Long-Run Market Configurations in a Dynamic Quality-Ladder Model with Externalities*

Mario Samano† Marc Santugini‡

February 10, 2020

Abstract

We study the impact of standard-setting by introducing an externality that increases product compatibility in the presence of asymmetric returns to investment in a dynamic quality-ladder-type model. We classify the long-run, multi-modal probability distributions over different market structures that arise from this model. In some cases, the lagging firm may remain in the market in the long-run depending on the strength of the externality. In the case where only the laggard invests in compatibility, it is possible that the laggard becomes a monopolist if the leader has a relatively low R&D capability and the two firms are almost symmetric in this same regard. This variety of multi-modal long-run distributions may have important consequences for the estimation and the simulation of this class of dynamic models.

Keywords: Quality-ladder model, industry dynamics, market structures, standard-setting, product compatibility.

JEL Classifications: C61, C73, L13.

*We thank participants at the SEA 2015, the IIOC 2015, the EARIE 2017, and the CEF 2019 conferences, and at the University of Alberta and Laval University seminar participants for helpful comments. M. Samano acknowledges financial support from the SSHRC and the FQRSC.

†HEC Montreal, Department of Applied Economics. E-mail: mario.samano@hec.ca

‡University of Virginia, Department of Economics. E-mail: ms3ae@virginia.edu

1 Introduction

We study a class of business strategies in which a leading firm causes a positive externality on the perceived quality of the good produced by its competitors. For instance, the decision of a firm to increase the degree of compatibility with its competitor's good, where the leading firm is the one with a higher capability of turning investments into larger installed bases. This strategy has been implemented in different industries in the form of standard-setting protocols.

Our research goal is to understand how the presence of this positive externality −due either to a firm's unilateral decision or to regulation− affects the leading firm and the industry overall. To that end, we model the quality of a product as a function of the sum of its own installed base and the spillover, which is the degree of compatibility, with a fraction of the competitor's installed base.[1] This aggregate of the installed bases enters directly into the utility function, which captures higher compatibility among the products in the market. This modelling approach can be interpreted as an endogeneous spillover process version of the model in Chen et al. (2009) with a cost to achieve compatibility and the decision can be unilateral. Firms decide the optimal amount of compatibility through a function that maps the competitor's installed base into the amount of that base that can be added to its own. We embed the associated maximization utility problem into a dynamic quality-ladder model in which firms differ in their return to investment on their own installed bases. That is, for a given level of investment, one firm has a higher probability to increase its installed base size. We show that such a model can generate different types of long-run market configurations (market collapse, monopolies, duopoly, and combinations of these cases). We find the array of possible market structures that can arise from this game for different parameter values. Although we are not the first ones to acknowledge the multi-modal nature of these long-run probability distributions in a dynamic quality-ladder-type model, we are the first to classify them and to analyze the impact of compatibility externalities on them.

R&D spillovers have been studied in the context of process innovation. In this case, the externality directly decreases the marginal cost of production (see for instance D'Aspremont and Jacquemin (1988), Kamien et al. (1992), and Samano et al. (2017)).[2] For product

[1]We follow the terminology used by Chen et al. (2009).

[2]D'Aspremont and Jacquemin (1988) and Kamien et al. (1992) consider deterministic reductions to marginal costs and Samano et al. (2017) extend those models to the case of dynamic and stochastic decisions in investment.

innovation, an important case is when the spillover increases the firm's capacity to transform investment into a higher probability of achieving a higher product's quality (see Song (2011), Goettler and Gordon (2011)). In this paper we take a different approach and consider only the cases where the spillover causes an increase in the compatibility of the goods in the industry by increasing their perceived quality by consumers as it would occur in a standard-setting process and we do not consider reductions in production costs.

Empirical evidence on the existence of technological spillovers has been documented by Bloom et al. (2013). They separate the technology spillovers from the product rivalry effect of R&D and document that even when taking into account these two effects, industries such as pharmaceuticals, computers, and telecommunications exhibit technology spillovers. But even in the absence of positive externalities on quality, firms have different likelihoods of success of investment. Goettler and Gordon (2011) estimated a dynamic quality-ladder model for the computer processors industry. They find evidence for heterogeneity in the likelihood of success of investment, which can explain differences in the levels of investment and ultimately differences in the levels of quality between the goods.[3] Another example of estimation of this class of models is found in Gowrisankaran and Town (1997). They consider two types of hospitals, for-profits and non-profits. The ratio of the number of these two types of hospitals is endogenous in their model. The parameter governing the probability of success of investment is restricted to be the same for the two hospital types, and yet, the observed market configurations in the data are not symmetric.[4]

Motivated by those findings, we ask the following questions. What is the effect of heterogeneity in the firms' ability to invest in their installed base on long-run market configurations? How does this form of heterogeneity interact with asymmetric compatibility spillovers? To answer them, we adapt the quality ladder model introduced by Ericson and Pakes (1995) and the algorithms to numerically solve for its equilibrium such as in Pakes and McGuire (1994) and in a particular case in Levhari and Mirman (1980) to the case of heterogeneous likelihood of success of investment with spillovers that affect compatibility.

We restrict our attention to the quality-ladder model without entry or exit. This is not

[3]In their adaptation of the Ericson-Pakes model, the source of this heterogeneity in the model is twofold: specific parameters for each firm and the quality distance between the leader and the follower. Specifically, they find a parameter value for the likelihood of success of investment of 0.0010 for Intel and 0.0019 for AMD. The estimated parameters are different for each firm, which captures the observed heterogeneity of firm dominance in their data.

[4]In the data, this ratio was observed to be 16, meaning that the for-profits hospitals dominate the market. That parameter of the likelihood of success of investment is estimated to be 0.51, well within our parameter space specification.

a strong assumption since we allow for installed base levels of zero which are equivalent to exit. However, that does not prohibit the same firm from becoming active again if it achieves to increase its installed base to a positive level in the next period. We also note that in our motivating example from Goettler and Gordon (2011), they do not consider entry and exit since the industry they study does not exhibit such behavior during the time window in their data.[5] In our second motivating example on the estimation of a quality-ladder model, Gowrisankaran and Town (1997) consider the possibility of entry and exit, however all hospitals belong to one of two firm types, and thus if all firms of one type exit, this is equivalent in our two-firm model to having an installed base of zero for one type of firm.[6]

Heterogeneity in the quality-ladder dynamic models has been studied in the context of capacity games. Besanko and Doraszelski (2004) conclude that asymmetries of firm size can be due to the effects of price competition which in turn lead to long-run distributions that exhibit positive probabilities on outcomes that represent only one firm surviving.[7] Their analysis keeps parameters symmetric across the two firms. We also find such configurations in cases of symmetric firms, but those configurations can arise from other parameter combinations as well. The asymmetries in price competition in their model arise because of small asymmetries in capacity accumulation that occur accidentally which makes one firm slightly dominant over the other, making the other firm to give up if investment is highly reversible. In Borkovsky et al. (2010) and Borkovsky et al. (2012), it is shown that the dynamic quality-ladder model can exhibit multiplicity of equilibria even in the absence of entry or exit if the investment is highly permanent. We take a different approach and allow firms to have different parameter values in their investment success function and study the long-run distribution over the installed base space given the unique equilibrium policies.[8] We also abstract from collaborations in R&D such as research joint ventures that could lead to a different type of externalities (see Samano et al. (2017) and Cellini and Lambertini (2009)) and from interactions between free-riding firms (surfers) and R&D-oriented firms (see Abdelaziz et al.

[5]Goettler and Gordon (2011) pp. 1151., although extensions with entry and exit can be found in Goettler and Gordon (2014).

[6]Another class of models has also been inspired by the quality-ladder model in which the focus is on entry when firms have multiple locations and there is a trade-off between cannibalization and preemption (see Igami and Yang (2016) and the references therein).

[7]This behavior was not found under quantity competition.

[8]In Figure 5 from Borkovsky et al. (2010), they provide evidence on the existence of multiple equilibria for depreciation rates below 0.1. Our analysis uses depreciation rates above or equal to that level and we check for potential multiplicity of equilibria solving the game in consecutive finite time horizons versions of the model a la Levhari and Mirman (1980).

(2008)).

We find that asymmetries in the likelihood of success of investment can have relevant effects on long-run market configurations, which highlight the richness of the baseline model. Even though the externality may be beneficial to decrease the outside good market share, it could harm the leader and allow the lagging firm to remain in the market if the asymmetry in the externality is above certain level or even make the laggard to become the monopolist if the leader does not invest in increasing compatibility but the laggard does. Therefore, increasing compatibility can eliminate the advantage of the leader for a certain range of industry parameters. This is important since it shows that helping to increase compatibility with competitors is not always harmful and it explains some of the motivations behind the existence of standard-setting protocols. We show that when the laggard is the only firm that can invest in absorbing the competitor's quality, the likelihood of observing duopolies in the long-run increases relative to the case of symmetric externalities. We also analyze the effects of depreciation rates that depend on the level of the installed base, the effects of the consumers' sensitivity to price levels, and the interaction between the depreciation rate and the curvature of the utility function.

This paper also has implications for the simulation of this type of models. Typically, one obtains data from an industry, say in a duopoly, and assume this is the equilibrium. Then a set of parameters is obtained by estimating a dynamic model of competition that belongs to the class of models we analyze here. We show that it is possible that when simulating the industry with the estimated parameters, additional market structures arise in the long-run. If we only report expected values for the different outcomes of the model as it is the most common practice, it is possible that salient information is being masked since the multiplicity of modes in the probability distribution is not properly reflected in those expected values.

The remainder of this article has the following structure. Section 2 discusses some standard-setting examples. Section 3 introduces the model. In Section 4 we provide computational details and the parametrization of the model. Section 5 presents the main results. We discuss further connections to the literature and conclude in Section 6.

2 Standard-setting Practices

Compatibility can be achieved through different business strategies. Generally they consist of an investment such that, if successful, the products from two or more different firms can function in a compatible manner. Firms may enter into formal agreements to implement such mechanism. For example, video compressing algorithms need to be compatible across different platforms, which has been achieved through agreements between companies such as Motorola and Microsoft. A standard-setting organization (SSO) is one type of such agreements. They agglomerate different firms that have similar technological needs and sell products that can benefit from higher compatibility, which can be achieved by sharing patents. Rysman and Simcoe (2008) identified 724 U.S. patents disclosed in these major SSOs: the American National Standard Institute (ANSI), the Institute for Electrical and Electronic Engineers (IEEE), the Internet Engineering Task Force (IETF) and the International Telecommunications Union (ITU). The IEEE for instance, developed the standard Wi-Fi. As noted by these authors, "linking a disclosure to a particular standard is often quite difficult" but eventually the firms reach a consensus.[9] In our model, whenever the agreement is reached, the compatibility spillover increases the consumers' utility through the function that aggregates the firm's own installed base level and that of the other firm. This is our stylized model of the *marginal effect* in Rysman and Simcoe (2008): by creating an open standard, firms start using certain patents that otherwise they would not which increases their overall installed base.[10]

This does not mean that all patents in a standard set by an SSO are free to be used nor that it is mandatory for a member of an SSO to make its patents available. If some parts of the standard depend on some protected patents by copyright, it is possible that the patent holder agrees to share the patent at a low cost. This was not exactly the case between Motorola and Microsoft in 2012. Motorola owned patents that needed to be used by Microsoft for the use of certain industry standards (IEEE 802.11 for WiFi communication and the H.264/MPEG-4 AVC, a video compression format). Motorola asked Microsoft for

[9] Bekkers et al. (2017) build a model to assess different rules of disclosure within an SSO. Using data on SSOs reveals that disclosure increases citations but also the probability that the patent ends up involved in a litigation process. Ganglmair and Tarantino (2014) model the interaction of two firms of which one of them has a private piece of information −a secret− and when it is optimal to release it so that the interaction does not stop. They argue for disclosure under Reasonable and Non-Discriminatory (RAND) terms as a way to decrease the inefficiencies in the equilibrium in their game. Lerner et al. (2016) develop a model of disclosure of patents under an SSO. Their dataset covers seven SSOs. The main result is that larger, downstream firms are less likely to disclose patents and this is an outcome in both the theory model and in the empirics.

[10] The delay to adopt the competitor's installed base is optimally determined by a function in our model that depends on the state at the which the firm is producing. Since at each period, the investment to expand on the installed base can change the state, so can the intensity of the compatibility. In some sense, the notion of delay in Chamley and Gale (1994) is present through the time that it may take a firm to move from one state in which there is no need to increase compatibility to a state in which it is optimal to allow for a positive value of the externality.

a royalty of 2.25% on the price of products using these patents but Microsoft complained and ultimately sued Motorola. The Court's decision was in favor of the plaintiff, setting a precedent on the low cost of patents covered by an agreement from an SSO.[11] This type of commitments is known as Fair, Reasonable and Non-Discriminatory (FRAND) terms.

Compatibility can also be achieved by luring the competitors to use the firm's own patents.[12] Recently, that has been the strategy taken by Tesla Motors, the manufacturer of electric vehicles (EVs). Its technology used in their recharging stations allows Tesla car owners to fully recharge the battery in about one hour, much below that of its competitors. In June 2014, Elon Musk, the CEO of that company, announced that they were releasing most of the company's patents because he said "We believe that Tesla, other companies making electric cars, and the world would all benefit from a common, rapidly-evolving technology platform."[13] In practice, this patent release is highly restrictive as it forces reciprocity in the use of the patents.[14] In 2015 and 2019, Toyota put in place a similar proposition but with more concrete details on how to operationalize the strategy. In particular, in 2019 Toyota made 2,200 charger-related patents available on a royalty-free basis on top of a few thousand more patents.[15] Most likely, what those two car manufacturers expect is that the industry converges to the same technology standard for recharging stations, implicitly creating higher product compatibility. This will increase this quality feature for all EV users, which will increase the market share of EVs, including Tesla Motors and Toyota's. Our model allows us to examine whether in equilibrium, this type of strategies can yield market dominance by the adopter of the technology or by the leader.

3 Model

In this section, we extend the Ericson-Pakes dynamic quality-latter model to the case in which each firm's valuation of the good sold depends not only on its own installed base level, but is

[11] Microsoft Corp. v. Motorola Inc., 854 F. Supp. 2d 993 (2012) Seattle, Washington.

[12] A notable case is the forced disclosure of patents by Bell Labs, see Watzinger et al. (2017).

[13] https://www.tesla.com/blog/all-our-patent-are-belong-you?

[14] The use of the term "good faith" in the pledge by Tesla Motors implies that if a competitor uses one of Tesla's patents, the competitor is implicitly accepting that Tesla Motors can also use the competitor's patents without any penalties: "[...] A party is "acting in good faith" so long as such party and its related or affiliated companies have not: asserted, helped others assert or had a financial stake in any assertion of (i) any patent or other intellectual property right against Tesla", https://www.tesla.com/about/legal#patent-pledge. We do not have knowledge of another company using any of these patents although Elon Musk has suggested in at least one interview without mentioning the names of such companies that this has happened.

[15] https://global.toyota/en/newsroom/corporate/27512455.html

also potentially influenced by the installed base level achieved by the other firm. For instance, in the case of the electric car industry, installed base refers to the availability and effectiveness of the technology in the recharging stations. An improvement in the compatibility across the different technologies for this purpose translates into a net expansion of the overall installed base and thus an increase in consumers' valuation for electric cars. There are two possible cases to study.

1. *No externalities.* There is no compatibility among the different firms. Then, an expansion in the installed base of one firm affects only consumers' valuation for its own good.

2. *Externalities.* There is imperfect compatibility. For instance, one firm's electric cars can recharge in almost any recharging station. Then, an expansion in the installed base of one firm's recharging stations affects (asymmetrically) consumers' valuation for all goods in that industry.

To study the long-run implications of such an industry, we must take account of heterogeneity. There are two kinds of heterogeneity worth considering.

1. *Compatibility externality.* The first layer of heterogeneity concerns the link between installed base and consumers' valuation. For instance, the leading firm might not benefit from installed base expansions of the lagging firm as much as the lagging firm would benefit from installed base expansions by the leading firm. This is represented by the function $\kappa_{\cdot}$ below.

2. *Likelihood of success of investment.* The technological ability to expand the installed base varies across firms, i.e., some firms are more capable than others of turning investment into a successful expansion. This is represented by the parameter $\alpha_{\cdot}$ below.

Relations to other models. The way we model the compatibility externality is a contribution to the literature. Our externality specification differs from that in Goettler and Gordon (2011) in that we consider an externality that enters the utility function only, whereas they use a specification for the probability of success of investment that encompasses an externality and it is relative to some frontier: it is more difficult to push the frontier forward than getting closer to it. Our specification is agnostic about the frontier process. The value of κ −the intensity of the spillover− can represent cases in which the laggard is simply riding off of the installed base expansions of the leader (if the spillover is high) or not at all (if $\kappa = 0$).

In addition, our specification also differs from that in Song (2011) because he models the externality in the probability function for investment success (as Goettler and Gordon (2011)) but in a multiplicative form, not relative to any frontier. Our externality specification does not affect the way in which the R&D process itself is done, but rather, on how it directly affects consumers' utility.

We now provide a detailed description of the Ericson-Pakes model under the presence of these two sources of heterogeneity. For simplicity, we restrict attention to the case of two firms and abstract from entry or exit.[16]

Demand. Each consumer has a utility function over quality and price of the good. Quality is given by a function $g(\cdot,\cdot;\cdot)$ that combines: (i) the firm's own installed base size and (ii) the competitor's base size that becomes compatible (a fraction of it). This ability to absorb changes in base size that are shared by others is measured by κ and represents the degree of compatibility. The function g is increasing in the installed base levels. Notice that this interpretation of the state variables in this game is similar to that in Chen et al. (2009) except that (i) we concentrate on the adoption of the installed base assuming that the other firm is always willing to accept and (ii) that there is a cost for this compatibility.

We consider a differentiated-product market in which two firms compete à la Bertrand as well as invest to increase their installed base and to increase compatibility. For $j = 1, 2$, let $\omega_j \in \{0, 1, 2, \ldots, M\}$ be firm j's installed base out of $M + 1$ possible levels. Given installed bases $\{\omega_1, \omega_2\}$ and prices $\{p_1, p_2\}$, firm j's demand is

$$D\left(p_j, p_{3-j}; \omega_j, \omega_{3-j}; \kappa_j, \kappa_{3-j}\right) = m \frac{e^{g_j(\omega_j,\omega_{3-j};\kappa_j)-\lambda p_j}}{1 + e^{g_j(\omega_j,\omega_{3-j};\kappa_j)-\lambda p_j} + e^{g_{3-j}(\omega_{3-j},\omega_j;\kappa_{3-j})-\lambda p_{3-j}}}$$

where $m > 0$ is the size of the market and

$$g_j(\omega_j, \omega_{3-j}; \kappa_j) = \begin{cases} -\infty, & \omega_j + \kappa_j\omega_{3-j} < 0 \\ \omega_j + \kappa_j\omega_{3-j}, & 0 \le \omega_j + \kappa_j\omega_{3-j} < \omega^* \\ \omega^* + \log(2 - \exp(\omega^* - \omega_j - \kappa_j\omega_{3-j}), & \omega^* \le \omega_j + \kappa_j\omega_{3-j} \le M \end{cases} \quad (1)$$

maps the firms' installed bases into consumer's valuation. The parameter $\omega^* \in \{0, 1, \ldots, M\}$

[16] As discussed in the introduction, one of our two empirical examples in the literature (Goettler and Gordon (2011)) does not consider entry or exit. Moreover, we allow for installed base levels of zero and the demand function in this case becomes null, this is equivalent to exiting the market.

reflects the level of the base after which there is a degree of satiation.[17] Equation 1 introduces heterogeneity in the model via the level of compatibility $\kappa_{\cdot}$. That is, consumers' valuation for good j depends on the base level achieved by firm $3-j$, i.e., ω_{3-j}. If $\kappa_1 = \kappa_2 = 0$, we obtain the baseline model without externalities.[18] When $\kappa_{\cdot} > 0$ there is a positive influence of the competitor's installed base level on the firm's own base. Differences between κ_1 and κ_2 mean that one firm benefits from the externality more than the other. If $\kappa_j < 0$ there is a negative effect on the compatibility of the two goods. Since our main motivation in this paper is compatibility and standardization business strategies, we concentrate on the case of positive externalities.

Compatibility. To determine the values of $\kappa_{\cdot}$ that are relevant for the analysis we take the derivative of the market share function with respect to its own installed base level and that of its competitor. Define $\tilde{D}_{\cdot} = D_{\cdot}/m$, then those expressions can be signed as follows

$$\frac{\partial D_j}{\partial \omega_j} = m \times \left((1-\tilde{D}_j)\frac{\partial g_j}{\partial \omega_j} - \tilde{D}_k \frac{\partial g_k}{\partial \omega_j} \right) \tilde{D}_j > 0 \text{ if } 0 \leq \kappa_k \leq 1$$
$$\frac{\partial D_j}{\partial \omega_k} = m \times \left((1-\tilde{D}_j)\frac{\partial g_j}{\partial \omega_k} - \tilde{D}_k \frac{\partial g_k}{\partial \omega_k} \right) \tilde{D}_j > 0 \text{ if } 0 \leq \kappa_j \leq 1$$

since $1 - \tilde{D}_j - \tilde{D}_k > 0$ and $\frac{\partial g_j}{\partial \omega_j} = \frac{\partial g_k}{\partial \omega_k} = 1$, $\frac{\partial g_k}{\partial \omega_j} = \kappa_k$, and $\frac{\partial g_j}{\partial \omega_k} = \kappa_j$ if $\omega_j + \kappa_j \omega_k < \omega^*$ for $j \neq k$. This is consistent with intuition: increases in its own base should reflect a higher number of customers and increases in the competitor's good's installed base should spillover in a positive way to also attract more customers. Therefore, to simplify the analysis we will focus on the unit square for the parameters $\kappa_{\cdot}$. However, when $\omega_j + \kappa_j \omega_{3-j} \geq \omega^*$ the derivatives of g with respect to $\omega_{\cdot}$ are negative and therefore, the effects of $\omega_{\cdot}$ on market shares can be positive or negative but since demand responds in a negligible way beyond ω^* (satiation property), we take an agnostic stand on what the effects are from κ for values of ω beyond ω^*.

In the definition of the equilibrium of this game we describe how we optimize over $\kappa_{\cdot}$ and its interaction with optimal prices.

[17]The last two lines in the specification of the function g reflect a satiation effect at high installed base levels (above ω^*). This creates decreasing marginal returns on the utility function. The derivative of g is the same from the left and from the right of ω^*. We provide a sensitivity analysis on this parameter after presenting the main results.

[18]Note also that our specification in Equation 1 is similar to Borkovsky et al. (2012) in that $\omega_j = 0$ drives firm j's demand to zero. Although entry or exit are not explicitly modeled here, the state $(\omega_1, \omega_2) = (0, 0)$ essentially leads to a temporary collapse of the market. We call it a temporary collapse since firms can still successfully invest in the next period to go back into the game. In other words, it is possible that for a particular set of parameters even if $(\omega_1, \omega_2) = (0, 0)$, firms' optimal policy functions are positive at that state.

Profits. Firm j's instantaneous profits are

$$\pi\left(p_j, p_{3-j}; \omega_j, \omega_{3-j}; \kappa_j, \kappa_{3-j}\right) = D\left(p_j, p_{3-j}; \omega_j, \omega_{3-j}; \kappa_j, \kappa_{3-j}\right)\left(p_j - c\right)$$

where $c > 0$ is the constant marginal cost of production, same across firms. Because market competition has no effect on the dynamics, the pricing game is static. Let $\Pi\left(\omega_j, \omega_{3-j}; \kappa_j, \kappa_{3-j}\right)$ be firm j's instantaneous profit corresponding to the static Bertrand game.[19]

Investment. At the same time, there is the effect from success of in-house investment. This is the process by which money is mapped into a probability of success of achieving an expansion of installed base. This is measured by the parameter α.

Each period, firm j invests an amount $x_j \geq 0$ intended to expand the base. This process is stochastic and subject to an industry-wide shock. Specifically, firm j's base evolves stochastically as

$$\omega_j'|\omega_j = \min\{\max\{\omega_j + \tau_j + \eta, 0\}, M\}$$

where τ_j is a firm-specific shock and η is an industry-wide depreciation shock. Each random variable is binary. The firm-specific shock η has support $\{0, 1\}$ and depends on the amount of investment, i.e.,

$$\Pr(\tau_j = 1|x_j) = \frac{\alpha_j x_j}{1 + \alpha_j x_j} \equiv \phi_j(x_j)$$

is firm j's probability of success conditional on investing $x_j \geq 0$. Here, $\alpha_j > 0$ is specific to firm j, which is our second source of parameter heterogeneity.[20] The industry-wide depreciation shock has support $\{-1, 0\}$ such that

$$\Pr(\eta = -1) = \delta \in [0, 1]$$

is the probability of base depreciation. Equivalently, δ is the rate of improvement of the outside good relative to the inside goods.

Value Function. Before proceeding with the definition and characterization of the equilibrium, it is useful to write down the firm's value function taking as given the behavior

[19]That is, for $j = 1, 2$, $\Pi\left(\omega_j, \omega_{3-j}; \kappa_j, \kappa_{3-j}\right) = D\left(p_j^*, p_{3-j}^*; \omega_j, \omega_{3-j}; \kappa_j, \kappa_{3-j}\right)\left(p_j^* - c\right)$ where the pair $\{p_1^*, p_2^*\}$ is the Bertrand equilibrium defined as $p_j^* = \arg\max_{p_j > 0} D_j\left(p_j, p_{3-j}^*; \omega_j, \omega_{3-j}; \kappa_j, \kappa_{3-j}\right)\left(p_j - c\right)$. For all $\{\omega_1, \omega_2\}$, there exists a unique Bertrand-Nash equilibrium (Caplin and Nalebuff (1991)).

[20]In Goettler and Gordon (2014), they parametrize α as a function of the distance to the technological frontier and mention that this can encompass the skills to reverse engineering, but it still means that compatibility is not captured by this function nor somewhere else. It also implies that reverse engineering is as costly as in-house R&D.

The specific values for α_j we use in our simulations lie well within those in the literature (Goettler and Gordon (2011), Gowrisankaran and Town (1997), Borkovsky et al. (2010)).

of the other firm. Specifically, for $j = 1, 2$, given x_{3-j}, firm j's infinite-horizon value function satisfies

$$V_j(\omega_j, \omega_{3-j}) = \max_{\substack{x_j \geq 0 \\ 0 \leq \kappa_j \leq 1}} \left\{ \Pi(\omega_j, \omega_{3-j}; \kappa_j, \kappa_{3-j}) - x_j - \gamma\kappa_j + \beta \mathbf{E}[V_j(\omega_j', \omega_{3-j}')|\omega_j, \omega_{3-j}, x_j, x_{3-j}] \right\}$$

where a prime sign indicates a variable in the subsequent period. The expected continuation value function is written as

$$\begin{aligned} &\mathbf{E}[V_j(\omega_j', \omega_{3-j}')|\omega_j, \omega_{3-j}, x_j, x_{3-j}] \\ &= \phi_j(x_j)\phi_{3-j}(x_{3-j}) \cdot \left(\delta V_j(\omega_j, \omega_{3-j}) + (1-\delta)V_j\left(\omega_j^+, \omega_{3-j}^+\right)\right) \\ &\quad + \phi_j(x_j)(1-\phi_{3-j}(x_{3-j})) \cdot \left(\delta V_j\left(\omega_j, \omega_{3-j}^-\right) + (1-\delta)V_j\left(\omega_j^+, \omega_{3-j}\right)\right) \\ &\quad + (1-\phi_j(x_j))\phi_{3-j}(x_{3-j}) \cdot \left(\delta V_j\left(\omega_j^-, \omega_{3-j}\right) + (1-\delta)V_j\left(\omega_j, \omega_{3-j}^+\right)\right) \\ &\quad + (1-\phi_j(x_j))(1-\phi_{3-j}(x_{3-j})) \cdot \left(\delta V_j\left(\omega_j^-, \omega_{3-j}^-\right) + (1-\delta)V_j(\omega_j, \omega_{3-j})\right) \end{aligned} \tag{2}$$

with

$$\omega_j^+ \equiv \min\{\omega_j + 1, M\}, \tag{3}$$

$$\omega_{3-j}^+ \equiv \min\{\omega_{3-j} + 1, M\}, \tag{4}$$

$$\omega_j^- \equiv \max\{\omega_j - 1, 0\}, \tag{5}$$

$$\omega_{3-j}^- \equiv \max\{\omega_{3-j} - 1, 0\}. \tag{6}$$

Given an initial state (ω_j, ω_{3-j}), Equation 2 summarizes all possible changes in the states corresponding to investment levels (x_j, x_{3-j}). Notice that there are two costs, one for the investment amount x_j and another for the cost of making the competitor's good compatible $\gamma\kappa_j$. The parameter γ is assumed to be the same for both firms and its only purpose is to transform the level of the externality (which is a value between 0 and 1) into monetary units. We abstract from the strategic behavior over the pricing of the licensing of standards to achieve compatibility (see Chiao et al. (2007)). We explain further details on the value of this parameter below.

Equilibrium. We restrict attention to Markov-perfect equilibrium (MPE) in pure strategies. The tuple $\{X_1(\omega_1, \omega_2), X_2(\omega_2, \omega_1), \kappa_1(\omega_1, \omega_2), \kappa_2(\omega_2, \omega_1)\}$ is an equilibrium if, for $j=1,2$, given $X_{3-j}(\omega_{3-j},\omega_j)$ and $\kappa_{3-j}(\omega_{3-j},\omega_j)$,

$$\begin{aligned}X_j(\omega_j,\omega_{3-j}) = \arg\max_{x_j\geq 0}\{&\Pi\left(\omega_j,\omega_{3-j};\kappa_j,\kappa_{3-j}\right) - x_j - \gamma\kappa_j \\ &+\beta\mathbf{E}[V_j(\omega_j',\omega_{3-j}')|\omega_j,\omega_{3-j},x_j,X_{3-j}(\omega_{3-j},\omega_j)]\}\,, \\ \kappa_j(\omega_j,\omega_{3-j}) = \arg\max_{0\leq\kappa_j\leq 1}\{&\Pi\left(\omega_j,\omega_{3-j};\kappa_j,\kappa_{3-j}\right) - x_j - \gamma\kappa_j \\ &+\beta\mathbf{E}[V_j(\omega_j',\omega_{3-j}')|\omega_j,\omega_{3-j},x_j,X_{3-j}(\omega_{3-j},\omega_j)]\}\end{aligned}$$

where for any $(\omega_j,\omega_{3-j})\in\{0,1,...,M\}^2$, the value function satisfies

$$\begin{aligned}V_j\left(\omega_j,\omega_{3-j}\right) = &\Pi\left(\omega_j,\omega_{3-j};\kappa_j,\kappa_{3-j}\right) - X_j(\omega_j,\omega_{3-j}) - \gamma\kappa_j(\omega_j,\omega_{3-j}) \\ &+\beta\mathbf{E}[V_j(\omega_j',\omega_{3-j}')|\omega_j,\omega_{3-j},X_j(\omega_j,\omega_{3-j}),X_{3-j}(\omega_{3-j},\omega_j)]\end{aligned} \quad (7)$$

and $\mathbf{E}[V_j(\omega_j',\omega_{3-j}')|\omega_j,\omega_{3-j},X_j(\omega_j,\omega_{3-j}),X_{3-j}(\omega_{3-j},\omega_j)]$ has the same form as Equation 2.

From the first order condition with respect to X_j we obtain for $j=1,2$,

$$X_j(\omega_j,\omega_{3-j}) = \max\left\{\frac{-1}{\alpha_j}+\sqrt{\frac{\beta}{\alpha_j}}\sqrt{\frac{\alpha_{3-j}X_{3-j}(\omega_{3-j},\omega_j)\Delta_j+\Psi_j}{1+\alpha_{3-j}X_{3-j}(\omega_{3-j},\omega_j)}},0\right\} \quad (8)$$

when $\frac{\alpha_{3-j}X_{3-j}(\omega_{3-j},\omega_j)\Delta_j+\Psi_j}{1+\alpha_{3-j}X_{3-j}(\omega_{3-j},\omega_j)}\geq 0$ and $X_j(\omega_j,\omega_{3-j})=0$ otherwise. Here, using Equation 3 - Equation 6,

$$\begin{aligned}\Delta_j \equiv\ &\delta\left[V_j\left(\omega_j,\omega_{3-j}\right) - V_j\left(\omega_j^-,\omega_{3-j}\right)\right] \\ &+(1-\delta)\left[V_j\left(\omega_j^+,\omega_{3-j}^+\right) - V_j\left(\omega_j,\omega_{3-j}^+\right)\right], \\ \Psi_j \equiv\ &\delta\left[V_j\left(\omega_j,\omega_{3-j}^-\right) - V_j\left(\omega_j^-,\omega_{3-j}^-\right)\right] \\ &+(1-\delta)\left[V_j\left(\omega_j^+,\omega_{3-j}\right) - V_j\left(\omega_j,\omega_{3-j}\right)\right].\end{aligned}$$

The derivative of V_j with respect to κ_j given κ_{3-j} is

$$\frac{\partial V_j}{\partial\kappa_j} = \begin{cases} m\times(p_j-c)\tilde{D}_j(1-\tilde{D}_j)\omega_{3-j}-\gamma, & \text{if } 0\leq\omega_j+\kappa_j\omega_{3-j}<\omega^*, \\ m\times(p_j-c)\tilde{D}_j(1-\tilde{D}_j)\frac{\omega_{3-j}}{2\exp(-(\omega^*-\omega_j-\kappa_j\omega_{3-j}))-1}-\gamma, & \text{if } \omega^*\leq\omega_j+\kappa_j\omega_{3-j}<M,\end{cases} \quad (9)$$

and the derivative when $\omega_j+\kappa_j\omega_{3-j}<0$ is undefined. From these expressions we can find the policy function for κ_j by setting $\partial V_j/\partial\kappa_j=0$, which is a function of the competitor's value for κ_{3-j} and therefore, it is the reaction function for this choice variable. Each of these expressions depends on the optimal prices through $\tilde{D}_.$. In addition, the actual choice of which of the two expressions from Equation 9 to use at each state depends on the value of $\kappa_{\cdot}$ itself. We find the solution to this system of equations by iterating with the solution to the system of equations for the optimal prices and by using the value of $\kappa_{\cdot}$ from the previous iteration to evaluate the inequalities in Equation 9. More specifically, we start with a value of 0 for the externalities at every state and solve for the optimal prices. Then we solve for the optimal values of $\kappa_{\cdot}$ using that first set of optimal prices. With the new value of the externalities at each state, we solve for the optimal prices. We repeat this until we obtain convergence of the solutions to both systems of equations and at each state. Although we do not have a mathematical proof that this iteration method will always converge, we have not found a case in which convergence is not attained as long as γ is sufficiently high. All this can be done "off line" in a similar manner than for the matrices of prices in the case of the classic model with no externalities.

3.1 Strategic Complementarity and Substitutability in Investment

From the first order condition for X_j we can also obtain the rate at which firm j's investment changes with respect to the other firm's level of investment:

$$\left.\frac{\partial X_j}{\partial X_{3-j}}\right|_{(\omega_j,\omega_{3-j})} = \frac{\beta \frac{\alpha_{3-j}}{(1+\alpha_{3-j}X_{3-j}(\omega_{3-j},\omega_j))^2}(\Delta_j - \Psi_j)}{2(1+\alpha_j X_j(\omega_j,\omega_{3-j}))}.$$

All the factors in that expression are always positive except for the term $\Delta_j - \Psi_j$. Therefore, we would observe strategic substitutes in investment if $\Delta_j - \Psi_j < 0$ and complements if $\Delta_j - \Psi_j > 0$. The sign of this term depends on the degree of concavity of V, which in turn depends on its initial condition: the static profits, which is the only object that contains the information on the externality. We are unable to determine the exact mechanism by which the level of the externalities changes the curvature of this function but as explained in section 5, we do not find evidence for strategic complements. This also suggests that there cannot be multiple solutions to the investment problem at each time period since the two reaction curves are decreasing and therefore they can only intersect each other at most once. If both reaction curves were increasing and concave, it would be possible to have more than one intersection. A simple inspection of the second derivative of the reaction function indicates that this could occur for any value of α_j and $\Delta_j - \Psi_j > 0$ (strategic complements).[21]

[21]By implicit differentiation, $X_j'' = -\frac{\alpha_j X_j'^2 + \alpha_k \beta(\Delta_j - \Psi_j)\frac{1}{(1+\alpha_{3-j}X_{3-j})^3}}{1+\alpha_j X_j}$.

4 Computation and Parametrization

We use the Pakes-McGuire (PM) algorithm to numerically solve for $\{X_1(\omega_1,\omega_2), X_2(\omega_2,\omega_1)\}$ and $\{V_1(\omega_1,\omega_2), V_2(\omega_2,\omega_1)\}$. Since firms can be heterogeneous, i.e., $\alpha_1 \neq \alpha_2$, the algorithm consists of iterating on the best response operators until convergence is reached. Specifically, at the initial iteration $t=0$, we set

$$\{X_1^0(\omega_1,\omega_2), X_2^0(\omega_2,\omega_1)\} = \{0,0\},$$

for all combinations of states (ω_1,ω_2) and the corresponding value functions

$$\{V_1^0(\omega_1,\omega_2), V_2^0(\omega_2,\omega_1)\} = \{\Pi(\omega_1,\omega_2;\kappa_1), \Pi(\omega_2,\omega_1;\kappa_2)\}.$$

For iteration $t = 1,2,...,$ given $\{X_1^{t-1}(\omega_1,\omega_2), X_2^{t-1}(\omega_2,\omega_1)\}$ and $\{V_1^{t-1}(\omega_1,\omega_2), V_2^{t-1}(\omega_2,\omega_1)\}$, we construct X_1^t and X_2^t according to the reaction function given by Equation 8 where the policy functions on the right hand side of that equation and the terms that depend on value functions are all indexed by the previous time period, that is X_j^{t-1} and V_j^{t-1}.

Moreover, the value functions are defined by Equation 7 and updated as follows:

$$\begin{aligned} V_1^t(\omega_1,\omega_2) &= \Pi(\omega_1,\omega_2;\kappa_1,\kappa_2) - X_1^t(\omega_1,\omega_2) - \gamma\kappa_1(\omega_1,\omega_2) \\ &\quad + \beta\mathbf{E}[V_1^{t-1}(\omega_1',\omega_2')|\omega_1,\omega_2,X_1^{t-1}(\omega_1,\omega_2),X_2^{t-1}(\omega_2,\omega_1)], \\ V_2^t(\omega_2,\omega_1) &= \Pi(\omega_2,\omega_1;\kappa_1,\kappa_2) - X_2^t(\omega_2,\omega_1) - \gamma\kappa_2(\omega_2,\omega_1) \\ &\quad + \beta\mathbf{E}[V_2^{t-1}(\omega_2',\omega_1')|\omega_2,\omega_1,X_2^{t-1}(\omega_2,\omega_1),X_1^{t-1}(\omega_1,\omega_2)]. \end{aligned}$$

The algorithm stops when some convergence criterion for the value functions and the policy functions is met.

In the PM algorithm, the computed levels of investment at each iteration do not necessarily constitute an equilibrium since the best responses (in terms of investment) at iteration t are in reaction to the investments computed at iteration $t-1$. However, stationary points of such iterations are MPEs. In addition to the PM algorithm, we also apply the algorithm suggested by Levhari and Mirman (1980) (LM) in a resource extraction dynamic game. The algorithm consists of computing the equilibrium for any finite horizon and increasing the horizon (making use of the computation for shorter horizons) until convergence is met. Unlike the PM algorithm, the levels of investment computed under the LM algorithm at each iteration constitute a Markov-perfect equilibrium. In our numerical analysis, we compute the equilibrium using both algorithms, which always leads to the same converged policy functions. The algorithm that computes the limit of a finite horizon game has been applied in the context of the Ericson-Pakes framework in Goettler and Gordon (2011) and in Chen et al. (2009). A description of the LM algorithm is relegated to the Online Appendix. We note that the PM algorithm is much faster than the LM algorithm. However, the latter allows us to guarantee that the reaction functions cross at most once which suggests uniqueness of equilibrium. We discuss this further in the next two sections.

We use the same parameter values as in Borkovsky et al. (2010) except that we use a slightly higher value for the price sensitivity λ, we use 1.2 and 1.7 instead of 1 simply to obtain a wider range of variation in long-run outcomes over our interval for $\alpha_\cdot$, equivalently we could have lowered the value of λ and increased the range of α.[22]

For the cost of implementing the compatibility to absorb the installed base from the competitor (the parameter γ) we choose a value that is comparable to the cost of increasing the installed base by one unit through the investment policy $X_\cdot$. If for a given state (ω_1, ω_2) the amount to be invested is $X_\cdot(\omega_1, \omega_2) = 20$, then the probability of success of this investment is approximately 99%, which we obtain by evaluating $\phi(20)$ with $\alpha_\cdot = 5$. If firm 1 decides to absorb the installed base of its competitor, ω_2, at the optimal level of compatibility κ_1, this should cost at least $20 \times \kappa_1 \times \omega_2$ because at this cost it would be almost certain that such an increase in the installed base could be done through the regular investment $X_\cdot$ function over 20 time periods if there is not an industry-wide shock. The maximum amount that could ever be spent in such an absorption of the competitor's installed base would be when $\omega_2 = M$. Therefore, we set $\gamma = 20 \times M$. Other parameterizations are possible but they do not largely affect the shape of the κ function. See the Appendix for a sensitivity analysis on this parameter.

Table 1: Parameter values

Parameter	M	m	c	ω^*	β	λ	α_j	γ	δ
value(s)	18	5	5	12	0.925	{1.2, 1.7}	[0.1, 5]	360	0.1

[22] One interpretation of the parameter M is due to Goettler and Gordon (2011) where the state space consists of discrete steps of log quality levels, and such levels are unbounded. In each time period, an upper bound is defined by the maximum level of quality in that same period, and all the points in the state space are then normalized with respect to that maximum. Then the model is re-written in terms of this normalized state space which then has the same structure as the model in Pakes and McGuire (1994).

4.1 Long-run Distributions

Let $\mathbf{a}_t = [a_{t,0}, \ldots, a_{t,(M+1)^2}]$ be a vector of size $1 \times (M+1)^2$ where $a_{t,s}$ is the probability that the industry is in state $s = (\omega_j, \omega_k)$ at time t such that $\sum_s a_{t,s} = 1$.

Let $\mathbf{P}$ be a $(M+1)^2 \times (M+1)^2$ transition matrix such that each element provides the probability to transition from one industry state to another, i.e., $\Pr[(\omega_j', \omega_k') | (\omega_j, \omega_k)]$.[23] We can obtain the transient distribution at each time t by using the sequence

$$\mathbf{a}_t = \mathbf{a}_{t-1}\mathbf{P} \tag{10}$$

and therefore

$$\mathbf{a}_t = \mathbf{a}_0\mathbf{P}^t$$

where $\mathbf{a}_0$ is the initial distribution over the state space. For a given set of parameters, we obtain the converged policy functions $X^*(\omega_j, \omega_k)$ and use them to calculate $\mathbf{P}$. When this matrix has only one left eigenvalue equal to one, the limiting distribution $\mathbf{a}^*$ exists and satisfies

$$\mathbf{a}^* = \mathbf{a}^*\mathbf{P},$$

which guarantees that the limiting distribution is independent of the initial condition $\mathbf{a}_0$ and there is only one recurrent class. However, for some parameter values, we find two or more recurrent classes by using the sequence given in Equation 10. Therefore, all of our results are obtained by using this sequential method starting from a uniform distribution over the state space. We stop when a converging criterion is met, refer to this distribution as the long run distribution $\mathbf{a}^*$. We also check that we obtain the same distribution if $\mathbf{a}_0$ is a degenerate distribution and we do not find large qualitative differences in our results with respect to a uniform distribution initial condition. Only for small regions of the parameter space for α do we find discrepancies in the number of modes of the long run distribution. It is worth remarking that other studies of this type of models have also opted for considering the long-run transient distributions instead of the limiting distributions (see Borkovsky et al. (2010) Section 5.3).

Once we obtain the distribution $\mathbf{a}^*$, we reshape this vector into an $(M+1) \times (M+1)$ matrix $\tilde{\mathbf{a}}$ and we count the number of modes. Each of these modes represents the maximum probability of a specific market configuration.[24] We will treat this distribution over the space

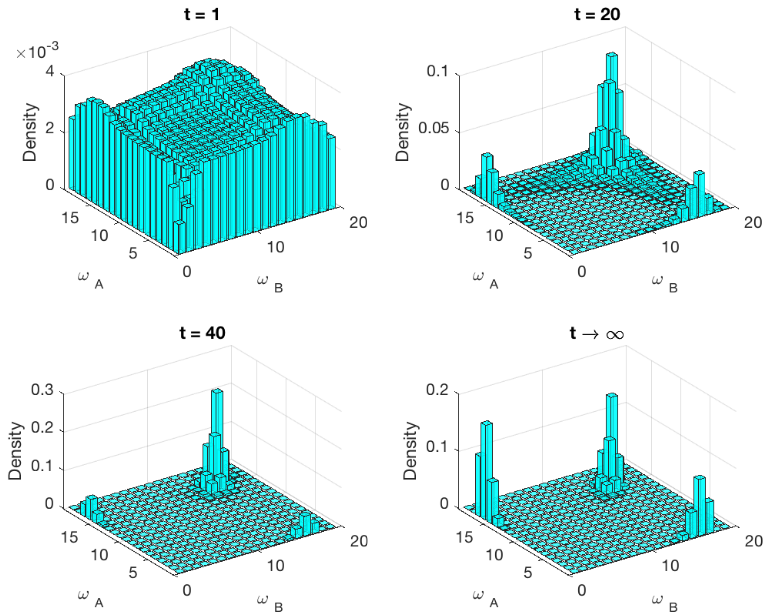
[23] The Online Appendix provides a detailed derivation of the transition matrix.

[24] We discard modes that have an associated probability of less than 10^{-3}. This threshold is equivalent to discard market structures that have an associated probability of less than 0.1% chance of occurring.

of installed base combinations as the full characterization of the market structures for a given set of parameters. It is worth discussing how this is equivalent to reporting the shapes of the distributions for the market shares for each firm. Equation 1 guarantees that all the values for the market shares are positive except when there is no installed base. When we multiply element by element the matrix of market shares over the (ω_1, ω_2) space for Firm 1 times $\tilde{\mathbf{a}}$ we obtain the same number of modes as in $\tilde{\mathbf{a}}$ except if there was a mode along the $q_1 = 0$ line. If there was such a mode, it will appear as a mode on the element by element product of the matrix of market shares for Firm 2 and $\tilde{\mathbf{a}}$.

In what follows we show that $\tilde{\mathbf{a}}$ might be unimodal (i.e., only one configuration occurs) or bimodal (two different market configurations are possible) or tri-modal (three different market structures can arise from the same set of parameters). Specifically, the market may collapse, i.e., installed base is driven to zero with probability one and firms do not sell anything. It is also possible to observe a duopoly. Finally, one firm may end up dominating, i.e., one firm owns a positive installed base amount, i.e., $\omega_j > 0$ whereas the other firm has no installed base, i.e., $\omega_{3-j} = 0$. For this case, it is possible to observe a realization in which the lagging firm dominates the market when there are externalities. Our main objective is to show that for a given set of parameter values, the resulting long-run distribution may imply a non-negligible probability for *different* market structures. Figure 1 shows an example in which $\tilde{\mathbf{a}}$ has three modes.

Figure 1: Transient distributions



Notes: Transient distributions from same policy function at different time periods. Initial distribution $\mathbf{a}_0$ is uniform.

In the next two sections, we provide a numerical analysis of the effect of heterogeneity on the long-run market structures. Also, to keep notation more tractable, instead of referring to firms by j and $3-j$, we will refer to them as A and B. We begin with the case of no externalities, i.e., $\kappa_A = \kappa_B = 0$ so that an expansion in the installed base of firm A has no effect on consumers' valuation for the good sold by firm B. In that case, we investigate how an advantage in the investment technology changes the equilibrium, which, in turn, affects the long-run market configurations. We then proceed with the case of externalities, i.e., $\kappa_A, \kappa_B > 0$. There, we show that the presence of an externality may make the leading firm to lose market dominance (vis-a-vis its competitor, the lagging firm) in order to increase the industry market share (with respect to the outside option). However, in some cases, the presence of the externality can also lead the lagging firm to dominate the market if the externality is more favorable to the lagging firm than to the leader.

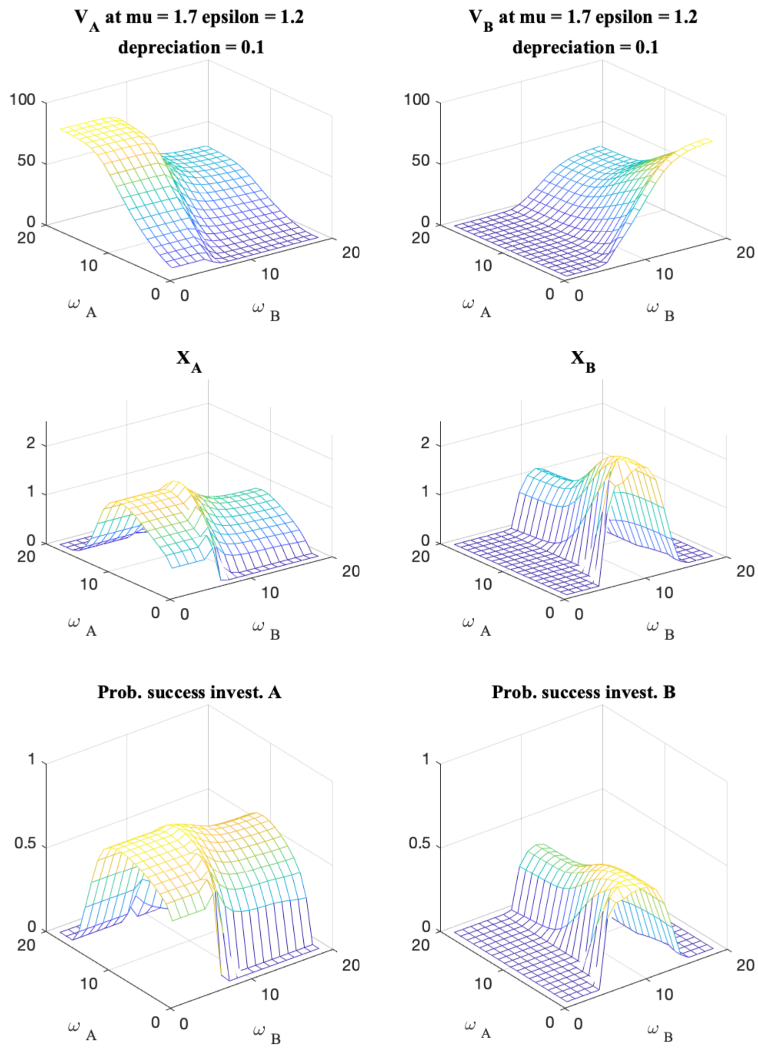
5 Market Structures in the Long-run

5.1 Market Structures in the Absence of Externalities

Suppose that $\kappa_A = \kappa_B = 0$ at every state, i.e., there are no externalities. We concentrate on the case where the probability of an industry-wide negative shock is $\delta = 0.1$ which does not cover the cases of multiplicity of equilibria discussed in Borkovsky et al. (2010). Higher values of δ simply increase the regions of a market collapse in our graphs below. To facilitate the discussion, we parametrize the heterogeneity in α_A and α_B as follows: $\alpha_A = \mu$ and $\alpha_B = \mu - \varepsilon$ where $\varepsilon \in [0, \mu]$ measures the heterogeneity in returns to investment between the two firms. Firm A is the leading firm and firm B is the lagging firm.

Differences between α_A and α_B have an effect on the equilibrium investment policy functions and the corresponding probabilities of success. Figure 2 shows the converged value and policy functions as well as the corresponding probabilities of success for each firm in the absence of externalities and for specific values of μ and ε.

Figure 2: Value, policy, and probability of success functions without externalities

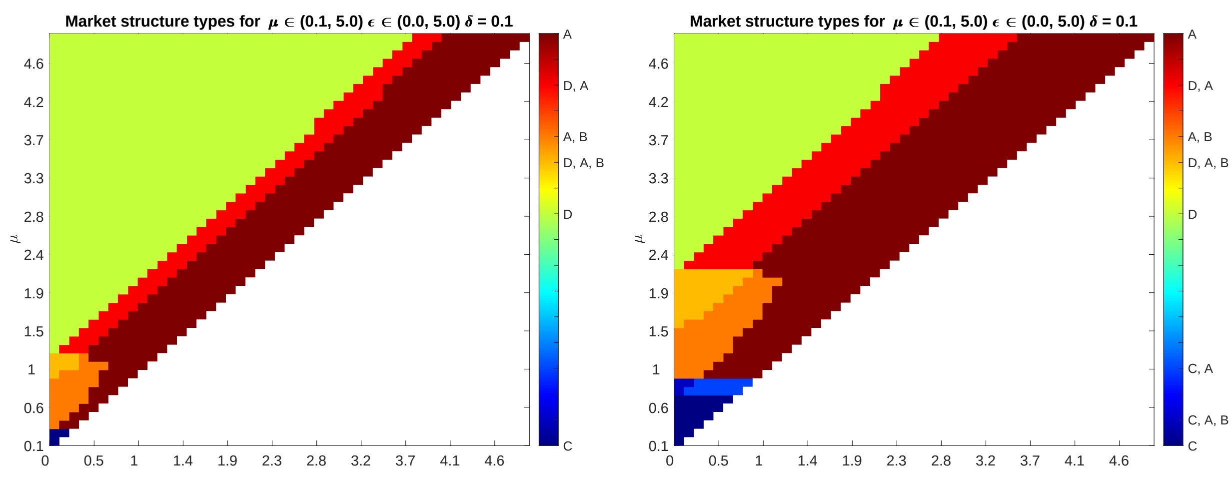


Notes: Asymmetric R&D capabilities with $\alpha_A = 1.7$ and $\alpha_B = 0.5$. Here $\lambda = 1.7$ and $\delta = 0.1$.

When the likelihood of success of investment is the same for both firms, the policy and value functions are identical (not shown on the graphs). However, when this likelihood is not the same across the two firms (in both Figures $\alpha_A > \alpha_B$), the lagging firm (the one with a lower α value, firm B on the graph) invests more in some states to compensate for this low likelihood of success. Because of the low probability of success and the higher amount of money spent in the investment, firm B receives in the long run a lower stream of cash flows and ends up having lower values for its value function compared to firm A. This is even true when firm B owns a high level of installed base and firm A is absent (A's installed base is equal to 0). The reason for this is that the depreciation effect is strong enough to counteract the possibility of higher installed base levels, thus leading to low net discounted profits. We discuss the impact of the depreciation rate in a subsection below.

Figure 3 provides a general overview of the long-run market configurations for different values of α_A and α_B with $\alpha_A = \mu$ and $\alpha_B = \mu - \varepsilon$: it summarizes all market configurations for different combinations of μ and ε when the rate of depreciation is $\delta = 0.1$. The left panel corresponds to $\lambda = 1.2$ and the right panel to $\lambda = 1.7$. Each point (μ, ε) is associated with one entire probability distribution in the long run such as the one depicted in the right lower panel in Figure 1. Points on the vertical axis represent the cases where both firms are identical ($\alpha_A = \alpha_B$). Any point to the right of the vertical axis represents a case of heterogeneity in which firm B is the laggard ($\varepsilon > 0$). Since below the diagonal the difference $\mu - \varepsilon$ is negative, none of those points is associated with any model specification and they are left in blank. The farther to the right from the vertical axis, the higher the degree of heterogeneity in the likelihood of success of investment. The different colors represent each of the market structure types as indicated by the color bar on the right of the graphs.

Figure 3: Market structures when there are no externalities



Notes: The letter A represents that firm A becomes the monopolist. The letter D refers to duopoly. The term A, B means that the limiting distribution for installed base levels is bimodal, i.e., either firm may take over as a monopoly. Finally, the letter C indicates that the market collapses. Left panel represents the outcomes when $\lambda = 1.2$ and the right panel when $\lambda = 1.7$.

As investment becomes more reversible (higher depreciation rate δ) the region for duopoly shrinks from occupying a large portion of the parameter space studied to no presence at all.[25] As the price sensitivity λ increases the outside good market share expands in all cases. This

[25]Further results on the role of δ can be made available upon request.

effect dominates when μ is low (below 0.6 on the right panel) and the only market structure prevailing is both firms providing a good with an installed base of zero (market collapse). For values of μ greater than a certain value (0.6 on the right panel), positive installed base levels are observed in the long-run but each non-market collapse structure requires a higher value of μ to counteract the higher price sensitivity. The advantage of the leader firm increases in this case.

Finally, we discuss how an increase in heterogeneity (an increase in ε keeping μ constant) leads to changes in the long-run market structures. Figure 3 shows the effect of heterogeneity when a more capable firm (i.e., $\varepsilon > 0$), leads the market to change from a duopoly structure to a duopoly and monopoly of the leading firm, and to a monopoly of the leading firm only (for $\mu > 1.5$ on left panel and for $\mu > 2.4$ on the right panel). The effect of heterogeneity is even stronger when the price sensitivity is higher (right panel). Notice that if the leading firm has a relatively low capacity of transforming investment into base expansions (low α), then the effect of the heterogeneity is weaker and the probability of observing a monopoly from the laggard coexisting with positive probabilities for duopoly and monopoly from the leader firm is not negligible (μ approximately less than 1.5 on the left panel and less than 2.4 on the right panel). We summarize these insights in the following Observation.

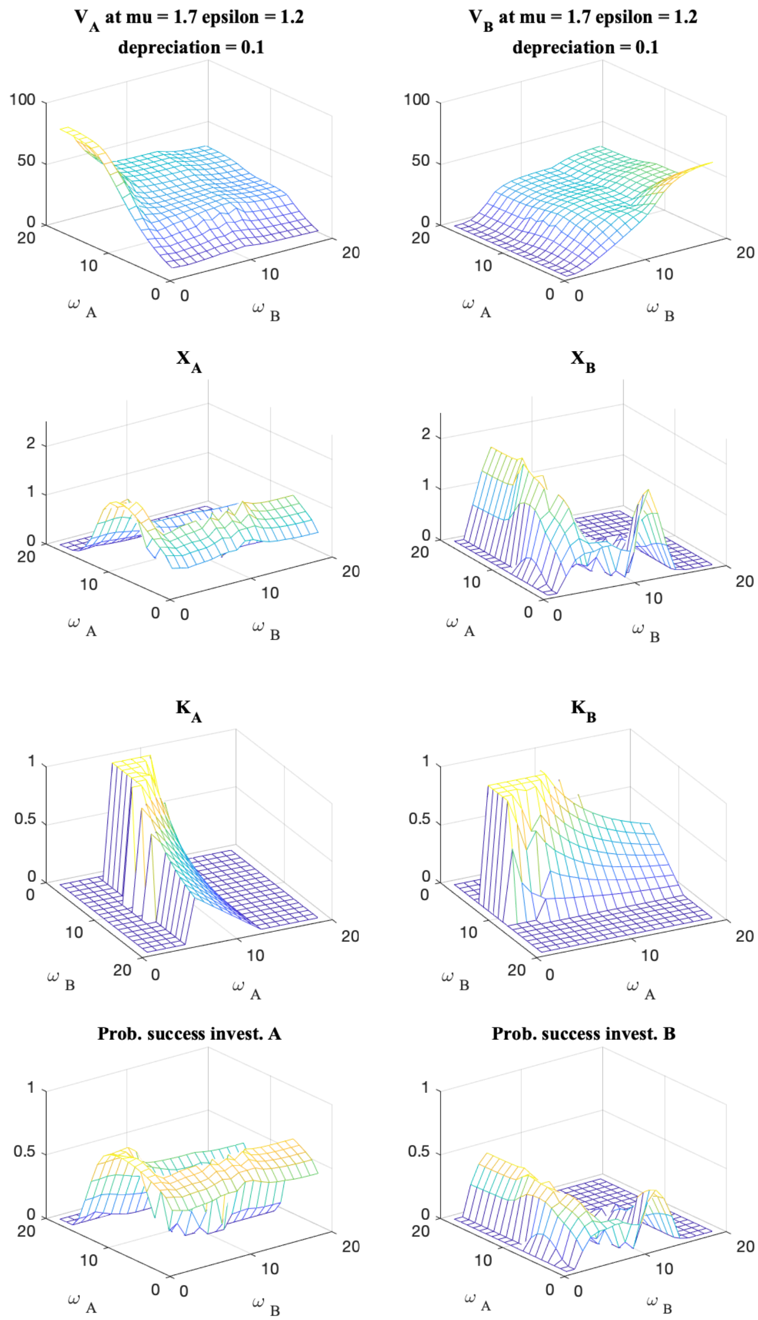
Observation 1: The model with heterogeneity in the likelihood of success of investment and externalities can exhibit different long-run distributions over market structures depending on parameter values. Those different structures are: market collapse, market dominance by either firm, duopoly, and combinations of these structures.

5.2 Market Structures and Externalities

Figure 4 shows the converged value, investment policy, compatibility, and probability of success functions in the case of externalities. If we compare them against the case of no externalities (Figure 2), we observe that although both firms benefit as indicated by the value function at states where the own-installed base levels are low, the larger relative gains are for the laggard firm. This is a direct consequence of the compatibility externality, which is symmetric for both firms.[26] The maximum value of the investment policy function shifts over lower values of the competitor's installed base because firms can rely also on the compatibility externality and not only on the investment to increase installed base levels through R&D.

[26]We discuss later in the paper the case where the laggard is the only firm that can absorb its competitor's installed base.

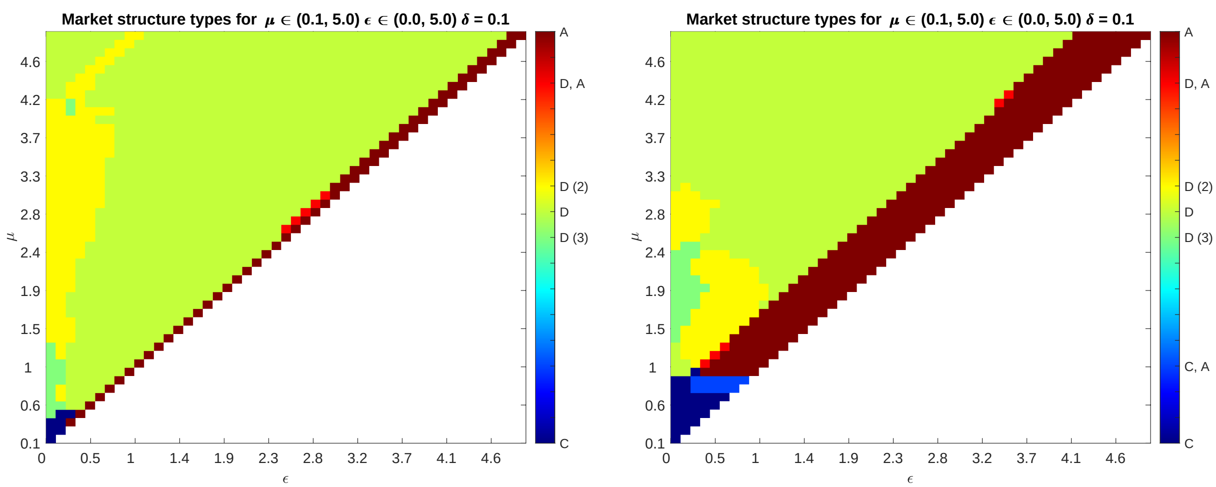
Figure 4: Value, investment, compatibility, and probability of success functions when there are externalities



Notes: Asymmetric R&D capabilities with $\alpha_A = 1.7$ and $\alpha_B = 0.5$. κ_A and κ_B are the functions describing the optimal amounts of the externalities at each state, note that the axes for these functions are reversed with respect to the rest of the graphs simply to ease their visualization and that these functions are the same regardless of the values of α_A and α_B. Here $\lambda = 1.7$ and $\delta = 0.1$.

Figure 5 shows the different combinations of market structures that arise from this model with externalities. The left panel shows the results when $\lambda = 1.2$ and the right panel when $\lambda = 1.7$. The presence of compatibility shrinks the region in which firm A dominates and makes the duopoly outcome more prevalent. In addition, there are new long run market configurations relative to the case with no externalities. For small levels of asymmetry in R&D capabilities we do not only observe a unimodal duopoly distribution, but for some parameter combinations we observe up to three modes, each one associated with positive amounts of installed base levels. This is one way in which the mutual adoption of the competitor's installed base creates fiercer competition leading to a higher frequency of market outcomes in which both firms can be present in the market in the long-run.

Figure 5: Market structures when there are externalities



Notes: The letter A represents that firm A becomes the monopolist. The letter D refers to duopoly. The term A, B means that the limiting distribution for installed base levels is bimodal, i.e., either firm may take over as a monopoly. Finally, the letter C indicates that the market collapses. $D(2)$ and $D(3)$ represent probability distributions in which there are two and three modes, respectively, each corresponding to a duopoly. Left panel represents the outcomes when $\lambda = 1.2$ and the right panel when $\lambda = 1.7$.

Although the presence of a market externality might be beneficial to the industry, this trade-off can be harmful to the leader if the competitors take advantage of this positive externality to a point where the lagging firm ends up dominating the market. This occurs if the benefit of the lagging firm from the leading firm's expansion is strong, for example when $\kappa_A = 0.3$ and $\kappa_B = 0.7$ for all states.[27] Increasing compatibility is detrimental for the leading firm as measured by the probability of observing $\omega_A > 0$ in the long run, which

[27]These results are available upon request.

leads to a total loss of market share by A as the lagging firm B becomes a duopolist for many more parameter combinations than before. We summarize this insight in the following Observation.

> **Observation 2:** Allowing for externalities, removes market dominance of the leading firm by allowing the lagging firm to benefit from the leading firm's investment. This leads to a higher frequency of market outcomes where both firms are present relative to the case without externalities.

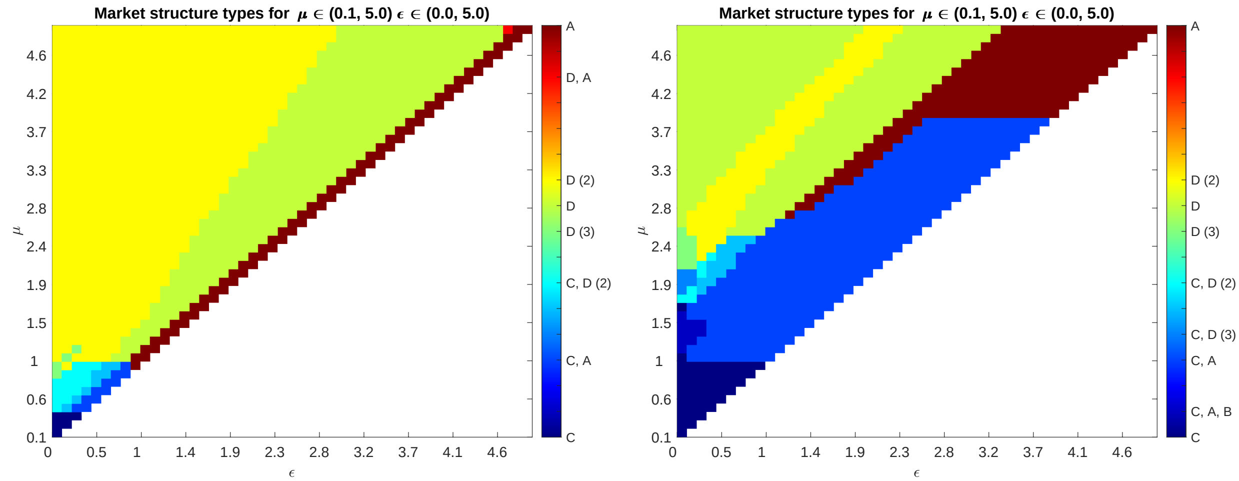
It is common when working with this type of models to report the expected market shares as opposed to the entire distribution over states: the sum of the element-by-element product of the matrix containing the long-run probability distribution and the matrix of market shares from the static game, by doing so, we mask the number of modes in the distribution by providing one single number that may confound different configurations. The analysis provided here allows us to distinguish the sources of the contributions towards the value of the weighted market share.

As we concluded in section 3, the curvature of the investment reaction functions depends on the sign of $\Delta_j - \Psi_j$. For all the cases analyzed here we find that the investments are strategic substitutes.

To end the analysis of symmetric externalities, we present in Figure 6 the map of market structures when the depreciation rate is not constant but equal to $1-(1-\delta)^{\omega_i}$ with $\delta = 0.1$.[28] For a fixed value of δ, this is an increasing function of the size of the installed base. As a consequence, it is harder to sustain high levels of ω in the long run and the possibility of long run distributions with two duopolies expands.

[28] This is the same functional form as in Chen et al. (2009).

Figure 6: Market structures when there are externalities and the depreciation is not constant



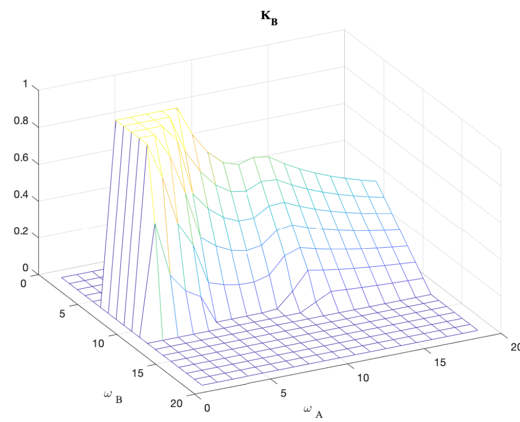
Notes: The letter A represents that firm A becomes the monopolist. The letter D refers to duopoly. The term A, B means that the limiting distribution for installed base levels is bimodal, i.e., either firm may take over as a monopoly. Finally, the letter C indicates that the market collapses. $D(2)$ and $D(3)$ represent probability distributions in which there are two and three modes, respectively, each corresponding to a duopoly. Left panel represents the outcomes when $\lambda = 1.2$ and the right panel when $\lambda = 1.7$. In both panels, the depreciation rate is $1 - (1 - \delta)^{\omega_i}$ with $\delta = 0.1$.

5.3 Asymmetric Absorption of the Competitor's Installed Base

Now consider the case where the firm that has a lower capability to transform investments into higher levels of installed base (the laggard) has the capability to make its good compatible with its competitor's. However, the other firm decides not to engage in attempting to absorb additional installed base through compatibility at any point in time. In symbols, $\kappa_A(\omega_A, \omega_B) = 0$ for all (ω_A, ω_B) and firm B solves Equation 9 together with the search for optimal prices. The resulting κ_B function is shown in Figure 7. κ_A is not shown since it is a constant function. Recall that those functions can be computed "off line".

Figure 8 shows the map of market structures when using those two different κ functions. By showing the different long-run market structures for different combinations of $\alpha_A = \mu$ and $\alpha_B = \mu - \varepsilon$ we examine the relationships between the leader (higher α) and the laggard ($\varepsilon > 0$) with the optimal compatibility strategy. Note that there are more structures where B can be present in the market than in the cases where both firms can adopt the competitor's technology. Moreover, there is a region where firm B, the laggard, is the only firm left in the market (approximately for $\mu \in [0.6, 0.7]$ and $\varepsilon < 0.2$) when the price sensitivity is low. This shows that it is possible that in equilibrium, a laggard invests in adopting its competitor's

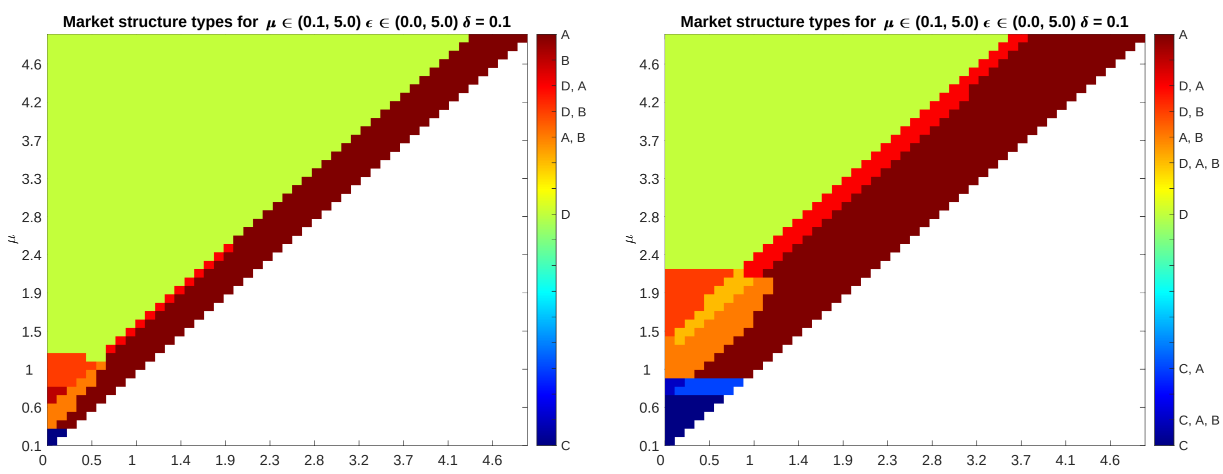
Figure 7: Function κ_B when $\kappa_A = 0$ for all states.



Notes: All the parameters values as in the baseline case and $\lambda = 1.7$. Note that the axes are reversed simply to ease the visualization.

installed base and overcompensates for its low ability to increase its own installed base. This special case is a stylized model of what a small car manufacturer would need to do to absorb certain patents from Toyota or other more advanced manufacturers, particularly affecting the perceived usability of their cars as discussed in section 2.

Figure 8: Market structures when only firm B can adopt its competitor's installed base



Notes: All the parameters values as in the baseline case. $\lambda = 1.2$ in the left panel and $\lambda = 1.7$ in the right panel. The lagging firm may become the monopolist (left lower corner of first panel).

5.4 The effects from ω^* and the depreciation rate δ

It has been documented in the literature the importance of the choice of the parameter ω^*.[29] Its role in the definition of the demand function in Equation 1 is that for installed base levels beyond ω^*, the utility function associated with this demand function becomes practically flat, which reduces the incentives for firms to innovate. Thus increasing the value of this parameter increases the region in the state space over which both firms have incentives to invest to increase installed base levels.

Equally important is the choice of the parameter δ, the probability of an industry-wide depreciation shock. An increase in this value makes it harder for firms to succeed in their investments. Since this parameter also measures the installed base level of the outside good relative to the inside goods, an increase in δ increases the incentive for firms to invest but up to a point where is no longer cost effective to keep doing so. Altogether, for a given value of ω^*, increases in δ will increase observing market structures in the long run with low installed base levels or even a market collapse. This can be seen in Figure 9 as measured by the integral of the market share using the transient distribution in the case of symmetric firms. The graph stresses the fact that ω^* belongs to a discrete support and uses an ad-hoc discretization for δ. As pointed throughout this paper, reporting outcomes of the game by taking the integral can mask the multiplicity of modes from the long-run probability distribution. Figure 12 in the Appendix shows how the number of modes associated with each choice of δ and ω^*.

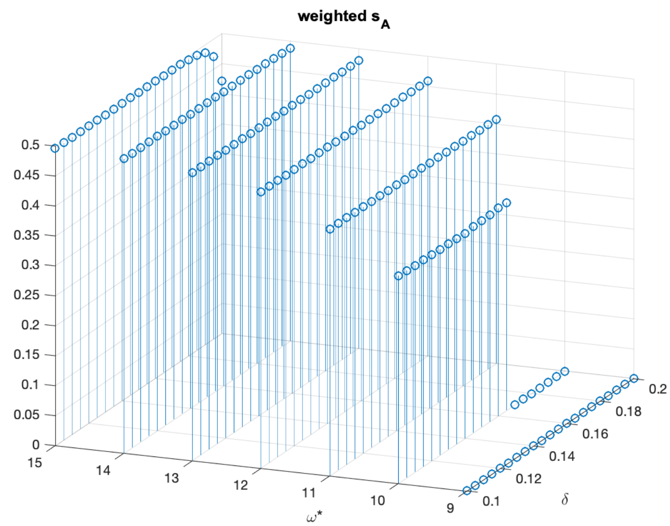
The different values of ω^* affect directly the static profits, the prices, and the κ function. The different values of δ only affect the dynamics of the game through the transition matrix. Figure 13 in the Appendix makes this evident by showing the value of the optimal level of the externality at $(\omega^* - 1, \omega^* - 1)$ for different values of ω^*. A greater value of ω^* requires a lower absorption level of the competitor's installed base.

6 Final Remarks

Although the distribution of long-run market structure outcomes has been documented in a few papers, we are the first to fully document the variety of these distributions that arise from a quality-ladder-type model with and without compatibility externalities. Pakes and McGuire (1994) report mean and standard deviations which acknowledge the multiplicity of modes in the long term distributions but no further mention of the characteristics of those

[29] See results in the Appendix from Goettler and Gordon (2014).

Figure 9: Market share for different values of δ and ω^*



Notes: Each point represents the value of s_A weighted by the long run probability distribution over market outcomes. For each combination of δ and ω^* shown we used $\lambda = 1.7$ and $\alpha_A = \alpha_B = 1.5$.

distributions. Goettler and Gordon (2014) wrote "although policymakers may be concerned with other moments of the outcome distribution, we follow the literature and focus on expected outcomes." Figure 4 in Borkovsky et al. (2012) and Figures 5-7 in Besanko et al. (2010) also suggest the multiplicity of modes in these long-run distributions. These few examples suggest that this richness of market outcomes has been underappreciated in the literature.

We agree with the previous mentions in the literature that additional moments of these distributions may be relevant for policymakers. For instance, a decision maker might contemplate whether to incentivize compatibility to implement a particular technology. With the long-term distribution of market outcomes in hand, the decision maker can rank the different equilibrium price distributions using stochastic dominance. This has a direct implication for consumer surplus. Along the same lines, the distribution on prices (and other market outcomes) has a direct connection with risk aversion: the decision maker can rank these distributions over prices if she has increasing and concave utility functions (second order stochastic dominance).

This paper does not attempt to assess the multiplicity of equilibria that Borkovsky et al. (2010) discuss. In the case of multiplicity, different policy functions and hence different transient distributions arise from the same parameter values (see for instance Table 7 in that paper). We abstract from such situations by focusing on a region of the parameter space where there has not been documented any multiplicity of equilibria. We are agnostic as to whether that phenomenon occurs in the case of externalities but our analysis of the investment reaction functions suggests there is a unique equilibrium.

As mentioned in the introduction, the estimation of the dynamic quality-ladder model is challenging. One important application is the one in Goettler and Gordon (2011). There, the policy experiments consist of simulating 10,000 industries under some specific counterfactual scenario, each industry is simulated 300 time periods given the initial condition given by the data. Then different outcomes are provided: expected profits, consumer surplus, and investments. It is unlikely but possible that the transient distribution over the state space exhibits multiple modes, which would indicate the possibility of a positive probability of different market structures similar to those in our paper.

Our main application of these insights is to study the effect of standard-setting. We model this as a compatibility externality in the presence of asymmetric returns to investment levels. Such externality is a function of the competitor's installed base and affects the consumer's utility for the other good. We examine the long-run distributions over market structures obtained by simulating the industry using the converged policy and value functions. We show that a single vector of model parameters can generate probability distributions that can lead to positive probabilities for one or more market structures. In particular, we show that it is possible for the laggard to keep its presence on the market in the long-run, in which case allowing for standard-setting practices may not be desirable by the leader. In the case where only the laggard invests in compatibility, it is possible for it to become a monopolist if the leader has a relatively low R&D capability and the two firms are almost symmetric in this same regard. Finally, the variety of multi-modal long-run distributions may have important consequences on the estimation and simulation of this type of models.

References

Abdelaziz, F. B., Brahim, M. B., and Zaccour, G. (2008). R&D Equilibrium Strategies with Surfers. *Journal of Optimization Theory and Applications*, 136(1):1–13.

Bekkers, R., Catalini, C., Martinelli, A., Righi, C., and Simcoe, T. (2017). Disclosure rules and declared essential patents. Technical report, National Bureau of Economic Research.

Besanko, D. and Doraszelski, U. (2004). Capacity Dynamics and Endogenous Asymmetries in Firm Size. *RAND Journal of Economics*, pages 23–49.

Besanko, D., Doraszelski, U., Kryukov, Y., and Satterthwaite, M. (2010). Learning-by-doing, Organizational Forgetting, and Industry Dynamics. *Econometrica*, 78(2):453–508.

Bloom, N., Schankerman, M., and Van Reenen, J. (2013). Identifying Technology Spillovers and Product Market Rivalry. *Econometrica*, 81(4):1347–1393.

Borkovsky, R. N., Doraszelski, U., and Kryukov, Y. (2010). A User's Guide to Solving Dynamic Stochastic Games Using the Homotopy Method. *Operations Research*, 58:1116–1132.

Borkovsky, R. N., Doraszelski, U., and Kryukov, Y. (2012). A Dynamic Quality Ladder Model with Entry and Exit: Exploring the Equilibrium Correspondence Using the Homotopy Method. *Quantitative Marketing and Economics*, 10(2):197–229.

Caplin, A. and Nalebuff, B. (1991). Aggregation and Imperfect Competition: On the Existence of Equilibrium. *Econometrica*, pages 25–59.

Cellini, R. and Lambertini, L. (2009). Dynamic R&D with Spillovers: Competition vs Cooperation. *Journal of Economic Dynamics and Control*, 33(3):568–582.

Chamley, C. and Gale, D. (1994). Information revelation and strategic delay in a model of investment. *Econometrica (1986-1998)*, 62(5):1065.

Chen, J., Doraszelski, U., and Harrington Jr, J. E. (2009). Avoiding Market Dominance: Product Compatibility in Markets with Network Effects. *The RAND Journal of Economics*, 40(3):455–485.

Chiao, B., Lerner, J., and Tirole, J. (2007). The Rules of Standard-setting Organizations: an Empirical Analysis. *The RAND Journal of Economics*, 38(4):905–930.

D'Aspremont, C. and Jacquemin, A. (1988). Cooperative and Non cooperative R&D in Duopoly with Spillovers. *American Economic Review*, pages 1133–1137.

Ericson, R. and Pakes, A. (1995). Markov-Perfect Industry Dynamics: A Framework for Empirical Work. *The Review of Economic Studies*, 62(1):53–82.

Ganglmair, B. and Tarantino, E. (2014). Conversation with secrets. *The RAND Journal of Economics*, 45(2):273–302.

Goettler, R. L. and Gordon, B. R. (2011). Does AMD Spur Intel to Innovate More? *Journal of Political Economy*, 119(6):1141–1200.

Goettler, R. L. and Gordon, B. R. (2014). Competition and Product Innovation in Dynamic Oligopoly. *Quantitative Marketing and Economics*, 12(1):1–42.

Gowrisankaran, G. and Town, R. J. (1997). Dynamic Equilibrium in the Hospital Industry. *Journal of Economics & Management Strategy*, 6(1):45–74.

Igami, M. and Yang, N. (2016). Unobserved Heterogeneity in Dynamic Games: Cannibalization and preemptive entry of hamburger chains in Canada. *Quantitative Economics*, 7(2):483–521.

Kamien, M. I., Muller, E., and Zang, I. (1992). Research Joint Ventures and R&D Cartels. *American Economic Review*, pages 1293–1306.

Lerner, J., Tabakovic, H., and Tirole, J. (2016). Patent disclosures and standard-setting. Technical report, National Bureau of Economic Research.

Levhari, D. and Mirman, L. J. (1980). The Great Fish War: an Example Using a Dynamic Cournot-Nash Solution. *The Bell Journal of Economics*, pages 322–334.

Pakes, A. and McGuire, P. (1994). Computing Markov-perfect Nash equilibria: Numerical implications of a dynamic differentiated product model. *The Rand Journal of Economics*, 25(4):555.

Rysman, M. and Simcoe, T. (2008). Patents and the Performance of Voluntary Standard-setting Organizations. *Management science*, 54(11):1920–1934.

Samano, M., Santugini, M., and Zaccour, G. (2017). Dynamics in Research Joint Ventures and R&D Collaborations. *Journal of Economic Dynamics and Control*, 77:70–92.

Song, M. (2011). A Dynamic Analysis of Cooperative Research in the Semiconductor Industry. *International Economic Review*, 52(4):1157–1177.

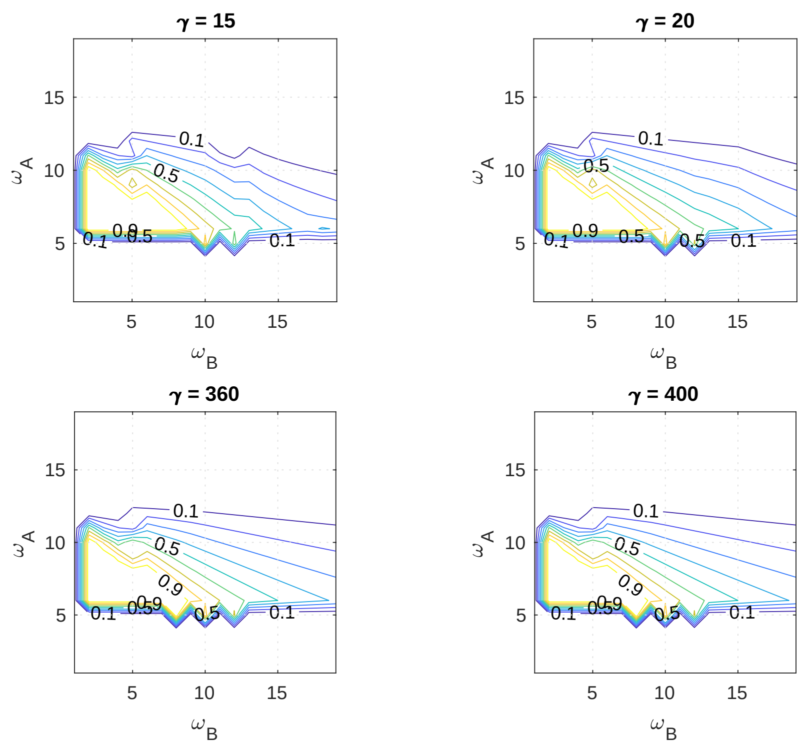
Watzinger, M., Fackler, T. A., Nagler, M., and Schnitzer, M. (2017). How Antitrust Enforcement Can Spur Innovation: Bell Labs and the 1956 Consent Decree. *Working paper.*

Appendix

The role of γ

We show how the function $\kappa_{.}$ changes as we change the parameter γ. First for our main specification for λ, 1.7, and then for a lower sensitivity to price changes, $\lambda = 1.2$. We are aware that the $\kappa_{.}$ function has a discrete support (the grid of states) but to facilitate the comparison across the different cases we opt for a set of contour plots that shows interpolations of the original function. As Figure 10 shows, the shape of the κ function remains mostly unchanged as we change the value of γ. Recall that our baseline value is $\gamma = 360$. This gives us confidence about the robustness of our main results with respect to this parameter.

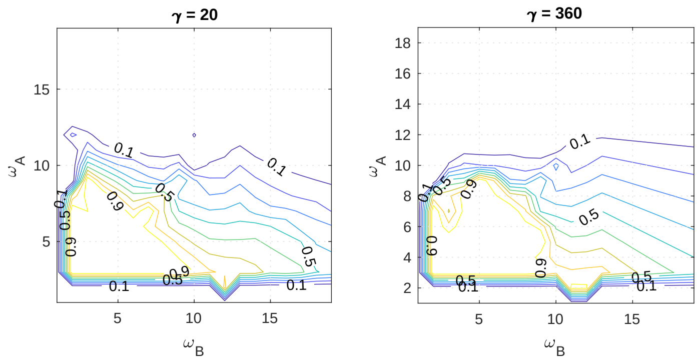
Figure 10: Contour plots for κ_A for different values of the parameter γ



Notes: All the parameters values as in the baseline case and $\lambda = 1.7$. The graph for κ_B is the transpose of that for κ_A.

However, there is an interaction with λ. Figure 10 shows the shape of the function κ but now using a lower sensitivity to price, $\lambda = 1.2$. The graph has still the same shape but its non-zero values are shifted all towards lower values of ω_A. This is the case because at each state, $\kappa_{.}$ is the solution of a system of four equations: two for the externalities and two for the prices, and the latter are functions that depend on the value of λ. By lowering the sensitivity to the price, prices are higher and it becomes optimal to absorb more of the competitor's quality level. We obtain convergence for the solution of the system of equations for all the values of γ tested except when $\gamma < 10$. Recall that the κ function is computed "off line" and therefore, the parameters α play no role in its computation.

Figure 11: Contour plots for κ_A for different values of the parameter γ

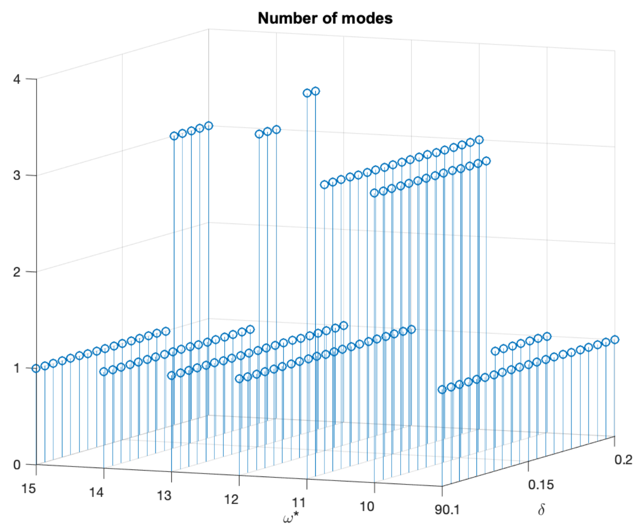


Notes: All the parameter values as in the baseline case and $\lambda = 1.2$. The graph for κ_B is the transpose of that for κ_A.

Interaction between δ and ω^*

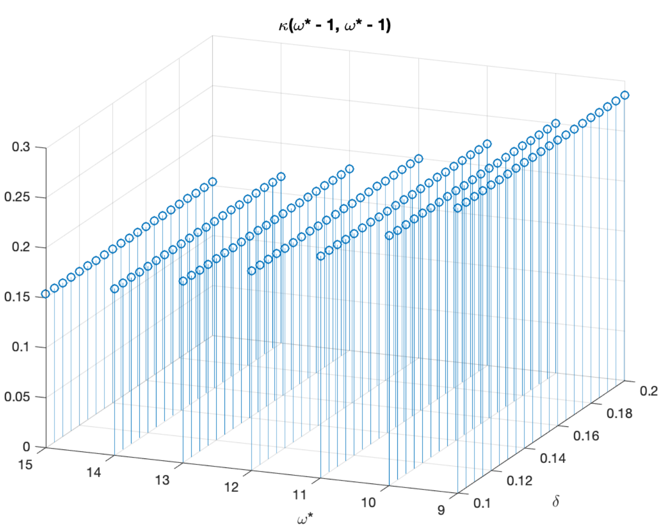
The following graphs complement the analysis from subsection 5.4.

Figure 12: Number of different market structures for different values of δ and ω^*



Notes: Each point represents the number of modes in the long-run probability distribution of market outcomes. For each combination of δ and ω^* shown we used $\lambda = 1.7$ and $\alpha_A = \alpha_B = 1.5$.

Figure 13: Strength of the optimal externality for different values of δ and ω^*



Notes: Each point represents the value of the κ function at $(\omega^* - 1, \omega^* - 1)$. Note that the values are constant for a given value of ω^* regardless of the value of δ. For each combination of δ and ω^* shown we used $\lambda = 1.7$ and $\alpha_A = \alpha_B = 1.5$.

Online Appendix (not for publication)

In this appendix, we describe the Levhari-Mirman (LM) (1980) algorithm and the transition matrix. To ease the notation we suppress the arguments κ_j and κ_{3-j}.

LM Algorithm

Value Function, Finite Programs. For $j = 1, 2$, consider firm j's maximization problem for a horizon of t periods, $t = 0, 1, ...$ For $j = 1, 2$, given $x_{3-j} \geq 0$, firm j's value function for a t-period horizon is

$$V_j^t(\omega_j, \omega_{3-j}) = \max_{x_j \geq 0} \left\{ \Pi_j(\omega_j, \omega_{3-j}) - x_j + \beta_j \mathbf{E}[V_j^{t-1}(\omega'_j, \omega'_{3-j}) | \omega_j, \omega_{3-j}, x_j, x_{3-j}] \right\} \tag{11}$$

where $\mathbf{E}[\cdot]$ is the expectation operator with respect to $\{\omega'_j, \omega'_{3-j}\}$ according to the transition probabilities. The value function for the static game (i.e., when $t = 0$) is

$$V_j^0(\omega_j, \omega_{3-j}) = \max_{x_j \geq 0} \{\Pi_j(\omega_j, \omega_{3-j}) - x_j\}. \tag{12}$$

Consistent with Equation 11, firm j's value function for the infinite-period horizon is thus

$$V_j^\infty(\omega_j, \omega_{3-j}) = \max_{x_j \geq 0} \left\{ \Pi_j(\omega_j, \omega_{3-j}) - x_j + \beta_j \mathbf{E}[V_j^\infty(\omega'_j, \omega'_{3-j}) | \omega_j, \omega_{3-j}, x_j, x_{3-j}] \right\}. \tag{13}$$

Equilibrium. Next, we define the Markov-perfect equilibrium for a game lasting $T + 1$ periods, i.e., a horizon of T periods, $T = 0, 1, ..., \infty$. The equilibrium consists of the strategies of the two firms for every horizon from the first period (when there are T periods left) to the last period (when there is no horizon).

Definition The tuple $\{X_1^t(\omega_1, \omega_2), X_2^t(\omega_2, \omega_1)\}_{t=0}^T$ is a Markov-perfect Nash equilibrium for a game of T-period horizons if, for all $\{\omega_1, \omega_2\}$,

1. For $t = 0$, for $j = 1, 2$, given $X_{3-j}^0(\omega_{3-j}, \omega_j)$,

$$X_j^0(\omega_{3-j}, \omega_j) = \arg\max_{x_j \geq 0} \{\Pi_j(\omega_j, \omega_{3-j}) - x_j\}. \tag{14}$$

2. For $t = 1, 2, \ldots, T$, for $j = 1, 2$, given $X^t_{3-j}(\omega_{3-j}, \omega_j)$ and $\{X^t_1(\omega_1, \omega_2), X^t_2(\omega_2, \omega_1)\}^{t-1}_{t=0}$,

$$
\begin{aligned}
&X^t_j(\omega_{3-j}, \omega_j) \\
&= \arg\max_{x_j \geq 0} \{\Pi_j(\omega_j, \omega_{3-j}) - x_j \\
&+ \beta_j \phi_j(x_j)\phi_{3-j}(X^t_{3-j}(\omega_{3-j}, \omega_j)) \cdot \left(\delta V^{t-1}_j(\omega_j, \omega_{3-j}) + (1-\delta)V^{t-1}_j(\omega_j + 1, \omega_{3-j} + 1)\right) \\
&+ \beta_j \phi_j(x_j)\left(1 - \phi_{3-j}\left(X^t_{3-j}(\omega_{3-j}, \omega_j)\right)\right) \cdot \left(\delta V^{t-1}_j(\omega_j, \omega_{3-j} - 1) + (1-\delta)V^{t-1}_j(\omega_j + 1, \omega_{3-j})\right) \\
&+ \beta_j(1 - \phi_j(x_j))\phi_{3-j}\left(X^t_{3-j}(\omega_{3-j}, \omega_j)\right) \cdot \left(\delta V^{t-1}_j(\omega_j - 1, \omega_{3-j}) + (1-\delta)V^{t-1}_j(\omega_j, \omega_{3-j} + 1)\right) \\
&+ \beta_j(1 - \phi_j(x_j))\left(1 - \phi_{3-j}\left(X^t_{3-j}(\omega_{3-j}, \omega_j)\right)\right) \cdot \left(\delta V^{t-1}_j(\omega_j - 1, \omega_{3-j} - 1) + (1-\delta)V^{t-1}_j(\omega_j, \omega_{3-j})\right)\}
\end{aligned} \tag{15}
$$

where, for any $y, z \in \{1, 2, ..., M\}$,

$$
V^{t'-1}_j(y, z) = \begin{cases} \Pi_j(y, z) - X^0_j(y, z) & t' = 1 \\ \Pi_j(y, z) - X^{t'-1}_j(y, z) + \beta_j \cdot \Gamma^{t'-2}_j(X^{t'-1}_j(y, z), X^{t'-1}_{3-j}(z, y)) & t' = 2, 3, ..., T \end{cases} \tag{16}
$$

is the value function for a $t' - 1$ period horizon for any state vector (y, z) with

$$
\begin{aligned}
&\Gamma^{t'-2}_j(X^{t'-1}_j(y, z), X^{t'-1}_{3-j}(z, y)) \\
&= \phi_j(X^{t'-1}_j(y, z))\phi_{3-j}(X^{t'-1}_{3-j}(z, y) \cdot \left(\delta V^{t'-2}_j(y, z) + (1-\delta)V^{t'-2}_j(y + 1, z + 1)\right) \\
&+ \phi_j(X^{t'-1}_j(y, z))(1 - \phi_{3-j}(X^{t'-1}_{3-j}(z, y))) \cdot \left(\delta V^{t'-2}_j(y, z - 1) + (1-\delta)V^{t'-2}_j(y + 1, z)\right) \\
&+ (1 - \phi_j(X^{t'-1}_j(y, z)))\phi_{3-j}(X^{t'-1}_{3-j}(z, y)) \cdot \left(\delta V^{t'-2}_j(y - 1, z) + (1-\delta)V^{t'-2}_j(y, z + 1)\right) \\
&+ (1 - \phi_j(X^{t'-1}_j(y, z)))(1 - \phi_{3-j}(X^{t'-1}_{3-j}(z, y)) \cdot \left(\delta V^{t'-2}_j(y - 1, z - 1) + (1-\delta)V^{t'-2}_j(y, z)\right)
\end{aligned} \tag{17}
$$

is the expected continuation value function corresponding to the equilibrium for a horizon of $t' - 2$ periods.

Condition 1 defines the Nash equilibrium in the static game. Note that in fact, there is no externality since $X^0_{3-j}(\omega_{3-j}, \omega_j)$ has no effect on the zero-period-horizon objective function for firm j. Condition 2 states the equilibrium for every higher horizon of the game. For $t = 1, 2, 3, ..., T$, Equation 16 and Equation 17 reflect the recursive nature of the equilibrium in which the equilibrium continuation value function for a $(t-1)$-period horizon depends on the equilibrium strategies for t'-period horizons, $(t - 1) > t' \geq 0$.

Proposition states the Markov-perfect Nash equilibrium for each horizon of the game.

Proposition 6.1. *Consider a game of T-period horizons.*

1. *For* $t = 0$,
$$\left\{X_1^0\left(\omega_1, \omega_2\right), X_2^0\left(\omega_2, \omega_1\right)\right\} = \{0, 0\},$$
with the corresponding value function is
$$V_j^0\left(\omega_j, \omega_{3-j}\right) = \Pi_j\left(\omega_j, \omega_{3-j}\right).$$

2. *For* $t \geq 1$, *given* $\{V_1^{t-1}\left(\omega_1, \omega_2\right), V_2^{t-1}(\omega_2, \omega_1)$, $\{X_1^t\left(\omega_1, \omega_2\right), X_2^t\left(\omega_1, \omega_2\right)\}$ *is defined by*
$$X_1^t\left(\omega_1, \omega_2\right) = \max\left\{-\frac{1}{\alpha_1} + \sqrt{\frac{\beta_1}{\alpha_1}}\sqrt{\frac{\alpha_2 X_2^t\left(\omega_2, \omega_1\right)\Delta_1^{t-1} + \Psi_1^{t-1}}{1 + \alpha_2 X_2^t\left(\omega_2, \omega_1\right)}}, 0\right\}, \tag{18}$$
$$X_2^t\left(\omega_2, \omega_1\right) = \max\left\{-\frac{1}{\alpha_2} + \sqrt{\frac{\beta_2}{\alpha_2}}\sqrt{\frac{\alpha_1 X_1^t\left(\omega_1, \omega_2\right)\Delta_2^{t-1} + \Psi_2^{t-1}}{1 + \alpha_1 X_1^t\left(\omega_1, \omega_2\right)}}, 0\right\}, \tag{19}$$
where for $j = 1, 2$,
$$\begin{aligned}\Delta_j^{t-1} &\equiv \delta\left[V_j^{t-1}\left(\omega_j, \omega_{3-j}\right) - V_j^{t-1}\left(\omega_j - 1, \omega_{3-j}\right)\right] \\ &\quad + (1-\delta)\left[V_j^{t-1}\left(\omega_j + 1, \omega_{3-j} + 1\right) - V_j^{t-1}\left(\omega_j, \omega_{3-j} + 1\right)\right], \\ \Psi_j^{t-1} &\equiv \delta\left[V_j^{t-1}\left(\omega_j, \omega_{3-j} - 1\right) - V_j^{t-1}\left(\omega_j - 1, \omega_{3-j} - 1\right)\right] \\ &\quad + (1-\delta)\left[V_j^{t-1}\left(\omega_j + 1, \omega_{3-j}\right) - V_j^{t-1}\left(\omega_j, \omega_{3-j}\right)\right].\end{aligned}$$

Proof The first-order condition corresponding to Equation 15 is

$$\begin{aligned}&-d_j + \beta_j\frac{\alpha_j}{\left(1+\alpha_j x_j\right)^2}\phi_{3-j}(X_{3-j}^t(\omega_{3-j}, \omega_j)) \cdot \left(\delta V_j^{t-1}\left(\omega_j, \omega_{3-j}\right) + (1-\delta)V_j^{t-1}\left(\omega_j + 1, \omega_{3-j} + 1\right)\right) \\ &+ \beta_j\frac{\alpha_j}{\left(1+\alpha_j x_j\right)^2}\left(1 - \phi_{3-j}\left(X_{3-j}^t(\omega_{3-j}, \omega_j)\right)\right) \cdot \left(\delta V_j^{t-1}\left(\omega_j, \omega_{3-j} - 1\right) + (1-\delta)V_j^{t-1}\left(\omega_j + 1, \omega_{3-j}\right)\right) \\ &- \beta_j\frac{\alpha_j}{\left(1+\alpha_j x_j\right)^2}\phi_{3-j}\left(X_{3-j}^t(\omega_{3-j}, \omega_j)\right) \cdot \left(\delta V_j^{t-1}\left(\omega_j - 1, \omega_{3-j}\right) + (1-\delta)V_j^{t-1}\left(\omega_j, \omega_{3-j} + 1\right)\right) \\ &- \beta_j\frac{\alpha_j}{\left(1+\alpha_j x_j\right)^2}\left(1 - \phi_{3-j}\left(X_{3-j}^t(\omega_{3-j}, \omega_j)\right)\right) \cdot \left(\delta V_j^{t-1}\left(\omega_j - 1, \omega_{3-j} - 1\right) + (1-\delta)V_j^{t-1}\left(\omega_j, \omega_{3-j}\right)\right) \\ &= 0\end{aligned}$$

which yields Equation 8 and thus Equation 19, as long as the second-order condition is satisfied, i.e., for $j, 3-j = 1, 2, j \neq 3-j$,

$$
\begin{aligned}
&-\beta_j \frac{2\alpha_j^2}{(1+\alpha_j x_j)^3} \frac{\alpha_{3-j} x_{3-j}}{1+\alpha_{3-j} x_{3-j}} \cdot \left(\delta V_j^{t-1}(\omega_j, \omega_{3-j}) + (1-\delta) V_j^{t-1}(\omega_j + 1, \omega_{3-j} + 1)\right) \\
&-\beta_j \frac{2\alpha_j^2}{(1+\alpha_j x_j)^3} \frac{1}{1+\alpha_{3-j} x_{3-j}} \cdot \left(\delta V_j^{t-1}(\omega_j, \omega_{3-j} - 1) + (1-\delta) V_j^{t-1}(\omega_j + 1, \omega_{3-j})\right) \\
&+\beta_j \frac{2\alpha_j^2}{(1+\alpha_j x_j)^3} \frac{\alpha_{3-j} x_{3-j}}{1+\alpha_{3-j} x_{3-j}} \cdot \left(\delta V_j^{t-1}(\omega_j - 1, \omega_{3-j}) + (1-\delta) V_j^{t-1}(\omega_j, \omega_{3-j} + 1)\right) \\
&+\beta_j \frac{2\alpha_j^2}{(1+\alpha_j x_j)^3} \frac{1}{1+\alpha_{3-j} x_{3-j}} \cdot \left(\delta V_j^{t-1}(\omega_j - 1, \omega_{3-j} - 1) + (1-\delta) V_j^{t-1}(\omega_j, \omega_{3-j})\right) < 0
\end{aligned}
$$

Algorithm. Having described the model and defined the equilibrium. We now proceed with the characterization of the MPE. Here, we solve the equilibrium recursively as in Levhari and Mirman (1980). Consider first the static game of investment, i.e., $t = 0$. Then, there is no externality, and no firm has an incentive to invest, i.e., the Markov-perfect equilibrium for a game of 0-periods horizon is simply

$$\left\{X_1^1(\omega_1, \omega_2), X_2^1(\omega_1, \omega_2)\right\} = \{0, 0\},$$

with the corresponding value function

$$V_j^0(\omega_j, \omega_{3-j}) = \Pi_j(\omega_j, \omega_{3-j}).$$

Hence, there is a unique equilibrium for the no-horizon game in which the firms do not invest and the value function is equal to the profit function corresponding to the Bertrand game.

Consistent with the solution of the equilibrium, we characterize the equilibrium for each horizon. Each iteration is an horizon with the caveat that at each iteration, the solution to the reaction function is a Markov-perfect Nash equilibrium (and not an approximation). Hence, wherever we stop, we have an equilibrium. The question remains whether we converge to the stationary Markov-perfect Nash equilibrium (in infinite horizons).

1. For $t = 0$,

   $$\left\{X_1^0(\omega_1, \omega_2), X_2^0(\omega_2, \omega_1)\right\} = \{0, 0\},$$

   with the corresponding value function is

   $$V_j^0(\omega_j, \omega_{3-j}) = \Pi_j(\omega_j, \omega_{3-j}).$$

2. For $t \geq 1$, given $\{V_1^{t-1}(\omega_1, \omega_2), V_2^{t-1}(\omega_2, \omega_1)\}$, firm j's reaction functions are given by Equation 19 and the solution to this system of equations is the equilibrium at this iteration.

Transition Probability Matrix

Using the converged policy functions, for $j = 1, 2$,

$$\omega_j' | \omega_j = \min\{\max\{\omega_j + \tau_j + \eta, 1\}, M\}$$

where $\tau_j \in \{1, 0\}$ such that $\Pr[\tau_j = 1] = \phi_j(\omega_1, \omega_2) = \frac{\alpha_j X_j(\omega_j, \omega_{3-j})}{1+\alpha_j X_j(\omega_j, \omega_{3-j})}$ and $\eta \in \{-1, 0\}$ such that $\Pr[\eta = -1] = \delta$.

We want to calculate all transition probabilities such as $\Pr[(\omega_1', \omega_2') \,|\, (\omega_1, \omega_2)]$. We consider each case separately.

1. Suppose that (ω_1, ω_2) is such that $\omega_1, \omega_2 \notin \{0, M\}$. Given (ω_1, ω_2), there are $(M + 1)^2$ conditional probabilities to calculate. All of them are zero except

$$\begin{aligned}
\Pr[(\omega_1, \omega_2) \,|\, (\omega_1, \omega_2)] &= \delta\phi_1(\omega_1, \omega_2)\,\phi_2(\omega_2, \omega_1) \\
&\quad + (1-\delta)(1-\phi_1(\omega_1, \omega_2))(1-\phi_2(\omega_2, \omega_1)), \\
\Pr[(\omega_1 + 1, \omega_2) \,|\, (\omega_1, \omega_2)] &= (1-\delta)\phi_1(\omega_1, \omega_2)(1-\phi_2(\omega_2, \omega_1)), \\
\Pr[(\omega_1 - 1, \omega_2) \,|\, (\omega_1, \omega_2)] &= \delta(1-\phi_1(\omega_1, \omega_2))\,\phi_2(\omega_2, \omega_1), \\
\Pr[(\omega_1, \omega_2 - 1) \,|\, (\omega_1, \omega_2)] &= \delta\phi_1(\omega_1, \omega_2)(1-\phi_2(\omega_2, \omega_1)), \\
\Pr[(\omega_1 - 1, \omega_2 - 1) \,|\, (\omega_1, \omega_2)] &= \delta(1-\phi_1(\omega_1, \omega_2))(1-\phi_2(\omega_2, \omega_1)), \\
\Pr[(\omega_1, \omega_2 + 1) \,|\, (\omega_1, \omega_2)] &= (1-\delta)(1-\phi_1(\omega_1, \omega_2))\,\phi_2(\omega_2, \omega_1), \\
\Pr[(\omega_1 + 1, \omega_2 + 1) \,|\, (\omega_1, \omega_2)] &= (1-\delta)\,\phi_1(\omega_1, \omega_2)\phi_2(\omega_2, \omega_1).
\end{aligned}$$

2. Suppose that $(\omega_1, \omega_2) = (0, 0)$. Given (ω_1, ω_2), there are $(M + 1)^2$ conditional probabilities to calculate. All of them are zero except

$$\begin{aligned}
\Pr[(0,0) \,|(0,0)] &= 1 - (1-\delta)(\phi_1(0,0) + \phi_2(0,0) - \phi_1(0,0)\phi_2(0,0)), \\
\Pr[(1,0) \,|(0,0)] &= (1-\delta)\phi_1(0,0)(1-\phi_2(0,0)), \\
\Pr[(0,1) \,|(0,0)] &= (1-\delta)(1-\phi_1(0,0))\,\phi_2(0,0), \\
\Pr[(1,1) \,|(0,0)] &= (1-\delta)\,\phi_1(0,0)\phi_2(0,0).
\end{aligned}$$

3. Suppose that $(\omega_1, \omega_2) = (M, M)$. Given (ω_1, ω_2), there are $(M + 1)^2$ conditional prob-

abilities to calculate. All of them are zero except

$$\begin{aligned}
\Pr[(M,M)\,|\,(M,M)] &= 1-\delta(1-\phi_1(M,M)\,\phi_2(M,M)),\\
\Pr[(M-1,M)\,|\,(M,M)] &= \delta\,(1-\phi_1(M,M))\,\phi_2(M,M)\,,\\
\Pr[(M,M-1)\,|\,(M,M)] &= \delta\phi_1(M,M)\,(1-\phi_2(M,M))\,,\\
\Pr[(M-1,M-1)\,|\,(M,M)] &= \delta\,(1-\phi_1(M,M))\,(1-\phi_2(M,M))\,.
\end{aligned}$$

4. Suppose that $(\omega_1,\omega_2)=(0,M)$. Given (ω_1,ω_2), there are $(M+1)^2$ conditional probabilities to calculate. All of them are zero except

$$\begin{aligned}
\Pr[(0,M)\,|\,(0,M)] &= 1-(1-\delta)\phi_1(0,M)-\delta\,(1-\phi_2(0,M))\,,\\
\Pr[(1,M)\,|\,(0,M)] &= (1-\delta)\phi_1(0,M)\,,\\
\Pr[(0,M-1)\,|\,(0,M)] &= \delta\,(1-\phi_2(0,M))\,.
\end{aligned}$$

5. Suppose that $(\omega_1,\omega_2)=(M,0)$. Given (ω_1,ω_2), there are $(M+1)^2$ conditional probabilities to calculate. All of them are zero except

$$\begin{aligned}
\Pr[(M,0)\,|\,(M,0)] &= 1-(1-\delta)\phi_2(0,M)-\delta\,(1-\phi_1(M,0))\,,\\
\Pr[(M,1)\,|\,(M,0)] &= (1-\delta)\phi_2(0,M)\,,\\
\Pr[(M-1,0)\,|\,(M,0)] &= \delta\,(1-\phi_1(M,0))\,.
\end{aligned}$$

6. Suppose that (ω_1,ω_2) is such that $\omega_1=0$ and $\omega_2\notin\{0,M\}$. Given (ω_1,ω_2), there are $(M+1)^2$ conditional probabilities to calculate. All of them are zero except

$$\begin{aligned}
\Pr[(0,\omega_2)\,|\,(0,\omega_2)] &= \delta\phi_2(\omega_2,0)\\
&\qquad+(1-\delta)\,(1-\phi_1(0,\omega_2))\,(1-\phi_2(\omega_2,0))\,,\\
\Pr[(1,\omega_2)\,|\,(0,\omega_2)] &= (1-\delta)\phi_1(0,\omega_2)\,(1-\phi_2(\omega_2,0))\,,\\
\Pr[(0,\omega_2-1)\,|\,(0,\omega_2)] &= \delta\,(1-\phi_2(\omega_2,0))\,,\\
\Pr[(0,\omega_2+1)\,|\,(0,\omega_2)] &= (1-\delta)\,(1-\phi_1(0,\omega_2))\,\phi_2(\omega_2,0)\,,\\
\Pr[(1,\omega_2+1)\,|\,(0,\omega_2)] &= (1-\delta)\,\phi_1(0,\omega_2)\,\phi_2(\omega_2,0)\,.
\end{aligned}$$

7. Suppose that (ω_1,ω_2) is such that $\omega_1\notin\{0,M\}$ and $\omega_2=0$. Given (ω_1,ω_2), there are

$(M+1)^2$ conditional probabilities to calculate. All of them are zero except

$$\begin{aligned}
\Pr[(\omega_1, 0) \,|\, (\omega_1, 0)] &= \delta\phi_1(\omega_1, 0) \\
&\quad + (1-\delta)(1-\phi_2(0,\omega_1))(1-\phi_1(\omega_1,0)), \\
\Pr[(\omega_1, 1) \,|\, (\omega_1, 0)] &= (1-\delta)\phi_2(0,\omega_1)(1-\phi_1(\omega_1,0)), \\
\Pr[(\omega_1 - 1, 0) \,|\, (\omega_1, 0)] &= \delta(1-\phi_1(\omega_1,0)), \\
\Pr[(\omega_1 + 1, 0) \,|\, (\omega_1, 0)] &= (1-\delta)(1-\phi_2(0,\omega_1))\phi_1(\omega_1,0), \\
\Pr[(\omega_1 + 1, 1) \,|\, (\omega_1, 0)] &= (1-\delta)\phi_2(0,\omega_1)\phi_1(\omega_1,0).
\end{aligned}$$

8. Suppose that (ω_1, ω_2) is such that $\omega_1 = M$ and $\omega_2 \notin \{0, M\}$. Given (ω_1, ω_2), there are $(M+1)^2$ conditional probabilities to calculate. All of them are zero except

$$\begin{aligned}
\Pr[(M, \omega_2) \,|\, (M, \omega_2)] &= \delta\phi_1(M,\omega_2)\phi_2(\omega_2, M) \\
&\quad + (1-\delta)(1-\phi_2(\omega_2, M)) \\
\Pr[(M-1, \omega_2) \,|\, (M, \omega_2)] &= \delta(1-\phi_1(M,\omega_2))\phi_2(\omega_2, M) \\
\Pr[(M, \omega_2 - 1) \,|\, (M, \omega_2)] &= \delta\phi_1(M,\omega_2)(1-\phi_2(\omega_2, M)) \\
\Pr[(M-1, \omega_2 - 1) \,|\, (M, \omega_2)] &= \delta(1-\phi_1(M,\omega_2))(1-\phi_2(\omega_2, M)) \\
\Pr[(M, \omega_2 + 1) \,|\, (M, \omega_2)] &= (1-\delta)\phi_2(\omega_2, M)
\end{aligned}$$

9. Suppose that (ω_1, ω_2) is such that $\omega_1 \notin \{0, M\}$ and $\omega_2 = M$. Given (ω_1, ω_2), there are $(M+1)^2$ conditional probabilities to calculate. All of them are zero except

$$\begin{aligned}
\Pr[(M, \omega_1) \,|\, (\omega_1, M)] &= \delta\phi_2(M,\omega_1)\phi_1(\omega_1, M) \\
&\quad + (1-\delta)(1-\phi_1(\omega_1, M)) \\
\Pr[(\omega_1, M-1) \,|\, (\omega_1, M)] &= \delta(1-\phi_2(M,\omega_1))\phi_1(\omega_1, M) \\
\Pr[(\omega_1 - 1, M) \,|\, (\omega_1, M)] &= \delta\phi_2(M,\omega_1)(1-\phi_1(\omega_1, M)) \\
\Pr[(\omega_1 - 1, M-1) \,|\, (\omega_1, M)] &= \delta(1-\phi_2(M,\omega_1))(1-\phi_1(\omega_1, M)) \\
\Pr[(\omega_1 + 1, M) \,|\, (\omega_1, M)] &= (1-\delta)\phi_1(\omega_1, M).
\end{aligned}$$