [2021a, 2021b], and Banal-Estañol *et al.* [2020]. Huse *et al.* [2024] study both cross and common ownership in the global automobile industry, over the period 2007–2021, and find that common-ownership links constitute between 31% and 39% of the equity ownership of automobile manufacturers, while cross-ownership links amount to 6% and 9%; however, accounting for cross-ownership links can increase the average weight assigned by managers to the profit of competitors by between 33% and 68%.

[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 decisions. See Burnside and Kidane [2020]. Interestingly, the commission approved the two mergers subject to divestitures of major businesses and assets, including R&D organizations, 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 2351 and case M. 8084 – Bayer/Monsanto.

[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.

Hunold and Schlütter [2021]).[4] The picture that emerges from this literature is that PCO generally has adverse unilateral and coordinated competitive effects.[5] This picture has received some empirical support (e.g., Dietzenbacher *et al.* [2000]; Brito *et al.* [2014]; Nain and Wang [2016]; Heim *et al.* [2022]).

A more recent literature, which we discuss in detail below, studies the effects of horizontal cross and common ownership (the two are often labeled "overlapping ownership") on innovation. It shows that overlapping ownership may promote investments in innovation and thereby benefit consumers. Our paper belongs to this literature. 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.

Importantly, we allow firms to hold asymmetric PCO stakes in each other and 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.[6] In particular, we assume that firm i's stake in firm j, α_i, is (weakly) larger than firm j's stake in firm i, α_j, and establish sufficient conditions for an increase in α_i (which 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. The results are driven by two effects. The first effect, which we term the "price effect" of PCO, arises because an increase in α_i induces firm i to be softer as it internalizes part of the negative externality it imposes on firm j. This boosts firm j's profit, but lowers firm i's profit, and consequently firm j has a stronger incentive to invest, whereas firm i has a weaker incentive to invest.[7] At the same time, investment cannibalizes the rival's profit, so an increase in α_i weakens firm i's incentive to invest. We term this the *cannibalization effect* of PCO. We explore how the two effects play out in equilibrium and explore the implications for consumers' welfare.

[4] PCO among vertically related firms can also have pro-competitive effects. For instance, Flath [1989] shows that partial forward integration can relax the double marginalization problem and thereby lower prices and benefit consumers. By contrast partial backward integration may have the opposite effect.

[5] Ma *et al.* [2021] consider a Cournot setting where firms have asymmetric marginal costs and one firm invests in rivals. They show that while this one-sided PCO softens competition and hence harms consumers, it can also enhance total welfare if the acquiring firm has a high marginal cost. The reason is that the acquirer cuts its output level whereas rivals expand, so overall production becomes more efficient.

[6] We study the effect on consumer surplus as most antitrust agencies, including the US, the EU, and the UK, use the consumer welfare standard (see OECD [2012], pp. 26-27).

[7] In our Bertrand setting, marginal cost is either c if a firm does not innovate or 0 if it does. Absent PCO, a firm earns a profit only if its investment succeeds and the rival's investment fails. The sole innovator then, charges c and (by a tie-breaking rule) serves the entire market. 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 c). As a result, firm i can raise its price in equilibrium above c. Firm j then shares some of this profit due to its PCO stake.

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 in our model directly on the size of the PCO stakes.

When the PCO stakes are sufficiently small to ensure that the innovation is non-drastic, an increase in the larger stake, α_i (which leads to a more asymmetric PCO structure), unambiguously harms consumers. By contrast, an increase in the smaller stake, α_j (which makes the PCO structure more symmetric), benefits consumers when the cost of innovation is sufficiently low, but harms consumers otherwise.

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 give rise to a price effect. An increase in α_i or α_j can benefit consumers in this case if the cost of innovation is sufficiently high or the lower bound on the PCO stakes is sufficiently large. When the cost of innovation is small, an increase in α_i surely harms consumers, but an increase in α_j can still benefit consumers. We also show that in the neighborhood of a symmetric PCO structure, an increase in α_i or α_j benefits consumers when the cost of innovation is large and harms them when it is low.

Our analysis highlights the fact that PCO may benefit consumers by softening price competition which in turn may promote innovation. One may then wonder how PCO performs relative to other arrangements 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 consider symmetric PCO which is sufficiently large to ensure that the innovation is drastic, and show that it leads to more investment than semi-collusion or an RJV and may also lead to more investment than a full merger. We also show that PCO benefits consumers more than semicollusion or a full merger and can also benefit them more than an RJV.

The rest of the paper is organized as follows. In Section II, we review related literature, and in Section III, we present the model and characterize the equilibrium absent PCO. In Section IV, we characterize the equilibrium with PCO, and in Section V, study the welfare implications of PCO. In Section VI, we consider the unit demand case in order to shed more light on the welfare implications of PCO. In Section VII, we compare PCO with semicollusion, RJV's, and full mergers. In Section VIII we conclude. All proofs are in the Appendix.

II. RELATED LITERATURE

As mentioned above, our paper belongs to the small literature that studies the effects of horizontal overlapping ownership (both cross and common ownership) on innovation. Unlike the early literature which studies the effects of overlapping ownership on price or quantity competition, here firms first invest in R&D and only then compete in the product market. 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.[8]

Stenbacka and Van Moer [2023] consider a duopoly model with stochastic product innovation, but like López and Vives [2019] assume that the overlapping ownership structure is symmetric. They show that an increase in the symmetric level 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 and firms cut output, they also have a weaker incentive to invest.

Bayona and López [2018] consider a Hotelling duopoly model with possibly asymmetric common ownership, but consider a deterministic quality-enhancing R&D process.[9] 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.[10] Moreover, consumer and total surplus may increase or decrease when only one controlling shareholder holds a stake in the rival, but they are always lower under symmetric common ownership.[11]

Antón *et al.* [2021] consider an n-firm Cournot oligopoly model with differentiated products and linear demand functions with possibly asymmetric

[8] 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.

[9] 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.

[10] 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 to the transportation cost is sufficiently low.

[11] Li and Zhang [2021] study a related model, 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.

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 patent citations of publicly listed US corporations.[12]

Ghosh and Morita [2017] also consider an n-firm Cournot oligopoly, but with homogeneous products. All firms have the same constant marginal cost c, except for firm 1 whose marginal cost is $c-x$. They show that if firm 1 acquires a sufficiently large ownership stake in firm 2, it has an incentive to transfer knowledge to firm 2, which lowers firm 2's marginal cost to $c-x$.[13] The PCO of firm 1 in firm 2 softens competition between firms 1 and 2, but may induce other firms to become more aggressive. They find that an endogenously determined level of PCO of firm 1 in firm 2 can increase total surplus and even consumer surplus.

Our paper differs from the above papers in that we consider a fairly simple model of product market competition (albeit under PCO, the Bertrand model is less simple than one may think) but allow the ownership stakes that the two firms hold in each other to be asymmetric and consider a stochastic R&D process which could be viewed as either process or product innovation. Our modeling choice is motivated by the following considerations. 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 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. 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 equal increase in both stakes always harms consumers, whereas a unilateral increase in only the smaller stake can enhance welfare if the marginal cost of investment in R&D is sufficiently small.

[12] 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] find that mergers of financial institution caused substantial and lasting increases in common ownership, but they find no significant effects of common ownership on either firm profitability or firm R&D.

[13] While their paper is not, strictly speaking, about innovation, one can view the technology transfer as a cost-reducing "innovation" because it allows firm 2 to lower its cost.

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], Antón *et al.* [2021], and Ghosh and Morita [2017]. This difference is important because a deterministic R&D process leads to lower costs (or higher quality) and unambiguously benefits consumers. By contrast, in the presence of PCO, 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 *et al.* [2018]; Jullien and Lefouili [2018]; 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.

III. 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. The parameter c then reflects the size of the innovation.[14] We assume that λ_i is a choice variable for the firm and refer to it as "firm i's investment level."[15] 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 two 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.[16] If firms are equally efficient, consumers randomize between them.

[14] Alternatively, we can normalize marginal cost to 0 and assume that if the 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 or product innovation. While the two formulations are isomorphic, we will use the process innovation interpretation for the sake of concreteness.

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

[16] 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.

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 quasi-concave and has a unique maximizer p^m

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

A3 $c < p^m < 2c$

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.[17] As we shall see below, with PCO, the innovation becomes drastic for sufficiently large PCO stakes. 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).

Our Bertrand setting is a special case of the Aoki and Spiegel [2009] model of stochastic R&D competition, where the stage 2 profit of each firm is π_{yy} if both firms innovate successfully, 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.[18] In our Bertrand setting, $\pi_{yn} = \pi(p) > 0 = \pi_{yy} = \pi_{nn} = \pi_{ny}$. We chose to work with this setting because in the more general setting, PCO affects all stage 2 profits, π_{yn}, π_{yy}, π_{nn}, and π_{ny}, so the model becomes too complex to analyze, especially since we focus on asymmetric PCO structure and cannot invoke symmetry to simplify the analysis.[19]

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.[20] 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

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

[17] For example, when demand is linear and given by $Q = A - p$, the monopoly price when the firm innovates is $A/2$, so Assumption A3 implies that $c < A/2 < 2c$ (in particular if the innovating firm charges $A/2$ the rival will be able to profitably undercut it).

[18] Jullien and Lefouili [2018] and Stenbacka and Van Moer [2023] consider a similar setting, where $\pi_{ny} = \pi_{nn} = 0$.

[19] Aoki and Spiegel [2009] show that so long as $\pi_{yn} + \pi_{ny} > \pi_{yy} + \pi_{nn}$, the R&D investments of the two firms are strategic substitutes, which is also true in our Bertrand setting, where $\pi_{yn} > 0 = \pi_{ny} = \pi_{yy} = \pi_{nn}$.

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

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.[21]

IV. 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

$$V_i = \frac{\Pi_i + \alpha_i \Pi_j}{1 - \alpha_i \alpha_j}, \qquad V_j = \frac{\Pi_j + \alpha_j \Pi_i}{1 - \alpha_i \alpha_j}. \tag{2}$$

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 $\left(1 - \alpha_j\right) V_i + \left(1 - \alpha_i\right) V_j = \Pi_i + \Pi_j$.

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 $\beta_i V_i$. 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.

Before proceeding, it is worth clarifying the difference between PCO and common ownership. The latter arises when the stakes in rivals are held by the controlling shareholders of the two firms rather than the firms themselves. To see why it matters, suppose that firms do not hold stakes in one another, but

[21] 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 $\left(\lambda_i, \lambda_j\right)$ 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 $\left(\lambda_i, \lambda_j\right)$ 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 $\left(\lambda_i, \lambda_j\right)$ space.

the controlling shareholder of firm i holds a stake β_{ij} in firm j (in addition to his stake β_i in firm i) and the controlling shareholder of firm j holds a stake β_{ji} in firm i. Then, the objective functions of the two shareholders (who choose the firms' strategies) are given by

(3)
$$V_i = \beta_i \Pi_i + \beta_{ij}\Pi_j = \beta_i\left(\Pi_i + \frac{\beta_{ij}}{\beta_i}\Pi_j\right), \quad V_j = \beta_j \Pi_j + \beta_{ji}\Pi_i = \beta_j\left(\Pi_j + \frac{\beta_{ji}}{\beta_j}\Pi_i\right).$$

With only two firms, (2) and (3) are essentially equivalent.[22] This equivalence however generally breaks down when there are three firms or more and the ownership stakes are asymmetric. Then, as Gilo *et al.* [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.[23]

It is also worth pointing out that, if in addition to PCO, the controlling shareholder of firm i holds a stake β_{ij} in firm j, the shareholder's objective function specified in (2) becomes $\beta_i V_i + \beta_{ij} V_j = \left(\beta_i + \alpha_j \beta_{ij}\right)\left(\Pi_i + \frac{\alpha_i\beta_i+\beta_{ij}}{\beta_i+\alpha_j\beta_{ij}}\Pi_j\right)$. Although the weight assigned to firm j's profit now exceeds α_i, conceptually nothing else changes, which is why we set $\beta_{ij} = \beta_{ji} = 0$.

As in the standard Bertrand model, when both firms innovate in stage 1 or both fail, competition drives their values to 0. To see why, suppose that in stage 2 firm i charges a price p. If firm j undercuts p, its profit approaches $\Pi_j = \left(p - \hat{c}\right) Q(p)$, where $\hat{c} = 0$ if both firms innovate and $\hat{c} = c$ if both firms fail. Since $\Pi_i = 0$, firm j's value is $V_j = \frac{(p-\hat{c})Q(p)}{1-\alpha_i\alpha_j}$. If firm j sets a price above p, firm i serves the entire market, so $\Pi_j = 0$ and $\Pi_i = \left(p - \hat{c}\right) Q(p)$, in which case, $V_j = \frac{\alpha_j(p-\hat{c})Q(p)}{1-\alpha_i\alpha_j}$. 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 $V_j = \frac{\alpha_j p Q(p)}{1-\alpha_i\alpha_j}$. Firm j will not undercut firm i if

$$\frac{(p-c)\,Q(p)}{1-\alpha_i\alpha_j} \le \frac{\alpha_j p Q(p)}{1-\alpha_i\alpha_j}, \qquad \Rightarrow \qquad p \le \frac{c}{1-\alpha_j}.$$

[22] While common and cross ownership are isomorphic in our duopoly setting, we will refer to links between firms as cross ownership for the sake of concreteness.

[23] Some early papers on PCO (e.g., Reynolds and Snapp [1986]; Reitman [1994]) did not make the distinction between PCO and common ownership and while they claim that the stakes in rivals are held by firms, the objective functions they consider are essentially similar to (3), implying that what they actually study is common ownership rather than PCO.

When both firms charge the same price, then by assumption, consumers buy from the more efficient firm; hence firm i can charge $\frac{c}{1-\alpha_j}$ and serve the entire market. However, if $\alpha_j \geq \sigma \equiv \frac{p^m - c}{p^m}$, then $\frac{c}{1-\alpha_j} \geq p^m$, where p^m is the monopoly price when marginal cost is 0, so the innovation becomes drastic in the sense that firm i is better off charging p^m.[24] Assumption A3 guarantees that $0 < \sigma < 1/2$.[25]

Proposition 1 in Shelegia and Spiegel [2012] implies that when firm i has a lower cost than firm j, there exist 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 $p \in \left[0, p\left(\alpha_j\right)\right]$, where

$$p\left(\alpha_j\right) \equiv \begin{cases} \frac{c}{1-\alpha_j}, & \alpha_j < \sigma, \\ p^m, & \alpha_j \geq \sigma. \end{cases} \tag{4}$$

Of these Nash equilibria, the only equilibrium in which firm j does not play a weakly dominated strategy is the one where both firms charge $p\left(\alpha_j\right)$ and firm i serves the entire market.[26] In what follows, we will restrict attention to this equilibrium. The stage 2 profit of firm i in this equilibrium, as a function of firm j's stake, α_j, is

$$\pi\left(\alpha_j\right) \equiv \pi\left(p\left(\alpha_j\right)\right) = \begin{cases} \frac{c}{1-\alpha_j} Q\left(\frac{c}{1-\alpha_j}\right), & \alpha_j < \sigma, \\ \pi^m, & \alpha_j \geq \sigma. \end{cases} \tag{5}$$

The corresponding stage 2 profit of firm j is $\pi\left(\alpha_i\right)$ (note that the stage 2 profit of each firm depends on the rivals's stake in the firm).

Four comments are now in order. First, when firm j does not hold a stake in firm i, that is, $\alpha_j = 0$, then $p(0) = c$ and $\pi(0) = cQ(c)$, exactly as in the traditional Bertrand model.

[24] Note that 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.

[25] If Assumption A3 fails and $p^m < c$, then $\sigma < 0$, so the innovation is always drastic even without PCO; if $p^m > 2c$, then $\sigma > 1/2$, so the innovation cannot be drastic because by assumption, $\alpha_j \leq \alpha_i < \frac{1}{2}$.

[26] To see why, consider an equilibrium where $p_i = p_j = p^* \in \left[0, p\left(\alpha_j\right)\right)$. Since firm i's cost is 0 whereas firm j's cost is c, consumers buy from firm i; firm j makes no sales and its value is $\frac{\alpha_j p^* Q(p^*)}{1-\alpha_i \alpha_j}$. If firm i deviates upward from p^*, firm j serves the entire market and its value becomes $\frac{(p^*-c)Q(p^*)}{1-\alpha_i \alpha_j}$, which is below $\frac{\alpha_j p^* Q(p^*)}{1-\alpha_i \alpha_j}$ as $p^* < p\left(\alpha_j\right)$. Hence, p^* is weakly dominated for firm j by $p\left(\alpha_j\right)$, implying that a trembling hand argument will eliminate all such equilibria.

Second, $p'(\alpha_j) \geq 0$ and $\pi'(\alpha_j) \geq 0$: when firm i is the sole innovator, its price and stage 2 profit are weakly increasing with α_j. 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.

Third, the threshold above which the innovation becomes drastic, $\sigma \equiv \frac{p^m - c}{p^m}$, is inversely related to the size of the innovation, c. In particular, $\sigma \to 0$ as $c \to p^m$ (by Assumption A3, the innovation is then largest), and $\sigma \to 1/2$ as $p^m \to 2c$ (the innovation then is smallest). Since $\sigma < 1/2$, PCO levels that are sufficiently close to $1/2$ make the innovation drastic.

Fourth, $\alpha_j \leq \alpha_i$ implies that $p(\alpha_j) \leq p(\alpha_i)$; since $p(\alpha_i) \leq p^m$, it follows from Assumption A1 that $\pi(\alpha_j) \leq \pi(\alpha_i)$. That is, the stage 2 profit of the firm with the small PCO stake is higher than that of the firm with the larger PCO stake. Moreover, $\pi'(\alpha_i) \geq 0$ and $\pi'(\alpha_j) \geq 0$: the stage 2 profit of each firm increases with the rival's stake in the firm.

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 stage 2 equilibrium profits 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* $\alpha_j \leq \alpha_i$, $p(\alpha_j) \leq p(\alpha_i)$ and $\pi(\alpha_j) \leq \pi(\alpha_i)$.

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

$$\text{(6)} \qquad \mathrm{BR}_i\left(\lambda_j\right) = \begin{cases} 0 & \lambda_j > \frac{\pi(\alpha_j)}{\pi(\alpha_j)+\alpha_i\pi(\alpha_i)}, \\ \left(1-\lambda_j\right)\frac{\pi(\alpha_j)}{k} - \lambda_j\frac{\alpha_i\pi(\alpha_i)}{k} & \lambda_j \leq \frac{\pi(\alpha_j)}{\pi(\alpha_j)+\alpha_i\pi(\alpha_i)}. \end{cases}$$

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

Notice that $\mathrm{BR}_i'\left(\lambda_j\right) \leq 0$ and $\mathrm{BR}_j'\left(\lambda_i\right) \leq 0$, 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, $\pi'\left(\alpha_j\right) \geq 0$ and $\pi_j'\left(\alpha_i\right) \geq 0$, so $\mathrm{BR}_i\left(\lambda_j\right)$ is increasing with α_j and decreasing with α_i. Consequently, PCO has two distinct effects on the incentive to invest, which we will refer to as the "price effect" and the "cannibalization effect." The price effect is due to the effect of α_i and α_j on the profits of the two firms in stage 2 and hence their marginal benefit of investment. In our model, each firm makes a profit only when it is the sole innovator and this profit is increasing with the rival's stake in the firm (the rival then becomes softer). Specifically, an increase in α_j boosts $\left(1-\lambda_j\right)\pi\left(\alpha_j\right)$, which is firm i's extra profit when it innovates (the profit is realized only when firm j fails to innovate), while an increase in α_i boosts $-\lambda_j\alpha_i\pi\left(\alpha_i\right)$, which is firm i's share in the negative externality of its investment on firm j.[27] Hence, the price effect of α_j is positive and the price effect of α_i is negative. The cannibalization effect of PCO arises because an increase in α_i implies that firm i internalizes a larger fraction of the negative effect of its investment on firm j's chance to be a sole innovator. Hence, firm i's marginal cost of investment increases with α_i. In the Appendix we show that these properties also hold in more general settings.

A (subgame perfect) Nash equilibrium in stage 1 is a pair $(\lambda_i^*, \lambda_j^*)$, defined by the intersection of $\mathrm{BR}_i\left(\lambda_j\right)$ and $\mathrm{BR}_j\left(\lambda_i\right)$ 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: $k > \underline{k} \equiv \pi\left(\alpha_i\right)\left(1+\alpha_i\frac{\pi(\alpha_i)}{\pi(\alpha_j)}\right)$ for all $0 \leq \alpha_j \leq \alpha_i < 1/2$

[27] The negative externality reflects the idea that firm j innovates with probability λ_j and then earns a profit of $\pi\left(\alpha_i\right)$ conditional on firm i failing to innovate. When firm i succeeds, the expected profit $\lambda_j\pi\left(\alpha_i\right)$ is lost.

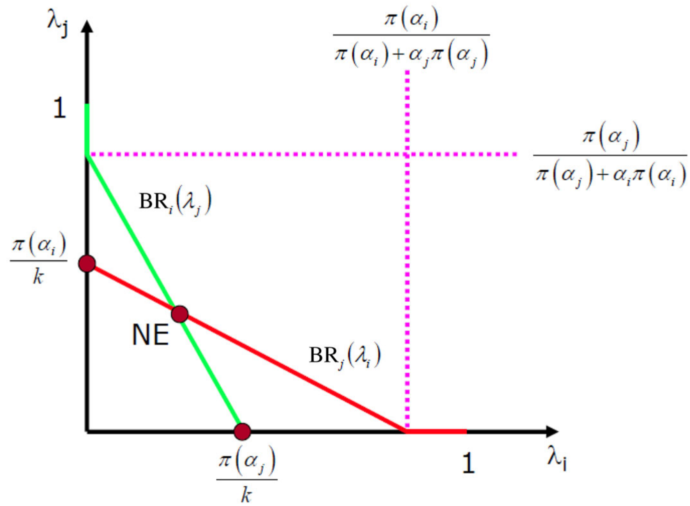

Figure 1
The Best Response Functions in Stage 1 and the Nash Equilibrium

Notes: [Colour figure can be viewed at wileyonlinelibrary.com]

Assumption A5 is stronger than Assumption A4. The reason is as follows. Recalling that $\pi\left(\alpha_i\right)$ is (weakly) increasing with α_i and $\pi\left(\alpha_j\right)$ 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 $\alpha_j = 0$. Its value then is $\pi\left(1/2\right)\left(1 + \frac{\pi(1/2)}{2\pi(0)}\right) > \pi\left(1/2\right) = \pi^m$, where the last equality follows from equation (5) because $\sigma < 1/2$ by Assumption A3.

The equilibrium in stage 1 is illustrated in Figure 1. Assumption A5 ensures that $\text{BR}_i\left(\lambda_j\right)$ crosses $\text{BR}_j\left(\lambda_i\right)$ in the interior of (λ_i, λ_j) space once and from above.[28] In the Appendix we also show that Assumption A5 ensures that the slope of $\text{BR}_i\left(\lambda_j\right)$ in the $\left(\lambda_i, \lambda_j\right)$ space exceeds 1 in absolute value, whereas the slope of $\text{BR}_j\left(\lambda_i\right)$ is below 1.

Lemma 2. The equilibrium investment levels chosen in stage 1 are given by

$$\lambda_i^* = \frac{\pi\left(\alpha_j\right)k - \pi\left(\alpha_i\right)\left(\pi\left(\alpha_j\right) + \alpha_i\pi\left(\alpha_i\right)\right)}{k^2 - \left(\pi\left(\alpha_j\right) + \alpha_i\pi\left(\alpha_i\right)\right)\left(\pi\left(\alpha_i\right) + \alpha_j\pi\left(\alpha_j\right)\right)}, \tag{7}$$

[28] 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.

and

$$\lambda_j^* = \frac{\pi(\alpha_i)k - \pi(\alpha_j)\left(\pi(\alpha_i) + \alpha_j\pi(\alpha_j)\right)}{k^2 - \left(\pi(\alpha_j) + \alpha_i\pi(\alpha_i)\right)\left(\pi(\alpha_i) + \alpha_j\pi(\alpha_j)\right)}, \tag{8}$$

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 shows that the investment level of firm i, which holds the larger PCO stake, is smaller than that of firm j (the Nash equilibrium in Figure 1 is attained above a 45° line that passes through the origin).[29] Moreover, firm i's investment level, λ_i^*, is bounded from above by $1/2$ (when $k \to \underline{k}$ and $\alpha_j = \alpha_i = 0$), firm j's investment level, λ_j^*, is bounded from above by 1 (when $k \to \underline{k}$, $\alpha_j = 0$ and $\alpha_i \to 0$), and 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 $\alpha_i \to 0$. The 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. The logic for this can be seen in Figure 1. As $k \to \underline{k}$, the vertical intercept of $\mathrm{BR}_j\left(\lambda_i\right)$, $\frac{\pi(\alpha_i)}{k}$, tends to $\frac{\pi(\alpha_j)}{\pi(\alpha_j)+\alpha_i\pi(\alpha_i)}$, which is also the vertical intercept of $\mathrm{BR}_i\left(\lambda_j\right)$; 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 stakes 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 and λ_j^* is lowest under a symmetric PCO structure where

[29] 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.

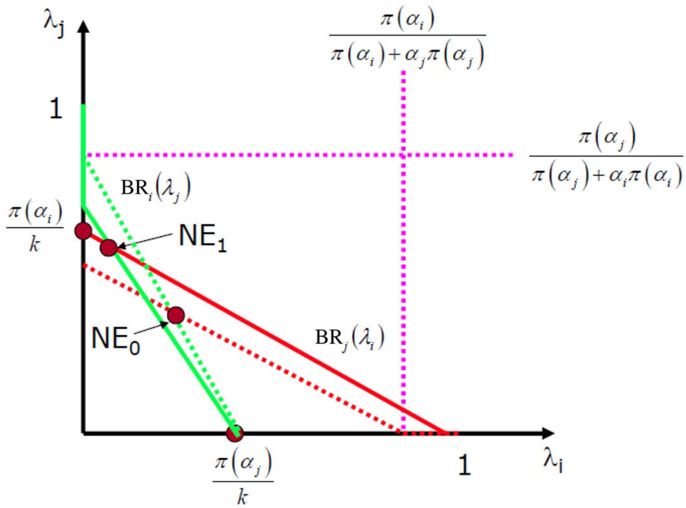

Figure 2
The Effect of an Increase in α on the Nash Equilibrium
Notes: [Colour figure can be viewed at wileyonlinelibrary.com]

$\alpha_i = \alpha_j$ and λ_i^* is lowest and λ_j^* is largest under a maximally asymmetric PCO structure where $\alpha_i \to \frac{1}{2}$ and $\alpha_j = 0$;

(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 $\sigma \le \alpha_j \le \alpha_i$ (the innovation is drastic) $\lambda_i^* + \lambda_j^*$ 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 $\mathrm{BR}_i\left(\lambda_j\right)$ 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 $\mathrm{BR}_j\left(\lambda_i\right)$ 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, $\lambda_i^* < \lambda_j^*$.[30]

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

While Proposition 1(i) shows that an increase in α_i decreases λ_i^* and increases λ_j^*, the fact that both $\mathrm{BR}_i\left(\lambda_j\right)$ and $\mathrm{BR}_j\left(\lambda_i\right)$ 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 $\pi\left(\alpha_i\right) = \pi\left(\alpha_j\right) = \pi^m$, a small increase in α_i decreases λ_i^* more than it increases λ_j^*. The reason for this is that when $\pi\left(\alpha_i\right) = \pi\left(\alpha_j\right) = \pi^m$, an increase α_i does not give rise to a price effect, so only $\mathrm{BR}_i\left(\lambda_j\right)$ rotates counterclockwise around its horizontal intercept, while $\mathrm{BR}_j\left(\lambda_i\right)$ stays intact. The new equilibrium then, lies on $\mathrm{BR}_j\left(\lambda_i\right)$; since the slope of $\mathrm{BR}_j\left(\lambda_i\right)$ 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 (the slope of the marginal cost of investment), c (the innovation size), and π^m (the monopoly profit that a sole innovator earns when the innovation drastic, i.e., when $\sigma \leq \alpha_j \leq \alpha_i$).

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

(iii) when $\sigma \leq \alpha_j \leq \alpha_i$ (the innovation is drastic), 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. When k increases, λ_j^* decreases and firm i becomes more likely to be 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 $k \to \underline{k}$, $\lambda_i^* \to 0$; hence the increase in firm i's marginal cost, $k\lambda_i^*$, is lower than the increase in its marginal benefit, so firm i invests more. By continuity, this is also true when k is not too far from $\underline{k}$. 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.[31] 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, $\mathrm{BR}_i\left(\lambda_j\right)$ and $\mathrm{BR}_j\left(\lambda_i\right)$, 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, $\mathrm{BR}_i\left(\lambda_j\right)$ rotates clockwise around its vertical intercept, while $\mathrm{BR}_j\left(\lambda_i\right)$ rotates counterclockwise around its horizontal intercept, so now $\mathrm{BR}_i\left(\lambda_j\right)$ and $\mathrm{BR}_j\left(\lambda_i\right)$ intersect at the interior of the $\left(\lambda_i, \lambda_j\right)$ space, implying that λ_i^* becomes positive, whereas λ_j^* falls. As $k \to \infty$, $\mathrm{BR}_i\left(\lambda_j\right)$ and $\mathrm{BR}_j\left(\lambda_i\right)$ 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 $\alpha_j = \alpha_i$, $\mathrm{BR}_i\left(\lambda_j\right)$ and $\mathrm{BR}_j\left(\lambda_i\right)$ intersect on a 45° line, but as k increases, they shift inward, so their intersection moves closer to the origin.

Proposition 2(ii) and (iii) shows 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 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 $\mathrm{BR}_i\left(\lambda_j\right)$ and $\mathrm{BR}_j\left(\lambda_i\right)$ 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 is drastic (i.e., $\sigma \leq \alpha_j \leq \alpha_i$), 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., $\sigma \leq \alpha_j \leq \alpha_i$), π^m affects λ_i^* and λ_j^* only through k/π^m, and hence has the

[31] 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^*, $\lambda_j^* > 0$ as $k \to \underline{k}$.

opposite effect of k.[32] That is, λ_i^* is first decreasing and then increasing with π^m . As in the case of k, this counterintuitive result arises because the increase in π^m when it is small 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.

V. 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, $p\left(\alpha_j\right)$ if only firm i innovates, and $p\left(\alpha_i\right)$ if only firm j innovates. Therefore, expected consumer surplus, as a function of the PCO stakes, α_i and α_j, is given by

$$\begin{aligned}\mathrm{CS}\left(\alpha_i,\alpha_j\right) &= \lambda_i^*\lambda_j^* S(0) + \lambda_i^*\left(1-\lambda_j^*\right) S\left(p\left(\alpha_j\right)\right) + \lambda_j^*\left(1-\lambda_i^*\right) S\left(p\left(\alpha_i\right)\right) \\ &\quad + \left(1-\lambda_i^*\right)\left(1-\lambda_j^*\right) S(c),\end{aligned} \tag{9}$$

where λ_i^* and λ_j^* are given by (7) and (8) and $S(p) = \int_p^\infty Q(x)dx$. Since $0 \leq \alpha_j \leq \alpha_i$, (4) implies that $S(0) > S(c) \geq S\left(p\left(\alpha_j\right)\right) \geq S\left(p\left(\alpha_i\right)\right)$, with strict inequalities when $0 < \alpha_j < \alpha_i$.

Absent PCO, where $\alpha_i = \alpha_j = 0$, $p(0) = c$, consumer surplus is given by

$$\mathrm{CS}(0,0) = S(c) + \lambda_i\lambda_j\left(S(0) - S(c)\right).$$

This expression is clearly increasing with λ_i and λ_j, implying that consumers benefit when firms invest more. Since in equilibrium $\lambda^* = \frac{\pi(c)}{k+\pi(c)} < 1$, there is too little investment from consumers' point of view.

In the presence of PCO, however, this is no longer necessarily true. To see why, note that

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

and

$$\frac{\partial \mathrm{CS}\left(\alpha_i,\alpha_j\right)}{\partial\lambda_j} = \lambda_i\underbrace{\left(S(0) - S\left(p\left(\alpha_j\right)\right)\right)}_{(+)} - \left(1-\lambda_i\right)\underbrace{\left(S(c) - S\left(p\left(\alpha_i\right)\right)\right)}_{(+)}. \tag{11}$$

[32] 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.

Equations (10) and (11) show that in the presence of PCO, 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 $p\left(\alpha_j\right)$ or $p\left(\alpha_i\right)$ instead of 0 or c. Hence, in the presence of PCO, innovation may either be insufficient or excessive from the perspective of consumers. 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 in equilibrium, λ_j^* is particularly large under maximal asymmetry of the PCO structure, whereas λ_i^* is particularly large when the PCO structure is symmetric, $\frac{\partial \mathrm{CS}(\alpha_i,\alpha_j)}{\partial \lambda_i}$ is more likely to be positive when the PCO structure is asymmetric and negative when it is symmetric, and conversely for $\frac{\partial \mathrm{CS}(\alpha_i,\alpha_j)}{\partial \lambda_j}$.

We now turn to the effect of PCO on expected consumer surplus. Straightforward differentiation reveals that

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

$$+\lambda_j^*\left(1-\lambda_i^*\right)S'\left(p\left(\alpha_i\right)\right)p'\left(\alpha_i\right). \tag{12}$$

The first term in (12) 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 $-\frac{\partial \mathrm{CS}(\alpha_i,\alpha_j)}{\partial \lambda_i}$. The second term in (12) 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 $\frac{\partial \mathrm{CS}(\alpha_i,\alpha_j)}{\partial \lambda_j}$. The third term in (12) reflects the price effect of PCO. Noting that $S'\left(p\left(\alpha_i\right)\right)=-Q\left(p\left(\alpha_i\right)\right)<0$, the price effect is negative when $\alpha_i<\sigma$ because then, by Lemma 1, $p'\left(\alpha_i\right)>0$, but it vanishes when $\alpha_i \geq \sigma$ because then, firm j already charges p^m when it serves the entire market, so there is no price effect when α_i increases further.

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

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

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

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

If at least one inequality in (i)–(iii) is strict or $p'\left(\alpha_i\right) < 0$, then $\frac{\partial \text{CS}(\alpha_i,\alpha_j)}{\partial \alpha_i} < 0$.

Proof. See the Appendix. ■

Proposition 3 shows that an increase in λ_i^* (the small investment) benefits consumers more, or harms them less, than an increase in λ_j^* (the large 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. The logic behind the three conditions is as follows. An increase in α_i causes a decrease in λ_i^* and an increase in λ_j^*. Condition (i) requires that both changes harm consumers. Condition (ii) requires that the increase in λ_j^* outweighs the decrease in λ_i^* and $\frac{\partial \text{CS}(\alpha_i,\alpha_j)}{\partial \lambda_i} \leq 0$. By (10), the latter condition is more likely to hold when λ_j^* is relatively small. Intuitively, by part (i), $\frac{\partial \text{CS}(\alpha_i,\alpha_j)}{\partial \lambda_i} \leq 0$ implies that $\frac{\partial \text{CS}(\alpha_i,\alpha_j)}{\partial \lambda_j} \leq 0$, 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 $\frac{\partial \text{CS}(\alpha_i,\alpha_j)}{\partial \lambda_j} \geq 0$, which by (11) holds when λ_i^* is relatively high. By part (i), $\frac{\partial \text{CS}(\alpha_i,\alpha_j)}{\partial \lambda_j} \geq 0$ implies that $\frac{\partial \text{CS}(\alpha_i,\alpha_j)}{\partial \lambda_i} \geq 0$, 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, $\frac{\partial \text{CS}(\alpha_i,\alpha_j)}{\partial \lambda_i} \leq 0$ is sufficient for $\frac{\partial \text{CS}(\alpha_i,\alpha_j)}{\partial \alpha_i} \leq 0$, with strict inequality when $\frac{\partial \text{CS}(\alpha_i,\alpha_j)}{\partial \lambda_i} < 0$;

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

Proof. See the Appendix. ■

Corollary 1 shows two cases where an increase in α_i surely harms consumers. In the first case, $\frac{\partial \mathrm{CS}(\alpha_i,\alpha_j)}{\partial \lambda_i} \leq 0$ implies $\frac{\partial \mathrm{CS}(\alpha_i,\alpha_j)}{\partial \lambda_j} \leq 0$, and in the neighborhood of symmetric PCO structure, the negative effect due to the increase in λ_j^* outweighs the beneficial effect of the decrease in λ_i^*. In the second case, $\frac{\partial \mathrm{CS}(\alpha_i,\alpha_j)}{\partial \lambda_j} \geq 0$ implies that $\frac{\partial \mathrm{CS}(\alpha_i,\alpha_j)}{\partial \lambda_i} \geq 0$, and 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 $\sigma \leq \alpha_j \leq \alpha_i$ (the innovation is drastic). Given the equilibrium investment levels, $\frac{\partial \mathrm{CS}(\alpha_i,\alpha_j)}{\partial \lambda_i} \leq 0$ is sufficient for $\frac{\partial \mathrm{CS}(\alpha_i,\alpha_j)}{\partial \alpha_j} \geq 0$, with strict inequality if $\frac{\partial \mathrm{CS}(\alpha_i,\alpha_j)}{\partial \lambda_i} < 0$.

Proof. See the Appendix. ■

VI. 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.[33] Since the monopoly price in this case is $p^m = B$, Assumption A3 requires that $c < B < 2c$. The equilibrium price of firm i, $p\left(\alpha_j\right)$, is still given by (4), with $p^m = B$. The equilibrium price of firm j, $p\left(\alpha_i\right)$, is analogous. With a unit demand function, $p\left(\alpha_j\right)$ and $p\left(\alpha_i\right)$ are also the standalone profits of firms i and j.

Note that now, the threshold of the PCO levels above which the innovation becomes drastic is $\sigma \equiv \frac{p^m - c}{p^m} = \frac{B-c}{B} \in (0, 1/2)$. We will now consider two cases: (i) $\alpha_j \leq \alpha_i < \sigma$ (the innovation is non-drastic), and (ii) $\sigma \leq \alpha_j \leq \alpha_i$ (the innovation is drastic).[34]

VI(i). *Non-drastic innovation:* $\alpha_j \leq \alpha_i < \sigma$

The equilibrium prices in this case are $p\left(\alpha_j\right) = \frac{c}{1-\alpha_j}$ and $p\left(\alpha_i\right) = \frac{c}{1-\alpha_i}$. An increase in PCO affects consumers both directly through the equilibrium

[33] 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].

[34] There is also an intermediate case where $\alpha_j < \sigma \leq \alpha_i$. This case is a hybrid of cases (i) and (ii) and we therefore do not study it.

prices, as well as indirectly through the equilibrium investment levels. With a unit demand function, the standalone profits in stage 2 are $\pi\left(\alpha_j\right)=\frac{c}{1-\alpha_j}$ and $\pi\left(\alpha_i\right)=\frac{c}{1-\alpha_i}$. Assumption A5 then requires that $k>\underline{k}=\frac{c\left(1-\alpha_i\alpha_j\right)}{\left(1-\alpha_i\right)^2}$. The inequality can be rewritten as $z>\underline{z}\equiv\frac{1-\alpha_i\alpha_j}{\left(1-\alpha_i\right)^2}$, where $z\equiv k/c$ is the ratio of the slope of marginal cost of investment, k, to the size of the innovation, c. In what follows, we will refer to z as the "relative cost of innovation." Note that the lower bound on the relative cost of innovation, $\underline{z}$, 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 $\pi\left(\alpha_j\right)=\frac{c}{1-\alpha_j}$ and $\pi\left(\alpha_i\right)=\frac{c}{1-\alpha_i}$ in (7) and (8), and using the definitions of z and $\underline{z}$, the equilibrium investment levels become

$$(13)\quad \lambda_i^*=\frac{\left(1-\alpha_j\right)\left(z-\underline{z}\right)}{z^2\left(1-\alpha_j\right)^2-\underline{z}^2\left(1-\alpha_i\right)^2},\qquad \lambda_j^*=\frac{\left(1-\alpha_i\right)\left(\frac{z\left(1-\alpha_j\right)^2}{\left(1-\alpha_i\right)^2}-\underline{z}\right)}{z^2\left(1-\alpha_j\right)^2-\underline{z}^2\left(1-\alpha_i\right)^2}.$$

λ_i^* and λ_j^* depend only on the PCO stakes, α_i and α_j, and on the relative cost of innovation, z. In Lemma 2(i) and in Proposition 1(i) we already established that $0<\lambda_i^*\le\lambda_j^*$, $\lambda_i^*<1/2$, $\lambda_j^*<1$, $\frac{\partial\lambda_i^*}{\partial\alpha_i}<0<\frac{\partial\lambda_j^*}{\partial\alpha_i}$ 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 $\alpha_j\le\alpha_i<\sigma$.

Lemma 3. Suppose that $\alpha_j\le\alpha_i<\sigma$. 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;
(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 examine the effects of α_i and α_j on $\lambda_i^*+\lambda_j^*$ only in the neighborhood of a symmetric PCO structure or when $\sigma\le\alpha_j\le\alpha_i$, where PCO has no price effect. In Lemma 3 by contrast, we can study the effects for all $0\le\alpha_j\le\alpha_i<\sigma$. 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.[35]

As in Lemma 2, we get here 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

$$\begin{aligned} \mathrm{CS}\left(\alpha_i, \alpha_j\right) = B - \left(1-\lambda_i^*\right)\left(1-\lambda_j^*\right)c \\ - \lambda_i^*\left(1-\lambda_j^*\right)\frac{c}{1-\alpha_j} - \lambda_j^*\left(1-\lambda_i^*\right)\frac{c}{1-\alpha_i}. \end{aligned} \tag{14}$$

We now prove the following result:

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

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

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

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

Proof. See the Appendix. ■

[35] The second effect cannot be negative and dominate the first positive effect for both firms. To see why, suppose by way of negation that it 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.

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 of the PCO structure as 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 of the PCO structure), benefits consumers when the relative cost of innovation, z, is small, but harms consumers when z is large. How small should z 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 (α_i increases relative to α_j), an increase in α_j benefits consumers for a larger set of values of z.[36]

Proposition 5 implies that when the PCO stakes are not too large, 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 smaller stake, provided that the relative cost of innovation, z, is low and that α_j is not too close to α_i (the PCO structure remains sufficiently asymmetric).[37]

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), $\frac{\partial \mathrm{CS}(\alpha_i,\alpha_j)}{\partial \lambda_i^*} \leq 0$ is sufficient for $\frac{\partial \mathrm{CS}(\alpha_i,\alpha_j)}{\partial \alpha_i} \leq 0$. Differentiating (14),

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

Recalling from Lemma 3(iii) that λ_j^* is decreasing with z, the above condition is more likely to hold when z is large, which is consistent with Proposition 5(i). Moreover, recall from Proposition 1(i) that λ_j^* is increasing with α_i and decreasing with α_j and note that $\frac{\alpha_j(1-\alpha_i)}{1-\alpha_i\alpha_j}$ is decreasing with α_i and increasing with α_j. Hence, the condition is more likely to hold when α_i is small and α_j is large. 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, $\frac{\partial \mathrm{CS}(\alpha_i,\alpha_j)}{\partial \alpha_i} < 0$. The reason for this is that the sufficient condition does not take into account the price effect of PCO: the fact

[36] For instance, if $\alpha_j = 0$, the largest z for which $\frac{\partial \mathrm{CS}(\alpha_i,\alpha_j)}{\partial \alpha_j} \geq 0$ 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 $\alpha_i \to 1/2$.

[37] 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.

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 II, 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 feature R&D spillovers), if we evaluate (14) at $\alpha_j = \alpha_i = \alpha < \sigma$ and differentiate $\text{CS}(\alpha, \alpha)$ with respect to α, we get

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

That is, an increase in a symmetric PCO stake, 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 and α_j is sufficiently below α_i. In other words, not every increase in PCO necessarily harms consumers absent R&D spillovers.

VI(ii). *Drastic innovation: $\sigma \leq \alpha_j \leq \alpha_i$*

This case can arise only when both firms hold stakes in one another and these stake are sufficiently large. The equilibrium price then is $p\left(\alpha_j\right) = p\left(\alpha_j\right) = p^m = B$, implying that an increase in PCO only affects the equilibrium investment levels, but does not give rise to a price effect. With a unit demand function, $\pi\left(\alpha_i\right) = \pi\left(\alpha_i\right) = B$. Assumption A5 then requires that $k > \underline{k} \equiv \left(1 + \alpha_i\right) B$, which is equivalent to $m > 1 + \alpha_i$, where $m \equiv k/B$ is the analog of z and reflects the relative cost of innovation when the innovation is drastic.

Substituting $\pi\left(\alpha_i\right) = \pi\left(\alpha_i\right) = B$ in (7) and (8), the equilibrium investment levels become

$$(15) \qquad \lambda_i^* = \frac{m - \left(1 + \alpha_i\right)}{m^2 - \left(1 + \alpha_i\right)\left(1 + \alpha_j\right)}, \qquad \lambda_j^* = \frac{m - \left(1 + \alpha_j\right)}{m^2 - \left(1 + \alpha_i\right)\left(1 + \alpha_j\right)}.$$

Note that λ_i^* and λ_j^* depend only on the PCO stakes, α_i and α_j, and on relative cost of innovation, m. 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^*. In the lemma and the proposition that follows we use the following threshold:

$$(16) \qquad \widehat{m} \equiv 1 + \alpha_i + \sqrt{\left(\alpha_i - \alpha_j\right)\left(1 + \alpha_i\right)}.$$

Lemma 4. Suppose that $\sigma \leq \alpha_j \leq \alpha_i$. 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 $m > \widehat{m}$ 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 a firm's stake in a rival has a bigger effect on the firm's own investment level than on that of the rival.

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 $m \to 1 + \alpha_i$. Note again the stark difference between symmetric and asymmetric PCO structures. In particular, if $m \to 1 + \alpha_i$, then $\lambda_i^* = \lambda_j^* \to 1/2\,(1+\sigma)$ if $\alpha_j = \alpha_i = \sigma$, but $\lambda_i^* \to 0$ and $\lambda_j^* \to 1/(1+\sigma)$ if $\alpha_j = \sigma$ and $\alpha_i \to \sigma$. Lemma 4(iii) also shows that so long as $\alpha_j < \alpha_i$, λ_i^* is an inverse U-shaped function of m. Noting that $\widehat{m}$ increases with α_i and decreases with α_j (and hence is larger when the PCO structure becomes more asymmetric), it follows that λ_i^* is increasing with α_i for a larger set of parameters 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, that is, $\alpha_i = \alpha_j = \alpha$, $\widehat{m} = 1 + \alpha < m$, 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, note that with probability $\lambda_i^* \lambda_j^*$, both firms innovate, so $p = 0$ and consumer surplus is B; with probability $\lambda_i^* \left(1 - \lambda_j^*\right) + \lambda_j^* \left(1 - \lambda_i^*\right)$, only one firm innovates, so $p = B$ and consumer surplus is 0; and with probability $\left(1 - \lambda_i^*\right)\left(1 - \lambda_j^*\right)$, both firms do not innovate, so $p = c$ and consumer surplus is $B - c$. Hence, expected consumer surplus is given by

$$\begin{aligned} \mathrm{CS}\left(\alpha_i, \alpha_j\right) &= \lambda_i^* \lambda_j^* B + \left(1 - \lambda_i^*\right)\left(1 - \lambda_j^*\right)(B - c) \\ &= B\left[\lambda_i^* \lambda_j^* + \left(1 - \lambda_i^*\right)\left(1 - \lambda_j^*\right)\sigma\right]. \end{aligned} \tag{17}$$

In the next proposition we examine how $\mathrm{CS}\left(\alpha_i, \alpha_j\right)$ is affected by unilateral increases in α_i and α_j.

Proposition 6. Suppose that $\sigma \leq \alpha_j \leq \alpha_i$. Then,

(i) $\frac{\partial \mathrm{CS}(\alpha_i,\alpha_j)}{\partial \alpha_i} > 0$ when m is sufficiently large or $\sigma \to 1/2$ and $\frac{\partial \mathrm{CS}(\alpha_i,\alpha_j)}{\partial \alpha_i} < 0$ when m is not too much above its lower bound, $1+\alpha_i$ or when $\sigma \to 0$ and $\sigma m \to 0$;

(ii) $\frac{\partial \mathrm{CS}(\alpha_i,\alpha_j)}{\partial \alpha_j} > 0$ when m is sufficiently large, when m is not too much above its lower bound, $1+\alpha_i$, and when σ is sufficiently close to $1/2$;

(iii) when $\sigma \to 0$ and $\sigma m \to 0$, $\frac{\partial \mathrm{CS}(\alpha_i,\alpha_j)}{\partial \alpha_j} > 0$ as $m < \widehat{m}$ and $\frac{\partial \mathrm{CS}(\alpha_i,\alpha_j)}{\partial \alpha_j} < 0$ as $m > \widehat{m}$;

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

Parts (i) and (ii) of Proposition 6 show that when the PCO stakes are sufficiently large to make the innovation drastic, an increase in α_i (which makes the PCO structure more asymmetric) benefits consumers when the relative cost of innovation, m, is sufficiently large, or the lower bound on the PCO values for which the innovation is drastic, σ, tends to its upper bound, $1/2$, but harms consumers when m or σ are small. As for α_j, an increase in α_j (which makes the PCO structure more symmetric) benefits consumers when m is sufficiently large or sufficiently close to its lower bound, $1+\alpha_i$, or when σ is sufficiently large (the range of PCO values for which the innovation is drastic, $[\sigma, 1/2)$, is then narrow). Interestingly, the effect of an increase in α_j on consumer surplus is not necessarily monotonic in m.[38] Notice that when m is large or $\sigma \to 1/2$, increases in α_j and α_i benefit consumers. Hence, 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 $\alpha_j = \alpha_i = \sigma$. In both case, a symmetric PCO structure benefits consumers.

Part (iii) of the proposition shows that when $\sigma \to 0$, an increase in α_j still benefits consumers when m is small, but harms consumers when m is large, and may even benefit consumers for all $m > 1+\alpha_i$, provided that σ is sufficiently large. Finally, part (iv) of Proposition 6 shows that an increase in the PCO

[38] To illustrate, suppose that $\alpha_i = 0.4$, $\alpha_j = 0.3$, and $\sigma = 0.25$. Then $m > 1.4$. Straightforward calculations show that $\frac{\partial \mathrm{CS}(\alpha_i,\alpha_j)}{\partial \alpha_j} > 0$ if $1.4 < m < 2.14$, $\frac{\partial \mathrm{CS}(\alpha_i,\alpha_j)}{\partial \alpha_j} < 0$ if $2.14 < m < 3.34$, and $\frac{\partial \mathrm{CS}(\alpha_i,\alpha_j)}{\partial \alpha_j} > 0$ if $m > 3.34$. However, if $\alpha_i = 0.4$, $\alpha_j = 0.3$, and $\sigma = 0.3$, then $\frac{\partial \mathrm{CS}(\alpha_i,\alpha_j)}{\partial \alpha_j} > 0$ for all $m > 1.4$.

stakes when they are (nearly) symmetric benefits consumers if m is large, but harms them if m is small.[39]

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, $\frac{\partial \mathrm{CS}(\alpha_i,\alpha_j)}{\partial \lambda_j^*} \geq 0$ is sufficient for $\frac{\partial \mathrm{CS}(\alpha_i,\alpha_j)}{\partial \alpha_i} \leq 0$. Differentiating (17),

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

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 $\frac{\partial \mathrm{CS}(\alpha_i,\alpha_j)}{\partial \alpha_i} < 0$ when $\sigma \to 0$ (so the condition holds) and $\frac{\partial \mathrm{CS}(\alpha_i,\alpha_j)}{\partial \alpha_i} > 0$ when $\sigma \to 1/2$ (the condition fails).

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 (17) at $\alpha_j = \alpha_i = \alpha$ and differentiate with respect to α:

$$\frac{\partial \mathrm{CS}(\alpha,\alpha)}{\partial \alpha} = \frac{2\sigma B\left(m - \left(\frac{1}{\sigma} - \alpha\right)\right)}{(m+\alpha+1)^3}. \tag{18}$$

The derivative is positive if $m > \frac{1}{\sigma} - \alpha$ and negative if $m < \frac{1}{\sigma} - \alpha$, similarly to Proposition 6(iv). That is, consumers are more likely to benefit from an increase in PCO when m and σ are large. While this result is consistent with Proposition 6(i), it is not necessarily consistent with Propositions 6(ii) or (iii) which show that an increase in α_j can benefit consumers even when m and σ are small.

VII. PCO VERSUS 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

[39] 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 $\frac{\partial \mathrm{CS}(\alpha_i,\alpha_j)}{\partial \alpha_i} = \frac{\partial \mathrm{CS}(\alpha_i,\alpha_j)}{\partial \alpha_j} > 0$, which is consistent with part (i).

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 United States, 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."[40] 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.[41]

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 give rise to 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 (7) and (8), the equilibrium investment levels are:

$$\lambda_i^* = \lambda_j^* = \lambda(\alpha) \equiv \frac{\pi^m}{k + (1+\alpha)\pi^m}. \tag{19}$$

Expected consumer surplus under PCO is

$$\mathrm{CS}(\alpha) = \lambda(\alpha)^2 S(0) + 2\lambda(\alpha)(1-\lambda(\alpha))\, S\left(p^m\right) + (1-\lambda(\alpha))^2 S(c). \tag{20}$$

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.

[40] Indeed the US 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).

[41] Similarly, Schinkel and Spiegel [2017] show that when firms invest in the sustainability of their respective products (which boosts the willingness of consumers to pay) before competing in the product market, semicollusion promotes investments and may benefit consumers if the slope of the marginal cost of investment is sufficiently low relative to the degree of product differentiation.

VII(i). *PCO versus 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.[42] 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 $\hat{p} \in \left[p^m, p_c^m\right]$ 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 $\pi_c^m \equiv \left(p_c^m - c\right) Q\left(p_c^m\right)$ 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.

The expected value of firm i in stage 1 of the game is,[43]

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

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. Then, PCO leads to more investment and yields a higher expected consumer surplus than semicollusion.

Proof. See the Appendix. ■

Proposition 7 says that a symmetric PCO which is large enough to make the innovation drastic boosts investment and benefits consumers more than semicollusion. Intuitively, under semicollusion, firms charge a price above cost and earn a profit even when they both innovate. By contrast, under PCO, a firm

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

[43] 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).

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 outweighs 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.

VII(ii). *PCO versus 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.[44]

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,

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

$$= \frac{\left(\lambda_i\left(1-\lambda_j\right) + \lambda_j(1-\lambda_i)\right)\pi^m - \frac{k\lambda_i^2}{2} - \frac{k\lambda_j^2}{2}}{1-\alpha}.$$

[44] In Choi [1993], the RJV also generates an information spillovers, which boost the profits of the two firms. We abstract from such spillovers. Kamien *et al.* [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.

The investment levels which maximize this expression are

$$\lambda_i = \lambda_j = \lambda^{RJV} \equiv \frac{\pi^m}{k + 2\pi^m}. \tag{21}$$

Comparing (21) and (19) reveals 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, 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 (20) with λ^{RJV} replacing $\lambda(\alpha)$. To compare consumer surplus under (pure) PCO and under RJV, it is useful to rewrite (20) as

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

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, so 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 $S(p)$ is decreasing and convex, $S(0) - S(c) > |S(p^m) - S(c)|$.[45] But then by Assumption A4, $k > \pi^m$, so from (19) 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 $\text{CS}(\alpha)$ 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.

We now establish the following result.

Proposition 8. Suppose that the two firms hold the same PCO stake, α, in each other. Then $\lambda(\alpha) > \lambda^{RJV}$ for all $\sigma \le \alpha < 1/2$. If $\alpha \ge \sigma$ and firms face a unit demand function, then consumer surplus is larger under PCO than under an RJV if $m < \frac{1}{\sigma} - \alpha$ and conversely if $m > \frac{1}{\sigma} - \alpha$.

Proof. See the Appendix. ■

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

VII(iii). *PCO versus full merger*

Finally, consider a full merger. Then, the two firms 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. 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 $p^m_c \equiv \arg\max(p-c)\,Q(p)$ when they both have a marginal cost c. Hence, in stage 1, the merged entity chooses λ_i and λ_j to maximize the sum of the firms' values, given by

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

where $\pi^m_c \equiv \left(p^m_c-c\right)Q\left(p^m_c\right)$. The investment levels which maximize this expression are,

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

We now prove the following result.

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

$$\alpha<\frac{\pi^m_c k}{\pi^m\left(\pi^m-\pi^m_c\right)}. \tag{22}$$

Expected consumer surplus is higher under PCO.

Proof. See the Appendix. ■

Proposition 9 shows that a sufficiently large symmetric PCO which ensures that the innovation is drastic 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 $\sigma \leq \alpha < \frac{\pi^m_c k}{\pi^m\left(\pi^m-\pi^m_c\right)}$.[46] 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

[46] To see that the set of parameters for which the condition holds is non empty, note that in the unit demand case, $p^m=p^m_c=B$, $\pi^m=B$, and $\pi^m_c=B-c$. Then, $\sigma \equiv \frac{B-c}{B}<\frac{(B-c)k}{Bc}=\frac{\pi^m_c k}{\pi^m\left(\pi^m-\pi^m_c\right)}$, where the inequality follows because by Assumptions A3 and A4, $k>B>c$.

for the merged entity to lower its cost), whereas under PCO it implies a loss of $\alpha\pi^m$, as 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 consumer surplus, note that prices under a full merger are equal to the monopoly prices, whereas under PCO they are equal to the competitive prices when both firms innovate or both firms fail. Hence, consumers are better off under PCO than under a full merger.

VIII. 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 profits. 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 PCO stake in the innovating firm, as the former is then reluctant to undercut the latter because it shares its profit. PCO then gives rise to a price effect, which boosts the incentive to invest and become the innovating firm. Apart from a price effect, PCO also creates a cannibalization effect: when a firm innovates, it cannibalizes the profit of the rival when it innovates. Due to PCO, the firm internalizes part of the resulting negative externality and hence its incentive to innovate becomes weaker, the larger its PCO stake is.

Importantly, more investment is not always good for consumers: although it 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 show that the effect of a unilateral increase in PCO on investments and on consumer surplus depends on the size of the PCO stakes and how symmetric they are. In particular, we provide sufficient conditions for such a unilateral increase to harm or benefit consumers and then explore the welfare implications in greater detail under the assumption that consumers have a unit demand function.

When the PCO stakes are relatively small, the innovation is non-drastic. Then, an increase in the large PCO stake (which makes the PCO structure more asymmetric) always harms consumers, whereas an increase in the small PCO stake (which makes the PCO structure more symmetric) can benefit consumers, provided that the relative cost of innovation is not too high; otherwise it also harms consumers. By contrast, 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 give rise to a price effect. An increase in either PCO stake can benefit consumers if the relative cost of innovation or the lower bound on the PCO stakes are sufficiently large. An increase in the small PCO can also benefit consumers if the relative cost of innovation or the lower bound on the PCO stakes are small, and when the PCO stakes are nearly symmetric and the relative cost of innovation is large.

Our analysis then shows that when the incentives of firms to invest are taken into account, a unilateral increase in PCO may benefit consumers, contrary to the insight from earlier literature which only examined the effects of PCO on price competition. In particular, our analysis highlights the fact that whether consumers benefit or not from a unilateral increase in PCO depends on the size of PCO stakes and on the extent to which they are symmetric.

APPENDIX

Following are technical proofs.

The price and cannibalization effects in a more general setting. In the more general Aoki and Spiegel [2009] setting where the stage 2 profit of firm i is π^i_{yy} if both firms innovate, π^i_{nn} if both firms fail, π^i_{yn} if the firm innovates and the rival fails, and π^i_{ny} if the firm fails but the rival innovates (the index i is needed because the profit of each firm depends on α_i and α_j which are not necessarily identical), the expected value of firm i when it chooses λ_i in stage 1 is

$$V_i = \frac{\lambda_i\left[\lambda_j\pi^i_{yy} + \left(1-\lambda_j\right)\pi^i_{yn}\right] + \left(1-\lambda_i\right)\left[\lambda_j\pi^i_{ny} + \left(1-\lambda_j\right)\pi^i_{nn}\right] - \frac{k\lambda_i^2}{2}}{1-\alpha_i\alpha_j}$$
$$+ \frac{\alpha_i\left(\lambda_j\left[\lambda_i\pi^j_{yy} + \left(1-\lambda_i\right)\pi^j_{yn}\right] + \left(1-\lambda_j\right)\left[\lambda_i\pi^j_{ny} + \left(1-\lambda_i\right)\pi^j_{nn}\right] - \frac{k\lambda_j^2}{2}\right)}{1-\alpha_i\alpha_j}.$$

Since V_i is quadratic in λ_i, the first-order condition for λ_i is necessary and sufficient. Using this condition, the best-response function of firm i against firm j is given by:

$$\mathrm{BR}_i\left(\lambda_j\right) = \frac{1}{k}\left(\overbrace{\left[\lambda_j\pi^i_{yy} + \left(1-\lambda_j\right)\pi^i_{yn}\right]}^{E\left(\pi^i|i\text{'s success}\right)} - \overbrace{\left[\lambda_j\pi^i_{ny} + \left(1-\lambda_j\right)\pi^i_{nn}\right]}^{E\left(\pi^i|i\text{'s failure}\right)}\right)$$
$$+ \frac{\alpha_i}{k}\left(\overbrace{\left[\lambda_j\pi^j_{yy} + \left(1-\lambda_j\right)\pi^j_{ny}\right]}^{E\left(\pi^j|i\text{'s success}\right)} - \overbrace{\left[\lambda_j\pi^j_{yn} + \left(1-\lambda_j\right)\pi^j_{nn}\right]}^{E\left(\pi^j|i\text{'s failure}\right)}\right).$$

In our Bertrand setting where $\pi^i_{yy} = \pi^i_{nn} = \pi^i_{ny} = 0$ and $\pi^i_{yn} = \pi\left(\alpha_j\right)$, $\mathrm{BR}_i\left(\lambda_j\right)$ simplifies to (6). The cannibalization effect arises because an increase in α_i boosts the second term, which is negative as firm j's profit is higher when firm i fails to innovate. It is easy to see that this effect is general and does not depend on our Bertrand setting.

The price effect is more involved and is due to the effect of α_i and α_j on the stage 2 profits of the two firms. To see how the effect works, it is useful to rewrite the best-response function of firm i as

$$\mathrm{BR}_i\left(\lambda_j\right) = \frac{1}{k}\left[\lambda_j\left(\left(\pi^i_{yy} - \pi^i_{ny}\right) + \alpha_i\left(\pi^j_{yy} - \pi^j_{yn}\right)\right)\right.$$
$$\left. + \left(1-\lambda_j\right)\left(\left(\pi^i_{yn} - \pi^i_{nn}\right) + \alpha_i\left(\pi^j_{ny} - \pi^j_{nn}\right)\right)\right].$$

Sufficient conditions for the price effect of α_j to be positive and the price effect of α_i to be negative are (i) $\pi^i_{yy} - \pi^i_{ny}$ and $\pi^i_{yn} - \pi^i_{nn}$ are increasing with α_j and decreasing with α_i, that is, an increase in α_j has a stronger positive effect on firm i's profit when it innovates than when it fails and an increase in α_i has a stronger negative effect on firm i's profit when it innovates than when it fails; and (ii) $\pi^j_{yy} - \pi^j_{yn}$ and $\pi^j_{ny} - \pi^j_{nn}$ are increasing with α_j and decreasing with α_i, that is, an increase in α_j has a stronger positive effect on firm j's profit when firm i innovates than when it fails and an increase in α_i has a stronger negative effect on firm j's profit when firm i innovates than when it fails. Analogously, condition (ii) can be stated as $\pi^i_{yy} - \pi^i_{yn}$ and $\pi^i_{ny} - \pi^i_{nn}$ are increasing with α_i and decreasing with α_j. That is, an increase in α_i has a stronger negative effect on firm i's profit when firm j fails to innovates than when it succeeds, and an increase in α_j has a stronger positive effect on firm i's profit when firm j fails to innovate than when it succeeds.

Intuitively, firm i becomes softer when α_i increases (the firm internalizes a larger fraction of the negative competitive externality it imposes on firm j). Consequently, firm i's profit decreases with α_i and increases with α_j. Condition (i) holds if the effect is larger when firm i innovates.[47] Condition (ii) holds if the effect is larger when firm j fails than when it innovates.

To illustrate the two conditions, consider a Cournot model with linear inverse demand $p = 1 - q_i - q_j$ and assume that firm i's cost is $\widehat{c}_i = 0$ if firm i innovates and $\widehat{c}_i = c$ if firm i fails to innovate. Then, the values of the two firms under the PCO are given by (2), with the standalone profit of each firm i given by $\Pi_i = \left(1 - q_i - q_j - \widehat{c}_i\right) q_i$. The Nash equilibrium when the two firms choose quantities simultaneously in stage 2 are

$$q_i = \frac{1 - \alpha_i - 2\widehat{c}_i + \left(1+\alpha_i\right)\widehat{c}_j}{3 - \alpha_i - \alpha_j - \alpha_i\alpha_j}, \qquad q_j = \frac{1 - \alpha_j - 2\widehat{c}_j + \left(1+\alpha_j\right)\widehat{c}_i}{3 - \alpha_i - \alpha_j - \alpha_i\alpha_j}.$$

Substituting these quantities in (2), the value function of firm i is

$$\pi^i_{yy} = \frac{\left(1 - \alpha_i\alpha_j\right)^2}{\left(3 - \alpha_i - \alpha_j - \alpha_i\alpha_j\right)^2},$$

[47] In models where firm i makes no profit if it fails to innovate, that is, if $\pi^i_{ny} = \pi^i_{nn} = 0$ (e.g., Jullien and Lefouili [2018], Stenbacka and Van Moer [2023] and our Bertand model), this condition surely holds.

when $\hat{c}_i = \hat{c}_j = 0$,

$$\pi^i_{yn} = \frac{(1+c)^2 + \alpha_i^2 (1-c)\left(\alpha_j^2 + \left(1+2\alpha_j\right)c\right) - \alpha_i\left(2\alpha_j + \left(5 - \left(2-\alpha_j\right)\alpha_j\right)c + 2\left(2-\alpha_j\right)c^2\right)}{\left(3-\alpha_i-\alpha_j-\alpha_i\alpha_j\right)^2},$$

when $\hat{c}_i = 0$ and $\hat{c}_j = c$,

$$\pi^i_{ny} = \frac{\left(1-\alpha_i\alpha_j\right)^2 - \left(4 - \alpha_i\left(1+\alpha_j\right)^2\right)\left(1-\alpha_i-c\right)c}{\left(3-\alpha_i-\alpha_j-\alpha_i\alpha_j\right)^2},$$

when $\hat{c}_i = c$ and $\hat{c}_j = 0$, and

$$\pi^i_{nn} = \frac{\left(1-\alpha_i\alpha_j\right)^2(1-c)^2}{\left(3-\alpha_i-\alpha_j-\alpha_i\alpha_j\right)^2}.$$

Using Mathematica, it is straightforward to verify that $\pi^i_{yy} - \pi^i_{ny}$ and $\pi^i_{yn} - \pi^i_{nn}$ are increasing with α_j and decreasing with α_i.[48] Hence, condition (i) above holds. As for condition (ii), note that

$$\pi^i_{yy} - \pi^i_{yn} = -\frac{k\left[2+k-\alpha_i\left(5-4k-\alpha_j(2-\alpha_j-2k)\ -\ \alpha_i\left((1-k+\alpha_j\left(2-\alpha_j-2k\right)\right)\right)\right]}{\left(3-\alpha_i-\alpha_j-\alpha_i\alpha_j\right)^2},$$

and

$$\pi^i_{ny} - \pi^i_{nn} = -\frac{k\left[2-3k-\alpha_i\left(5-k-\alpha_i-\alpha_j\left(1+\alpha_i\right)\left(2-\alpha_j(1-k)\right)\right)\right]}{\left(3-\alpha_i-\alpha_j-\alpha_i\alpha_j\right)^2}.$$

Using Mathematica again we verify that $\pi^i_{yy} - \pi^i_{yn}$ and $\pi^i_{ny} - \pi^i_{nn}$ are increasing with α_i and decreasing with α_j, implying that condition (ii) also holds. ■

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 $\mathrm{BR}_i\left(\lambda_j\right)$ in the $\left(\lambda_i, \lambda_j\right)$ space and $\frac{\pi(\alpha_i)}{k}$ is the vertical intercept of $\mathrm{BR}_j\left(\lambda_i\right)$.[49]

[48] For example, to check that $\pi^i_{yy} - \pi^i_{ny}$ is increasing with α_j, we define $d\Delta^i_y d\alpha_j \equiv \frac{\partial\left(\pi^i_{yy}-\pi^i_{ny}\right)}{\partial\alpha_j}$ and use the command Reduce[$d\Delta^i_y d\alpha_j \le 0$ && $0 < \alpha_i < 1/2$ && $0 \le \alpha_j < \alpha_i$&& $0 < k \le 1/3$, $\{\alpha_i, \alpha_j\}$]. The command returns the output "False", implying that, given the parameter restrictions, $\frac{\partial\left(\pi^i_{yy}-\pi^i_{ny}\right)}{\partial\alpha_j} > 0$. We repeat the same procedure for the other expressions.

[49] 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., $\mathrm{BR}_i\left(\frac{\pi(\alpha_j)}{\pi(\alpha_j)+\alpha_i\pi(\alpha_i)}\right) = 0$, and $\frac{\pi(\alpha_i)}{k}$ is firm j's best response against $\lambda_i = 0$, i.e., $\mathrm{BR}_j(0) = \frac{\pi(\alpha_i)}{k}$.

Moreover, recall that the assumption that $\alpha_i \geq \alpha_j$ implies that $\pi\left(\alpha_i\right) \geq \pi\left(\alpha_j\right)$, which implies in turn that

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

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 $\mathrm{BR}_i\left(\lambda_j\right)$ in the $\left(\lambda_i, \lambda_j\right)$ space and $\frac{\pi(\alpha_j)}{k}$ is the horizontal intercept of $\mathrm{BR}_j\left(\lambda_i\right)$.[50] Together with the fact that by (6), the slope of $\mathrm{BR}_i\left(\lambda_j\right)$ is constant, and likewise the slope of $\mathrm{BR}_j\left(\lambda_i\right)$ is constant, Assumption A5 ensures that $\mathrm{BR}_i\left(\lambda_j\right)$ crosses $\mathrm{BR}_j\left(\lambda_i\right)$ in the interior of the $\left(\lambda_i, \lambda_j\right)$ 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 $\mathrm{BR}_i\left(\lambda_j\right)$ in the $\left(\lambda_i, \lambda_j\right)$ space exceeds 1 in absolute value, whereas the slope of $\mathrm{BR}_j\left(\lambda_i\right)$ is below 1 in absolute value. To see why, note that fully differentiating (6), evaluating at $\lambda_i = \mathrm{BR}_i\left(\lambda_j\right)$, and using Assumption A5,

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

where the last inequality follows because $\alpha_i \geq \alpha_j$ implies that $\pi\left(\alpha_i\right) \geq \pi\left(\alpha_j\right)$. Similarly, evaluated at $\lambda_j = \mathrm{BR}_j\left(\lambda_i\right)$,

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

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

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

Proof of Lemma 2. The equilibrium investment levels in equations (7) and (8) are determined by the solution to the system $\lambda_i = \mathrm{BR}_i\left(\lambda_j\right)$ and $\lambda_j = \mathrm{BR}_j\left(\lambda_i\right)$.

(i) First, note that $\lambda_i^*, \lambda_j^* > 0$ because the equilibrium in stage 1 is interior as we have proved above. Another way to see it is to note that Assumption A5 ensures that

[50] 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 $\lambda_j = 0$.

the numerators of (7) and (8) are positive and also ensures that the denominators are positive as

$$k^2 - \left(\pi\left(\alpha_j\right) + \alpha_i \pi\left(\alpha_i\right)\right)\left(\pi\left(\alpha_i\right) + \alpha_j \pi\left(\alpha_j\right)\right)$$

$$> \underbrace{\pi\left(\alpha_i\right)^2 \left(1 + \alpha_i \frac{\pi\left(\alpha_i\right)}{\pi\left(\alpha_j\right)}\right)^2}_{\underline{k}^2} - \left(\pi\left(\alpha_j\right) + \alpha_i \pi\left(\alpha_i\right)\right)\left(\pi\left(\alpha_i\right) + \alpha_j \pi\left(\alpha_j\right)\right)$$

$$= \frac{\left(\pi\left(\alpha_j\right) + \alpha_i \pi\left(\alpha_i\right)\right)}{\pi\left(\alpha_j\right)^2}\left[\pi\left(\alpha_i\right)\pi\left(\alpha_j\right)\left(\pi\left(\alpha_i\right) - \pi\left(\alpha_j\right)\right) + \alpha_i \pi\left(\alpha_i\right)^3 - \alpha_j \pi\left(\alpha_j\right)^3\right] > 0,$$

where the first inequality follows because $k > \underline{k}$, and the last inequality follows because $\alpha_j \leq \alpha_i$ implies that $\pi\left(\alpha_j\right) \leq \pi\left(\alpha_i\right)$ by Lemma 1. Moreover, note that

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

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 (7) 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 $\alpha = 0$.

Third, in Proposition 2(i) below we prove that λ_j^* decreases with k. Then using (8),

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

where the upper bound on λ_j^* is its value as $k \to \underline{k} \equiv \pi\left(\alpha_i\right)\left(1 + \alpha_i \frac{\pi(\alpha_i)}{\pi(\alpha_j)}\right)$, 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 (7) and (8) 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 $k \to \infty$, (7) and (8) imply that $\lambda_i^* \to 0$ and $\lambda_j^* \to 0$.

■

Proof of Proposition 1.

(i) Recalling from Lemma 1 that $\pi'\left(\alpha_j\right) \geq 0$,

$$\frac{\partial \mathrm{BR}_i\left(\lambda_j\right)}{\partial \alpha_i} = -\frac{\lambda_j\left(\pi\left(\alpha_i\right) + \alpha_i \pi'\left(\alpha_i\right)\right)}{k} < 0, \quad \frac{\partial \mathrm{BR}_j\left(\lambda_i\right)}{\partial \alpha_i} = \frac{\left(1-\lambda_i\right)\pi'\left(\alpha_i\right)}{k} \geq 0.$$

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. The above implies that λ_i^* is largest under a symmetric PCO structure where $\alpha_i = \alpha_j$ and is smallest under a maximally asymmetric PCO structure where $\alpha_i \to \frac{1}{2}$ and $\alpha_j = 0$, and conversely for λ_j^*.

(ii) Using (7) and (8) again,

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

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

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

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

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

Hence,

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

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

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

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 (7) that

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

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

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

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

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

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 $\pi\left(\alpha_i\right) \geq \pi\left(\alpha_j\right)$ by Lemma 1, it follows that evaluated at $k \to \underline{k}$,

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

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

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

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

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

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

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

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

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

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

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

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

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 $\mathrm{BR}_i\left(\lambda_j\right)$, $\frac{\pi(\alpha_j)}{\pi(\alpha_j)+\alpha_i\pi(\alpha_i)}$, shifts up when c increases. Likewise, the horizontal intercept of $\mathrm{BR}_j\left(\lambda_i\right)$ shifts to the right. Given that both $\mathrm{BR}_i\left(\lambda_j\right)$ and $\mathrm{BR}_j\left(\lambda_i\right)$ shift outward, the sum of λ_i^* and λ_j^* increases.

(iii) When $\sigma \leq \alpha_j \leq \alpha_i$, $p\left(\alpha_i\right) = p\left(\alpha_j\right) = p^m$ and $\pi\left(\alpha_i\right) = \pi\left(\alpha_j\right) = \pi^m$. Substituting in (7) and (8), the equilibrium investments become

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

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 $\lambda_i^* \leq \lambda_j^*$,

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

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

Noting that $S'(p) < 0$ and $p'\left(\alpha_i\right) \geq 0$, it follows that $\frac{\partial \mathrm{CS}(\alpha_i,\alpha_j)}{\partial\alpha_i} \leq 0$ if the sum of the first two terms in (12) is negative, that is,

$$\text{(A1)} \qquad \frac{\partial \mathrm{CS}\left(\alpha_i,\alpha_j\right)}{\partial\alpha_i} = \frac{\partial \mathrm{CS}\left(\alpha_i,\alpha_j\right)}{\partial\lambda_i}\underbrace{\frac{\partial\lambda_i^*}{\partial\alpha_i}}_{(-)} + \frac{\partial \mathrm{CS}\left(\alpha_i,\alpha_j\right)}{\partial\lambda_j}\underbrace{\frac{\partial\lambda_j^*}{\partial\alpha_i}}_{(+)} \leq 0,$$

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 (A1) holds and are therefore sufficient for $\frac{\partial \mathrm{CS}(\alpha_i,\alpha_j)}{\partial\alpha_i} \leq 0$:

(i) If $\frac{\partial \mathrm{CS}(\alpha_i,\alpha_j)}{\partial\lambda_i} \geq 0 \geq \frac{\partial \mathrm{CS}(\alpha_i,\alpha_j)}{\partial\lambda_j}$, both terms of (A1) 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 (A1) and recalling from part (i) that $\frac{\partial CS(\alpha_i,\alpha_j)}{\partial \lambda_i} \geq \frac{\partial CS(\alpha_i,\alpha_j)}{\partial \lambda_j}$,

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

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

(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,

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

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

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

Proof of Corollary1. 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 $\sigma \leq \alpha_j \leq \alpha_i$. The result then follows immediately from Proposition 3(ii) and (iii). ■

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

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

$$+ \lambda_i^* \left(1 - \lambda_j^*\right) S'\left(p\left(\alpha_j\right)\right) p'\left(\alpha_j\right). \tag{A2}$$

If $\sigma \leq \alpha_j \leq \alpha_i$, then $p'\left(\alpha_j\right) = 0$ 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:

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

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

Proof of Lemma 3.

(i) It is clear from (13) 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 (13) 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 (13), 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,

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

where

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

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 $z > \underline{z}$,

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

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

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

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 (14) with respect to α_i, yields

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

where the sign of the derivative depends on the sign of

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

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 $0 \leq \alpha_j \leq \alpha_i < 1/2$.[51]

(ii) Differentiating (14) with respect to α_j, yields

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

[51] The command we use is Reduce[$H_i \geq 0$ && $z > \underline{z}$ && $0 \leq \alpha_i < 1/2$&& $0 \leq \alpha_j \leq \alpha_i$, $\{z, \alpha_i, \alpha_j\}$], where $\underline{z} \equiv \frac{1-\alpha_i\alpha_j}{\left(1-\alpha_i\right)^2}$. 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 $\frac{\partial \mathrm{CS}(\alpha_i,\alpha_j)}{\partial \alpha_j}$ is equal to the sign of H_j, where

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

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 $\alpha_j = \alpha_i = \alpha$,

$$\text{(A3)} \quad H_j = -(1-\alpha)^7\left(1+\left(z\left(1-\alpha\right)+\alpha\right)\left(z\left(1-\alpha^2\right)+\alpha^2\right)\right)\left(z-\frac{1+\alpha}{1-\alpha}\right)^3 < 0,$$

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 (15) that when $\alpha_j = \alpha_i$ and $m \to 1+\alpha_i$, $\lambda_i^* \to 0$, and

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

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,

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

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

(iii) By straightforward differentiation,

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

and

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

Recalling that $\alpha_j \leq \alpha_i$, $\frac{\partial \lambda_j^*}{\partial m} < 0$ for all $m > 1+\alpha_i$. Moreover, $\frac{\partial \lambda_i^*}{\partial m} \gtreqless 0$ if $m \lesseqgtr \widehat{m}$, where $\widehat{m}$ is defined by (16). Also note that when $\alpha_j = \alpha_i$, $\widehat{m} = 1+\alpha_i$, so $\frac{\partial \lambda_i^*}{\partial m} < 0$ for all feasible parameter values.

■

Proof of Proposition 6. Straightforward differentiation, using (15) and (17), yields

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

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

where

$$\begin{aligned} M_i \equiv & -\left(1+\alpha_i\right)\left(1+\alpha_j\right)\left(1-\alpha_j\sigma\right) \\ & + \left(1+\alpha_j\right)\left(2+\left(1-\alpha_i\right)\sigma\right) m - \left(1+\left(2+\alpha_j\right)\sigma\right) m^2 + \sigma m^3, \end{aligned}$$

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

(i) and (ii) M_i and M_j are a positive cubic function 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 $\alpha_j, \alpha_i \to 1/2$; by Assumption A5, then, $m > 1+\alpha_i \to 3/2$, implying that $M_i = M_j \to (m-3/2)^3/2 > 0$. Hence, in both cases, $\frac{\partial \mathrm{CS}(\alpha_i,\alpha_j)}{\partial \alpha_i} > 0$ and $\frac{\partial \mathrm{CS}(\alpha_i,\alpha_j)}{\partial \alpha_j} > 0$. At the other extreme, if $\sigma \to 0$ and $\sigma m \to 0$,

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

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

Now, note that

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

Using Mathematica, the derivative is strictly positive for $m > 1 + \alpha_i$ and all $0 \leq \alpha_j \leq \alpha_i < 1/2$;[52] Since $M_j > 0$ for $\sigma \to 1/2$, it follows by continuity that $M_j > 0$ for σ sufficiently large.

Next, note that,

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

Moreover, $\lim_{m \to 1+\alpha_i} \frac{\partial M_i}{\partial m} = -\left(\alpha_i - \alpha_j\right)\left(2 - \left(1 + 3\alpha_i\right)\sigma\right) < 0$, where the inequality follows because $\sigma \leq \alpha_i < 1/2$ implies that $2 - \left(1+3\alpha_i\right)\sigma > 2 - (1 + 3/2)/2 > 0$. 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,

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

and $\lim_{m \to 1+\alpha_i} \frac{\partial M_j}{\partial m} = \left(1+\alpha_i\right)\left(\alpha_i - \alpha_j\right)\sigma > 0$, $\lim_{m \to 1+\alpha_i} \frac{\partial^2 M_j}{\partial m^2} = -2\left(1 + \left(1 - 2\alpha_i\right)\sigma\right) < 0$, where the inequality follows because $\sigma \leq \alpha_i < 1/2$. 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 $\sigma m \to 0$,

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

which is negative if $m > \hat{m}$ and positive if $m < \hat{m}$. Hence, $\frac{\partial \mathrm{CS}\left(\alpha_i, \alpha_j\right)}{\partial \alpha_j} < 0$ if $m > \hat{m}$ and $\frac{\partial \mathrm{CS}\left(\alpha_i, \alpha_j\right)}{\partial \alpha_j} > 0$ if $m < \hat{m}$.

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

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

Hence, $\frac{\partial \mathrm{CS}\left(\alpha_i, \alpha_j\right)}{\partial \alpha_i} = \frac{\partial \mathrm{CS}\left(\alpha_i, \alpha_j\right)}{\partial \alpha_j} \lesseqgtr 0$ as $m \lesseqgtr \frac{1}{\sigma} - \alpha$.

■

Proof of Proposition 7. Recall that when the PCO structure is symmetric, $k > \pi(\alpha)(1+\alpha)$; since $\pi(\alpha) = \pi^m$ as $\alpha \geq \sigma$, (19) implies that $\lambda(\alpha) \equiv \frac{\pi^m}{k + (1+\alpha)\pi^m} > \frac{\pi^m}{2k}$. Hence,

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

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

$$= \frac{2k\left(\pi^m - \hat{\pi} + \pi_c^m\right) - \pi^m\left(\pi^m - \hat{\pi} + \pi_c^m - \hat{\pi}_c\right)}{2k\left(2k + \hat{\pi} + \hat{\pi}_c - \pi_c^m - \pi^m\right)}$$
$$= \frac{(2k - \pi^m)\left(\pi^m - \hat{\pi} + \pi_c^m\right) + \pi^m \hat{\pi}_c}{2k\left(2k + \hat{\pi} + \hat{\pi}_c - \pi_c^m - \pi^m\right)} > 0,$$

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

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

which is analogous to (20). Comparing CS^{**} with (20) and noting that $p_c^m > \widehat{p} > p^m > c > 0$, it is clear that $\mathrm{CS}(\alpha) > CS^{**}$. ■

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 (19) and (21), the equilibrium levels of investments under PCO are under RJV are given by

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

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

Using (17), expected consumer surplus under PCO and under RJV are given by

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

Notice that $\mathrm{CS}(1) = \mathrm{CS}^{\mathrm{RJV}}$ and that by (18), $\frac{\partial \mathrm{CS}(\alpha,\alpha)}{\partial \alpha} > 0$ if $m > \frac{1}{\sigma} - \alpha$ and $\frac{\partial \mathrm{CS}(\alpha,\alpha)}{\partial \alpha} < 0$ otherwise. Hence, for all $0 \leq \alpha < 1/2$, $\mathrm{CS}(\alpha) < \mathrm{CS}^{\mathrm{RJV}}$ if $m > \frac{1}{\sigma} - \alpha$ and conversely if $m < \frac{1}{\sigma} - \alpha$. ■

Proof of Proposition 9. Note that,

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

Since by definition $\pi^m > \pi_c^m$, $\lambda(\alpha) > \lambda^m$ if and only if (22) 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. ■

REFERENCES

Alley, W., 1997, 'Partial Ownership Arrangements and Collusion in the Automobile Industry,' *Journal of Industrial Economics*, 45, pp. 191–205.

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

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

Antón, M.; Ederer, F.; Giné, M. and Schmalz, M., 2022, 'Common Ownership, Competition, and Top Management Incentives,' *Journal of Political Economy*, 131(5), pp. 1294–1355.

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

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

Backus, M.; Conlon, C. and Sinkinson, M., 2021a, 'Common Ownership and Competition in the Ready-to-Eat Cereal Industry', Discussion Paper, National Bureau of Economic Research.

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

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

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

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

Bolle, F. and Güth, W., 1992, 'Competition Among Mutually Dependent Sellers,' *Journal of Institutional and Theoretical Economics*, 148, pp. 209–239.

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

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

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

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

Deneckere, R. and Kovenock, D., 1996. 'Bertrand-Edgeworth duopoly with unit cost asymmetry'. *Economic Theory*, 8, pp. 1–25.

Dietzenbacher, E.; Smid, B. and Volkerink, B., 2000, 'Horizontal Integration in the Dutch Financial Sector,' *International Journal of Industrial Organization*, 18, pp. 1223–1242.

European Commission, 2013, 'Towards More Effective EU Merger Control', Commission Staff Working Document. *http://ec.europa.eu/competition/consultations/2013_merger_control/*

Federico, G.; Langus, G. and Valletti, T., 2018, 'Horizontal Mergers and Product Innovation,' *International Journal of Industrial Organization*, 59, pp. 1–23.

Fershtman, C. and Gandal, N., 1994, 'Disadvantageous Semicollusion,' *International Journal of Industrial Organization*, 12, pp. 141–154.

Fiocco, R., 2016, 'The Strategic Value of Partial Vertical Integration,' *European Economic Review*, 89, pp. 284–302.

Flath, D., 1989, 'Vertical Integration by Means of Shareholding Interlocks,' *International Journal of Industrial Organization*, 7(3), pp. 369–380.

Flath, D., 1991, 'When is it Rational for Firms to Acquire Silent Interests in Rivals?' *International Journal of Industrial Organization*, 9, pp. 573–583.

Flath, D., 1992, 'Horizontal Shareholding Interlocks,' *Managerial and Decision Economics*, 13, pp. 75–77.

Ghosh, A. and Morita, H., 2017, 'Knowledge Transfer and Partial Equity Ownership,' *RAND Journal of Economics*, 48(4), pp. 1044–1067.

Gilo, D., 2000, 'The Anticompetitive Effect of Passive Investment,' *Michigan Law Review*, 99, pp. 1–47.

Gilo, D.; Moshe, Y. and Spiegel, Y., 2006, 'Partial Cross Ownership and Tacit Collusion,' *RAND Journal of Economics*, 37, pp. 81–99.

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

Heim, S.; Hüschelrath, K.; Laitenberger, U. and Spiegel, Y., 2022, 'The Anticompetitive Effect of Minority Share Acquisitions: Evidence from the Introduction of National Leniency Programs,' *American Economic Journal: Microeconomics*, 14, pp. 366–410.

Hunold, M. and Schlütter, F., 2021, 'Supply Contracts under Partial Forward Ownership,' LIDAM Discussion Paper CORE2022/03.

Hunold, M. and Stahl, K., 2016, 'Passive Vertical Integration and Strategic Delegation,' *RAND Journal of Economics*, 42, pp. 891–913.

Huse, C.; Ribeiro, R. and Verboven, F., 2024, 'Common-Ownership Versus Cross-Ownership: Evidence from the Automobile Industry,' *Journal of Industrial Economics*, 72, pp. 1360–1368.

Jullien B. and Lefouili, Y. 2018, 'Mergers and Investments in New Products', TSE Working Paper No. 18-949.

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

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

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

Li, Y. and Zhang, J., 2021, 'Product Positioning with Overlapping Ownership,' *Economics Letters*, 208, 110058.

López, Á. and Vives, X., 2019, 'Overlapping Ownership, R&D Spillovers, and Antitrust Policy,' *Journal of Political Economy*, 127(5), pp. 2394–2437.

Ma, H.; Qin, C. Z. and Zeng, C., 2021, 'Incentive and Welfare Implications of Cross-Holdings in Oligopoly,' *Economic Theory*, pp. 1–23.

Malueg, D., 1992, 'Collusive Behavior and Partial Ownership of Rivals,' *International Journal of Industrial Organization*, 10, pp. 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 Wang, Y., 2016, 'The Product Market Impact of Minority Stake Acquisitions,' *Management Science*, 64(2), pp. 825–844.

Nitta, K., 2008, Corporate ownership structure in Japan: recent trends and their impact. *NLI Research*, 2008(3), pp. 31.

OECD, 2012, *The Role of Efficiency Claims in Antitrust Proceedings*, Series Roundtables on Competition Policy No. 135, Available at *http://www.oecd.org/competition/EfficiencyClaims2012.pdf*.

Ono, H.; Nakazato, T.; Davis, C. and Alley, W., 2004, 'Partial Ownership Arrangements in the Japanese Automobile Industry; 1990-2000,' *Journal of Applied Economics*, 12(2), pp. 355–367.

Reiffen, D., 1998, 'Partial Ownership and Foreclosure: An Empirical Analysis,' *Journal of Regulatory Economics*, 13(3), pp. 227–244.

Reitman, D., 1994, 'Partial Ownership Arrangements and the Potential for Collusion,' *Journal of Industrial Economics*, 42, pp. 313–322.

Reynolds, R. and Snapp, B., 1986, 'The Competitive Effects of Partial Equity Interests and Joint Ventures,' *International Journal of Industrial Organization*, 4, pp. 141–153.

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

Schinkel, M. P. and Spiegel, Y., 2017, 'Can Collusion Promote Sustainable Production?' *International Journal of Industrial Organization*, 53, pp. 371–398.

Shelegia, S. and Spiegel, Y., 2012, 'Bertrand Competition when Firms Hold Passive Ownership Stakes in One Another,' *Economics Letters*, 114, pp. 136–138.

Spiegel, Y., 2013, 'Backward Integration, Forward Integration, and Vertical Foreclosure', CEPR Discussion Paper No. DP9617.

Stenbacka, R. and Van Moer, G., 2023, 'Overlapping Ownership and Product Innovation,' *International Journal of Industrial Organization*, 89, 102980.

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