Long-term Relationships: Static Gains and Dynamic Inefficiencies*

David Hémous† Morten Olsen‡

October 2016

Abstract

In the 1980s the Japanese 'keiretsu' system of interconnected business groups was praised as a model to emulate, but since then Japan has often been criticized for being less innovative than the United States. In this paper we connect the two views and argue that tight business relationships can create dynamic inefficiencies and reduce broad innovations. In particular, we consider the repeated interaction between final good producers and intermediate input suppliers, where the provision of the intermediate input is non-contractible. We build a cooperative equilibrium where producers can switch suppliers and start cooperation immediately with new suppliers. We first consider broad innovations: every period, one supplier has the opportunity to create a higher quality input that can be used by all producers. Since relationships are harder to break in the cooperative equilibrium the market size for potential innovators is smaller and the rate of innovation might be lower than in the non-cooperative equilibrium. We contrast this with a setting with relationship-specific innovations which we show are encouraged by the establishment of relational contracts. We illustrate the predictions of the model

*We are grateful to Daron Acemoglu, Philippe Aghion, Alberto Alesina, George-Marios Angeletos, Pol Antràs, Emmanuel Farhi, Oliver Hart, David Laibson, Jacob Leshno, Claire Lelarge, James Malcomson, Massimo Maoret, Nathan Nunn, Jennifer Page, Daniel Trefler and Timothy Van-Zandt for their thoughtful comments. We also thank seminar participants at Harvard University, the INSEAD-Georgetown University conference and CUNEF. Morten Olsen gratefully acknowledges the financial support of the European Commission under the Marie Curie Research Fellowship program (Grant Agreement PCIG11-GA-2012-321693) and the Spanish Ministry of Economy and Competitiveness (Project ref: ECO2012-38134).

†University of Zürich and CEPR

‡IESE Business School, University of Navarra

using the recent business history of the United States and Japan and further use patent data to show that U.S. patents are more general than Japanese and even more so in sectors using more differentiated inputs.

JEL: C73, K12, L14, O31, O43

KEYWORDS: contractibility, innovation, relational contracts, patents, repeated game

1 Introduction

An extensive literature exists on how relational contracts (that is implicit agreements relying on mutual trust) can overcome contractual incompleteness, encourage innovation and allow for risk sharing. A canonical example is the Japanese '*keiretsu*' system of interconnected business groups which, particularly in the 1980s, was widely seen as an economic system to be emulated by the rest of the world. Since then the pendulum has swung the other way and though Japan continues to have some of the most impressive firms in the world it is often criticized for not being as innovative as the United States (Dujarric and Hagiu, 2009). In this paper, we show that although the establishment of relational contracts can improve efficiency and encourage innovation within a relationship, the introduced rigidity can be detrimental to economic growth as it discourages broader innovation. In particular, firms engaged in relational contracting may be reluctant to switch to a new potential partner with a better technology, thus reducing the market size for a potential innovator and reducing the incentive to innovate. Based on this we compare the innovation pattern in the United States and Japan: while both countries are innovative, the U.S. innovations tend to be broader.

Figure 1.A provides an illustration of this using patents with at least one citation filed with the European Patent Office (EPO). It shows that whereas the share of patents that originate in the United States and Japan are about the same at around 13-14 per cent, when it comes to patents that are among the five per cent most cited, the U.S. share is 18 per cent compared with 13 for Japan. Trajtenberg, Henderson and Jaffe (1997) argue that the generality of a patent is better measured by the number of patent classes that cite the patent.[1] Figure 1.B demonstrates that generality only increases the difference between Japan and the United States and for the most general patents the share of the United States is almost three times as high as that of Japan. In section 5, we perform a more systematic analysis and show that the generality of US patents compared with Japan is more pronounced for products that are more differentiated (in the sense of Rauch, 1999). Our analysis aligns with Dujarric and Hagiu (2009) who study the case of Japan and argue that although Japan's keiretsu system has ensured a highly productive manufacturing sector, it has not been conducive to radically new innovations. As a consequence, Japan has failed to establish itself as a world leader in a

[1] They define generality of patent i is measured as $1 - \Sigma_{j=1}^{n_i} s_{i.j}^2$, where $s_{i,j}$ is the share of citations that patents i has received from patent class j and n_i is the total number of patent classes citations received by patent i. We regress this measure on dummies for the technological field and the year a patent was filed and use the residual in the figure. Nothing depends on these particular corrections.

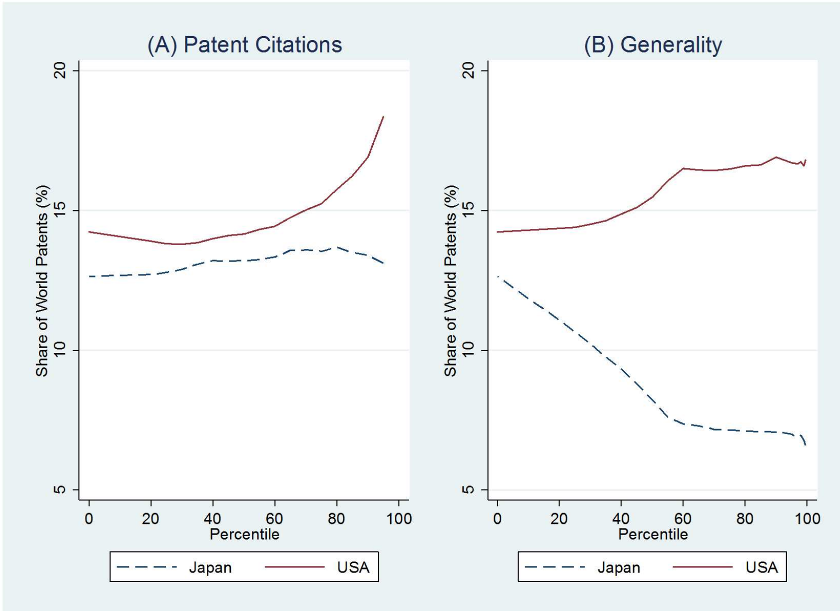


Figure 1: Relative Importance of Japanese and US Patents

number of new industries such as software and smartphones.

This paper provides a potential explanation for the differences in innovation patterns between the United States and Japan. More generally, it shows that relational contracts can be a poor substitute for good institutions because they transform contractibility issues from a static problem of inefficient allocation of resources into a dynamic problem of inefficient development of technologies. The paper focuses on growth and innovation, yet, relational contracts, requiring long-term relationships, can come at odds with economic efficiency, whenever the economy would benefit from flexible relationships.

We first consider broad innovations that are *not* specific to a relationship: We have in mind an industry with the following characteristics: (i) production requires the participation of producers and suppliers, where the suppliers provide complex inputs designed specifically for the final good producer, (ii) suppliers are competing with each other, and (iii) innovations allow them to "escape competition" and to increase their market share at the expense of their competitors. In a non-repeated framework, non-contractibility of the intermediate input typically creates an ex post hold-up situation leading to underinvestment by the supplier as in Grossman and Hart (1986). In a repeated framework, we rely on the existence of good and bad matches between producers and suppliers to build a "cooperative" equilibrium. Good matches are characterized by a higher productivity level. If a match turns out to be good, the value of the relationship in the following period is higher than the expected value of a new relationship. The supplier can capture the rents associated with this difference in values if cooperation with the producer continues, which induces her to invest more than the short-run interest would dictate. We contrast this case with two other cases: an economy with the same lack of contractibility, but where there is no cooperation in equilibrium (we refer to it as the "Nash case") and a setting in which inputs are fully contractible. The Nash and cooperative equilibria can be seen as two extremes on a spectrum, and we think of Japan as being closer to the cooperative case than the United States.

Every period, we let one supplier (the innovator) have the possibility to develop a new technology, which is imitated by her competitors after one period. Producers already engaged in a long-term relationship face a trade-off: switching to the innovator allows them to have access to a more productive technology, but at the risk of entering into a bad match. Entering into a bad match yields a lower productivity level no matter whether the input is contractible or not; but, when the input is noncontractible, bad matches are also characterized by more severe under-investment than good matches, since cooperation only occurs in the latter. Hence, bad matches become worse relative to good matches. This *worse bad match effect* is the main force behind our result that cooperation in a weak contractible setting magnifies rigidities in relationships.[2] Consequently, potential innovators have less of an incentive to develop technologies that require existing relationships to break up and fewer general innovations will be developed in countries where strong relational contracts are more widespread.[3]

We contrast this with a setting where innovation is done within an already established relationship. We show that relational contracts encourage this type of innovation both by improving the efficiency of production by overcoming the standard hold-up problem and because the introduction of relationship-specific innovation itself makes the parties more dependent on one another which further encourages cooperation.

Our model suggests that the Japanese economy did not loose steam *in spite* of the strong relational contracts, but perhaps *because* these strong relationships held back broader innovations. An extensive literature exists praising a superior Japanese economic

[2]Long-term relationships also present a barrier to entry in Aghion and Bolton (1987), who show that when an incumbent faces entry by potential competitors with superior technology, she will sign long-term contract that reduces the risk of entry. In our set-up, however, the relationship is of a different nature as the contract is implicit and we rule out explicit contracts that last more than a single period.

[3]The introduction of dynamic inefficiencies is not trivial and depends on the source of switching costs. In particular, in a model of exogenous fixed cost of switching, the cost of breaking up an existing relationship would be independent of the level of cooperation and relationships would not by themselves imply rigidity.

model. Dore (1983) discusses the Japanese economy as a whole and argues that relational contracts within the keiretsus overcome opportune behavior and allow for risk sharing. A prime example is the auto-industry (Helper,1991, and Helper and Henderson, 2014). In the heyday of the Japanese economy, there was little focus on the disadvantages of the Japanese economic system but the sluggish growth of Japan since the 1990s changed the tone of the literature. One such example is Dujarric and Hagiu (2009) who argue that "...[H]ierarchial industry organization can 'lock out' certain types of innovation indefinitely by perpetuating existing business practices". They argue that in the software industry in Japan, individual companies were part of keiretsus and developed advanced technological solutions for specific hardware producers. By contrast, in the United States a common platform developed which allowed for a competitive environment in which individual software developers had strong incentives to innovate to gain market share. Today the global software industry is dominated by American companies. We argue below that the history of the cellphone industry in Japan is similar. Collinson and Wilson (2006) describe Japanese chemical and steel production along the same lines.

Whether a country undertakes specific or broad innovations is important for welfare: With more actors building on an innovation, broader innovations tend to have disproportionately higher social returns (Bresnahan and Trajtenberg, 1995).[4] This is illustrated by the development of a broad software platforms that allowed for a subsequent host of products developed by third parties. Moreover, Jaffe, Trajtenberg and Henderson (1993) and Thompson and Fox-Kean (2005) show that knowledge spillovers as measured by patent citations are initially very localized, even for broader innovations.[5]

Our paper relates to an important literature on how the existence of relational contracts affect economic outcomes. Macaulay (1963) first showed that interactions between firms in most markets are repeated and that firms are engaged in relational contracts. More recently, the importance of relational contracts in developing countries has been highlighted by Banerjee and Duflo (2000) who show that in the Indian software industry,

[4]Bessen (2008) estimates that more general patents have more private value (controlling for the number of citations). The broadest of innovations are called "General Purpose Technology" (Bresnahan and Trajtenberg, 1995). An extensive literature exists on the importance of general purpose technology for both the level and the pattern of economic growth (see Jovanovic and Rousseau, 2005 and Helpman, 1998). Hall and Trajtenberg (2004) use generality as a measure to identify GPTs in patent data.

[5]A parallel can be made between our paper and Akcigit and Kerr (2016). They draw a distinction between external innovations to create new products (similar to general innovations in our framework) and internal innovations to improve a current product (similar to relationship-specific innovations here). They find that external innovations are more cited (indicating that they are more valuable). They calibrate an endogenous growth model and find that in the US, external innovation are responsible for 80% of economic growth.

reputation of firms matter for the kind of contracts they are offered. Macchiavello and Morjaria (2015) study relational contracts in the Kenyan rose export market and find that the value of a relationship increases with its age; while Macchiavello and Morjaria (2014) find that competition weakens relational contracts in Rwanda's coffee sector.[6]

To build our baseline model, we use the insights of Kranton (1996) and Ghosh and Ray (1996). They show that when a producer can switch suppliers at will, a cooperative equilibrium can only arise if there is a cost in switching partners (from the choice of equilibrium in Kranton, 1996, and from impatient players in Ghosh and Ray, 1996).

Two papers are close to our work: Board (2011) considers a simple hold-up problem where a principal invests in a supplier for the provision of an input. There are several suppliers and investment costs are stochastic. To prevent hold-up, a principal and a set of suppliers enter a relational contract where the principal is biased towards the suppliers with whom he has already worked. This implies that an outsider with a better technology is not systematically chosen, in line with our results.[7] Nevertheless, our paper goes further in several dimensions. First, we analyze how rigidities in turn affect the incentives to innovate. Second, in our set-up, cooperation can be welfare reducing, which it never is in his paper. Third, we also emphasize situations where the establishment of long-term relationships does not create rigidities.

Johnson, McMillan and Woodruff (2002) offer an empirical counterpart. They use an Eastern European survey of firms and show that in ongoing relationships, the belief in the efficiency of the court had very little impact on the level of trade credit, a proxy for the level of trust between firms. This suggests that firms engage in relational contracts. However, it matters a lot at the beginning of a relationship and for firms' incentives to try out new suppliers. Our model shares the same features, and may be understood as a rationalization of their results.

Our paper also relates to the literature on the impact of institutions on macroeconomic outcomes, both theoretical (Acemoglu, Antràs and Helpman, 2007) and empirical (Boehm, 2013, Cowan and Neut, 2007 and Nunn, 2007). Acemoglu, Aghion and Zilibotti (2003) show that institutions that favor the establishment of long-term relationships between firms and managers are appropriate far from the frontier but become a burden close to it. Bonfiglioli and Gancia (2014) present a similar trade-off. These papers,

[6] Allen, Qian, and Qian (2005) and Allen, Chakrabarti, De, Qian and Qian (2006, 2008) show in related papers that in India and China long-term relationships provide a way of financing firms.

[7] Similarly, Calzolari and Spagnolo (2009) find that a principal may want to limit the pool of agents that he chooses from to induce cooperation.

however do not allow for relational contracts. In contrast to our paper, Francois and Roberts (2003) study the impact of growth on contractual arrangements.

Finally, we can draw a parallel between our model and the Industrial Organization literature on buyer power and supplier innovation. For instance, Inderst and Shaffer (2007) find that following a merger, a retailer can increase its profits by reducing the number of its suppliers, which in return may lead suppliers to reduce the diversification of their products.[8] In our model, relational contracts push a buyer (the producer) to stay with the same supplier (and therefore to reduce the number of suppliers he works with) which reduces suppliers' innovation.

We start out by introducing the basic model in section 2, where we describe the cooperative equilibrium that we study and show that cooperation leads to rigid relationships. Section 3 studies the effect of cooperation on the rate of innovation. Section 4 demonstrates how cooperation encourages relationship-specific innovations. Section 5 discusses our results in light of a comparison between the United States and Japan and tests some of our results using patent data. Section 6 presents two extensions: a combination of both models and an extension of the baseline model where imitation need not occur after one period. Section 7 concludes. The proofs of the main results are available in Appendix A and the remaining proofs in the Online Appendix (Appendix B).

2 Cooperation and rigidity of relationships

In this section we show that relational contracts may lead to rigid relationships. We develop a general equilibrium model of repeated interaction between final good producers (he) and intermediate input suppliers (she), where some matches are exogenously more productive than others. We define a cooperative equilibrium in which the prospect of continuing the relationship in the following period provides suppliers in a good match with an incentive to invest more than they would do in the one-shot interaction. In this section, innovation occurs exogenously and grants one supplier a better technology. We show that a producer in a good relationship will be less willing to switch to the innovator when cooperation occurs relative to an equilibrium in which there is no cooperation (Nash case) or a setting with contractible inputs (contractible case).

[8] On the other hand, both Inderst and Wey (2011) and Fauli-Oller, Sandonis and Santamaria (2011) find that buyer power incentivizes upstream investment.

2.1 Preferences and production

A representative agent consumes a set of differentiated goods (denoted c_i) of measure 1, and a homogeneous outside good (denoted C_o) with a utility function given by:

$$U = \sum_{t=0}^{\infty} \frac{1}{(1+\rho)^t} \left(C_{o,t} + \frac{\sigma}{\sigma-1} \int_0^1 c_{jt}^{\frac{\sigma-1}{\sigma}} dj \right), \tag{1}$$

where ρ is the discount rate. We drop the subscript t when this does not lead to confusion.

The outside good is produced at constant returns to scale one for one with labor and we normalize its price and wages to 1 (we consider parameter values such that the outside good always remains active).[9] All the action in the model takes place in the production of differentiated goods. The demand for a variety j, c_j, and the quantity of variety j produced, q_j, can be written as a function solely of its own price:[10]

$$q_j = c_j = p_j^{-\sigma}. \tag{2}$$

There is a mass 1 of final good producers and each variety is associated with one producer who has the monopoly right over that variety. Final good producers die with probability δ^D every period and are replaced with new ones. Moreover, in every period, each final good producer must hire a single intermediate input supplier. There is a mass 1 of infinitely-lived intermediate input suppliers.[11] Each supplier can supply any number of final good producers without decreasing returns to scale.

More specifically, if the monopolist j hires the supplier k, the production technology is linear in the quantity of high quality inputs provided by the supplier:

$$q_j = (\theta_{jk} A_k)^{\frac{1}{\sigma-1}} X, \tag{3}$$

where θ_{jk} is a match specific and verifiable permanent level of productivity, A_k is the productivity of the intermediate input supplier k (with any producer) and X is the

[9]Because of technological progress, the differentiated sector will eventually become so productive, that the consumption of the homogeneous good is driven to 0. Hence, technically, we present an approximation which is only valid for sufficiently low productivity of the differentiated sector. Alternatively, if the productivity of the homogeneous good also grows at the rate of the technological frontier (through a knowledge externality), then our solution is exact. Nothing of substance depends on this.

[10]The functional form of the utility function allows us to avoid general equilibrium effects through wages (due to the homogeneous good) or the price index of the differentiated goods (as utility is separable). These features would complicate the analysis without changing any of our central results.

[11]We could equally have assumed that the intermediate good suppliers die with probability δ^D.

quantity of intermediate inputs of high quality provided by the supplier (we will refer to $\theta_{jk}A_k$ as productivity, although, strictly speaking productivity is given by $(\theta_{jk}A_k)^{\frac{1}{\sigma-1}}$). Producing one high quality intermediate input requires one unit of the homogeneous good, but the supplier can also produce an intermediate input with no value in production at 0 cost. The match-specific level of productivity θ_{jk} can take two values: $\theta_{jk} = 1$ in good matches or $\theta_{jk} = \theta < 1$ in bad matches. The quality of a match is revealed to both the supplier and the producer when they are matched (but before the supplier has incurred any investment) and is permanent.[12] A producer/supplier pair is a bad match with probability b. Once a supplier has been chosen, a period has to pass before the producer can form new relationships.

Throughout the paper we normalize the amount of high quality inputs provided by the supplier by $\theta_{jk}A_k$, and denote it x (such that $x \equiv X/(\theta_{jk}A_k)$). x also corresponds to the normalized investment level, as low quality inputs are produced costlessly. We can then express revenues as $\theta_{jk}A_kR(x)$, where $R(x) \equiv x^{\frac{\sigma-1}{\sigma}}$, and joint profits as $\theta_{jk}A_k\Pi(x)$ where $\Pi(x) \equiv x^{\frac{\sigma-1}{\sigma}} - x$. We think of a period as corresponding to several years. Hence, the quantity of intermediate inputs X captures not only an intermediate input per se, but different relationship-specific investments in physical or human capital.

2.2 Contractual incompleteness

We model contractual incompleteness as a classic hold-up problem (a simpler version of Grossman and Hart, 1986). More specifically, an input is specific to a particular producer and is useless to any other agent in the economy. Once a producer has chosen to work with a particular supplier, he cannot find another supplier for this period and the two are engaged in a bilateral monopoly. We briefly consider the one shot interaction in order to show the inefficiencies that repeated interactions can overcome.

If the input is contractible, the court can verify whether the input provided is of high or low quality. The producer and the supplier sign a contract where the normalized quantity of high quality inputs is at the first best level (m) given by:

$$m \equiv \arg\max_x R(x) - x = ((\sigma-1)/\sigma)^{\sigma}.$$

If the input is noncontractible, the court cannot verify the quality of the input.

[12] As explained in section 3.4, this is not a crucial assumption: the logic of our results would hold if the type of a match is revealed after investment has occurred.

We further make the classic assumption that revenues and expenditures of the parties are non-verifiable and cannot be part of a contract. There is a standard double hold-up problem: the producer can claim that the inputs are of low quality and refuse to pay and a supplier can costlessly deliver low quality inputs: any contract specifying the amount of inputs of high quality to be provided is worthless. Revenues are shared through ex-post Nash Bargaining, where $\beta \in (0,1)$ is the bargaining power of the supplier. Since she bears the full cost of the investment but is only paid a share β of the revenues she provides the amount of high quality input that maximizes her ex-post profits, the "Nash" normalized level of investment, n, given by:

$$n \equiv \arg\max_x \beta R(x) - x = \beta^\sigma m,$$

where there is naturally underinvestment: $n < m$.

Before the producer and the supplier start working together, an ex-ante cash transfer can be exchanged. If all suppliers are identical ex-ante ($A_k = 1$ for all k) Bertrand competition ensures that they make zero profits. Hence, the ex-ante transfer from the supplier to the producer is equal to $t = (1 - b + b\theta)(\beta R(m) - m)$ in the contractible case, and to $t = (1 - b + b\theta)(\beta R(n) - n)$ in the noncontractible one.

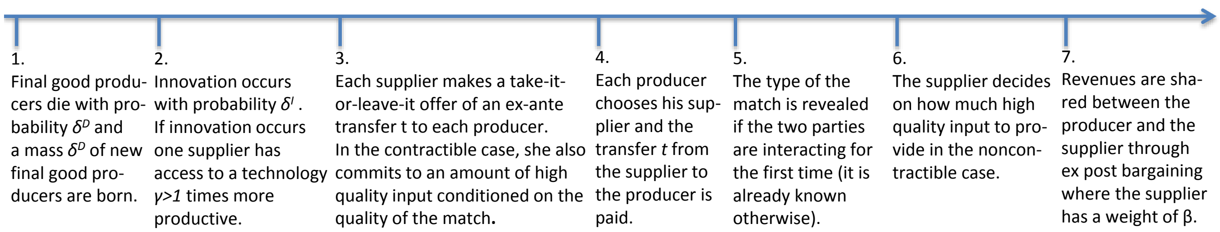
2.3 Innovation

We focus on "Schumpeterian" innovations where firms can capture a larger market share by improving the quality of their products (see Aghion, Akcigit and Howitt, 2015, for the relevance of Schumpeterian growth theory). We think of these innovations as representing broad innovations that can be adopted by several firms or sectors and study the case of relationship-specific innovations in section 4.

For the moment, we abstract from the innovation decision and assume that an innovation happens with probability $\delta^I \in (0,1)$—innovation is endogenized in section 3. When innovation occurs one of the suppliers gets access to a technology $\gamma > 1$ times more productive than the previous frontier technology, but, after a single period all suppliers have access to the new technology. This matches our view of each period corresponding to several years but can alternatively be viewed as reflecting relatively poor IPR protection. Section 6.2 presents a case where innovation diffuses. We denote by A the current frontier technology, so that, in periods without innovation all suppliers use technology A, and, in periods with innovation only the innovator uses the frontier technology while the other suppliers use $\gamma^{-1}A$.

2.4 Timeline

The overall timeline within each period is as follows:



Every stage game has three moves: in phase 3 suppliers make offers of ex-ante transfers, in phase 4 producers choose suppliers, and in phase 6 suppliers undertake the investment.[13] Since the transfer is in cash, it is verifiable and contractible.

2.5 Contractible, Nash and cooperative cases

Having described the model assumptions, we now characterize the two equilibria that we study when the input provision is noncontractible: *the Nash case* where the supplier's normalized investment level is fixed at the one-shot interaction level n and *the cooperative case* where in some matches the normalized investment level is above the Nash level n (that is the supplier "cooperates" with the producer). We contrast both equilibria with an alternative set-up: *the contractible case* where the input is fully contractible and the first best level of investment can be achieved even in a one-shot interaction.

Contractible and Nash cases. In both cases, a nascent producer switches suppliers until he finds a good match and stays with her in periods without innovation. The good match supplier offers an ex-ante transfer that allows her to capture the entire surplus of the ongoing relationship over any other relationship. In periods with innovation the producer optimally decides whether he should switch to the innovator (we study this in section 2.7). If the innovator turns out to be a bad match, the producer resumes working with his previous good match supplier in the following period. The only difference

[13] The assumption of suppliers making take it or leave it offers simplifies matters, but is not necessary. We could extend the model to include ex ante Nash bargaining over the surplus without affecting the incentive constraints (a similar result is demonstrated in MacLeod and Malcomson ,1989).

between the two cases is the investment level: it is given by the first-best level m in the contractible case and the Nash level n in the Nash case.[14]

Strategies of the cooperative equilibrium. There is a continuum of SPNEs featuring some level of cooperation. We consider strategies where the game is played independently for each producer. Our goal is to model a competitive industry where suppliers innovate in order to capture new customers. To capture this, we focus on a class of equilibria where (under some constraints) cooperation within new relationships is as high as possible from the beginning of the relationship— and as explained below our equilibrium satisfies a "bilateral rationality constraint". In other words, we consider a situation where relationships are relatively flexible because it is "easy" for a supplier to attract a producer, as she can offer him a high level of cooperation from the start.

We define a "cooperating good match supplier" as a good match supplier with whom no deviation triggering a punishment has occurred. Similarly a "non-cooperating good match supplier" is a good match supplier with whom a deviation (by either party) triggering a punishment has occurred. An "outdated" supplier is a supplier who does not have access to the frontier technology. We can now state the following proposition which characterizes the strategies played in our cooperative equilibrium (Appendix A.1 provides a formal proof and the following subsections explain the equilibrium intuitively):

Proposition 1. *The following strategies form a SPNE and apply for each producer independently:*

- S1) A cooperating good match supplier invests x^* if she has access to the frontier technology and y^* otherwise. A bad match and a non-cooperating good match invest n.

- S2) The values $x^*, y^* \in (n, m]$ are chosen so as to maximize the joint value of a relationship under the incentive compatibility constraint faced by a good match supplier.

- S3) The producer switches supplier until he finds a good match (in periods with innovation, he tries out the innovator), who then becomes a cooperating good match. Once the producer knows a cooperating good match, he sticks with her in periods without innovation. In periods with innovation, he optimally chooses between the outdated cooperating good match and the innovator depending on which relationship offers him the highest value. If the innovator turns out to be a good match, she becomes the new cooperating supplier.

[14] The contractible environment is still a world of limited contractibility as we do not allow for contracts across periods or between more than two parties. Hence, the equilibrium in the contractible case need not achieve the overall first best.

- S4) A cooperating good match supplier becomes non-cooperating when: i) she did not invest x^* or y^* when she should have, or ii) in a period without innovation, the producer picked a different supplier, or iii) in a period with innovation, the producer chose another outdated supplier, or iv) in a period with innovation, the producer chose the innovator and the innovator turned out to be a good match. A cooperating good match supplier does not become non-cooperating if the producer chose the innovator and the innovator turned out to be a bad match. All good match suppliers are initially cooperating.

- S5) If a deviation has occurred, the producer optimally chooses between starting a new relationship or sticking with the non-cooperating good match depending on which relationship offers the highest total discounted profits.

- S6) Ex-ante transfers are determined by Bertrand competition such that the producer is indifferent between his first best and second best option, and the second best supplier is indifferent between being chosen and not.

2.6 Characterizing the cooperative equilibrium

The equilibrium is characterized by two (normalized) investment levels: x^* in frontier good matches and y^* in outdated good matches. This subsection and the next explain how cooperation can be sustained, how x^* and y^* are determined, and the role played by the assumptions on the strategies in Proposition 1 S1-S6. Appendix A.2 gives a set of conditions on strategies which pin down the same SPNE.[15]

Value functions. To characterize the investment levels x^*, y^*, we must first derive the value functions. We normalize the value functions by the level of the frontier technology. In a period without innovation, we use the notation V_i^z to denote the beginning of the period normalized value of a producer ($z = p$), a supplier ($z = s$) or the total value ($z = T$) in a new relationship ($i = 0$) or a relationship with a cooperating good match supplier ($i = 1$). By "supplier value" here we only refer to the value that the supplier captures from working with that specific producer (since the game is played independently, we do not have to keep track of the value the supplier captures with other producers). In a period with innovation we similarly use W_i^z to denote the value with an outdated supplier ($i = 1$ for a cooperating good match and $i = 0$ for a new supplier).

[15]These are i) a symmetry and information condition, which ensures that equilibrium play does not depend on irrelevant information; ii) a forgiveness condition, which as discussed below, ensures that the supplier resumes working with a producer who switched to the innovator if the innovator turns out to be a bad match; iii) a bilateral rationality condition and iv) a no-cooperation in bad matches condition.

Consider the relationship between a producer and a cooperating good match supplier with the frontier technology. Their (normalized) joint value obeys:

$$V_1^T = \Pi(x^*) + \frac{1-\delta^D}{1+\rho}\left(\left(1-\delta^I\right)V_1^T + \delta^I \gamma W_1^T\right). \tag{4}$$

The current normalized profits are given by $\Pi(x^*)$ (per S1). If the producer survives, which happens with probability $1-\delta^D$, there are two possibilities for the next period. If no further innovation occurs, he keeps the same good match supplier. The situation is then identical to the current one, so the joint value is V_1^T. With probability δ^I, an innovation occurs, the frontier technology moves one step and the producer has to decide whether he should switch towards the innovator or stay with the now outdated good match supplier (per S3). This decision depends on parameters and is the subject of section 2.7. If he stays with the outdated good match, their joint value (normalized by the previous period's technology) is given by γW_1^T. If he switches, then per S6, Bertrand competition implies that the innovator captures the surplus over the producer's second best option, namely staying with the outdated good match; so that the joint value of the producer and the current good match is still given by γW_1^T.

The (normalized) joint value of a relationship between a producer and a cooperating, outdated, good match supplier in periods with innovation similarly obeys:

$$W_1^T = \frac{1}{\gamma}\Pi(y^*) + \frac{1-\delta^D}{1+\rho}\left(\left(1-\delta^I\right)V_1^T + \delta^I \gamma W_1^T\right). \tag{5}$$

The current normalized profit flow is now given by $\gamma^{-1}\Pi(y^*)$: the supplier's technology is one step below the frontier and the supplier invests y^* instead of x^* (per S1). If no innovation occurs in the following period, the supplier gets access to the frontier technology and the producer sticks to that supplier (unless a deviation has occurred), so that their normalized joint value is V_1^T. If another innovation occurs, the situation is the same as for equation (4) as the supplier remains just one step below the new frontier.

Consider now a producer that has not yet met a good match supplier. In a period without innovation, following S3, this producer will start a new relationship (he has no interest in staying with a bad match whose productivity is below average and who does not cooperate). The joint value of starting a new relationship obeys:

$$V_0^T = (1-b)V_1^T + b\theta\Pi(n) + b\frac{1-\delta^D}{1+\rho}\left(\left(1-\delta^I\right)V_0^T + \delta^I \gamma W_0^T\right). \tag{6}$$

With probability $1-b$, the supplier is a good match. Once this is revealed, the supplier can invest x^* and the joint value is simply V_1^T. With probability b, the supplier turns out to be a bad match, and current profits are only $\theta\Pi(n)$ as both productivity and cooperation are lower (per S1). The continuation value of a bad match supplier (with this producer) is 0, since the producer never returns to that supplier. Instead the producer will start another relationship in the next period. If no innovation occurs, several firms will have access to the frontier technology, so that through Bertrand competition (per S6), the producer captures the whole expected value of this new relationship: $V_0^p = V_0^T$. When an innovation occurs the producer gets the value of his second best option (per S6), namely a new relationship with an outdated supplier: γW_0^T.

The law of motion for W_0^T is similarly given by:

$$W_0^T = (1-b)W_1^T + b\frac{1}{\gamma}\theta\Pi\left(n\right) + b\frac{1-\delta^D}{1+\rho}\left(\left(1-\delta^I\right)V_0^T + \delta^I\gamma W_0^T\right). \tag{7}$$

With probability $1-b$, the outdated supplier turns out to be a good match delivering the joint value W_1^T. With probability b, the outdated supplier is a bad match so that the investment level is n (per S1), and productivity is γ times below the frontier. The following period is identical to the previous case where the producer meets a new supplier with the frontier technology since the frontier technology diffuses after one period.

To find the value functions for suppliers and producers, consider again a producer who knows a cooperating good match supplier. In a period without innovation, his second best option is to start a new relationship. Per S4, if he does so, and therefore deviates on his previous good match supplier, she becomes non-cooperating, but other potential good match suppliers are still willing to cooperate immediately, delivering the joint value V_0^T.[16] By Bertrand competition (see S6), on equilibrium path, the producer in a relationship with a cooperating good match supplier must capture his second best option, namely V_0^T, while the cooperating good match supplier can capture the surplus of a relationship with her over the producer's second best option. Therefore, we obtain:

$$V_1^p = V_0^T \text{ and } V_1^s = V_1^T - V_0^T. \tag{8}$$

[16]Technically, the presence of a non-cooperating good match supplier could affect the value of starting a new relationship so that in equation 8, one should replace V_0^T with $V_0^{T,n}$, the value of starting a relationship when the producer knows a non-cooperating good match. Nevertheless, if the producer always prefers trying a new supplier to staying with a non-cooperating good match, then $V_0^T = V_0^{T,n}$ and (8) holds (see Appendix A.1 which deals with this issue rigorously).

Incentive compatibility constraints. We can now describe the incentive compatibility (IC) constraints faced by suppliers. After the ex-ante payment, a cooperating frontier supplier has a short run incentive to deviate from x by investing n. She would then gain $\varphi(x) A_k$, where A_k is the technology used by the supplier and:

$$\varphi(x) \equiv (\beta R(n) - n) - (\beta R(x) - x). \tag{9}$$

Per S4, the cost of such deviation is that cooperation ceases, and the producer expects the now non-cooperating good match to only invest n from now on. Per S5, he optimally decides between trying a new supplier or sticking with this non-cooperating good match. Here, we focus on parameters for which the former occurs, so that the continuation value of a non-cooperating good match supplier (with that producer) is 0 (see Appendix B.1 for the other cases). On the other hand, if the supplier cooperates and the producer survives, their relationship continues (per S3) and her continuation value is V_1^s if there is no innovation and γW_1^s otherwise (W_1^s is derived in Appendix A.1). Therefore, a good match supplier who has access to the frontier technology faces an incentive compatibility constraint given by:

$$\varphi(x^*) \leq \frac{1-\delta^D}{1+\rho}\left(\left(1-\delta^I\right) V_1^s + \delta^I \gamma W_1^s\right). \tag{10}$$

This IC constraint applies to any cooperating good match supplier with the frontier technology: an old supplier in a period without innovation, a new supplier who turns out to be a good match in a period without innovation, or the innovator if she turns out to be a good match. Indeed, at the time of the investment decisions, all these suppliers face the same situation (in particular a good match innovator is not different because in the next period, all suppliers will have her technology). This justifies our assumption S1 that all good match suppliers with the frontier technology cooperate to the same extent (as long as no deviation has occurred for them).

Without innovation, the incentive constraint of an outdated good match supplier is:

$$\gamma^{-1}\varphi(y^*) \leq \frac{1-\delta^D}{1+\rho}\left(\left(1-\delta^I\right) V_1^s + \delta^I \gamma W_1^s\right). \tag{11}$$

The right-hand sides of (10) and (11) are identical but the left hand side of (11) is lower at equal levels of investment, which is why we had to allow for two different levels of investment x^* and y^* in S1. To understand this, consider a period with innovation, then the incentive to deviate in a good match is scaled by the technology currently used, which is lower for an outdated supplier than for the innovator, whereas the reward from cooperation is scaled by the technology available in the next period which is the same for an outdated supplier and the innovator since imitation occurs after one period.

Following S2, we focus on equilibria where investment levels x^* and y^* maximize joint profits under these incentive constraints, so that either x^* and y^* are equal to the first best or their respective incentive compatibility constraint (10) or (11) bind. By definition $\varphi(n) = 0$, so the investment levels are higher than the Nash level: $x^*, y^* > n$. Further, since (11) is a laxer constraint than (10), the level of investment with an outdated supplier must be weakly higher than with a frontier supplier, i.e. $y^* \geq x^*$ (with equality if $x^* = m$). Appendix A.1 derives explicitly the right-hand side of both IC constraints which then allows us to fully characterize the investment levels x^*, y^*.[17]

Bilateral rationality. Together, S1, S2 and S4 imply that our equilibrium satisfies a "bilateral rationality" condition for good matches, in that a producer and a new supplier's strategies are such that cooperation is maximized right from the beginning of the relationship (see Appendix A.2). In particular, if a producer had deviated on a supplier before, new suppliers are still willing to cooperate with that producer right away (in other words, suppliers do not coordinate to enforce cooperation). This feature of the equilibrium matches our goal capturing an industry where suppliers innovate to increase their market share and therefore should not start by "punishing" new customers.

It is well known in the literature that generating cooperation when players can switch partners at will requires a switching cost: otherwise, the threat of retaliation from the current partner does not carry any force and the cost from not cooperating is non-existing. The switching cost here is the risk of finding a bad match.[18] A good match supplier benefits from having been revealed as such. This informational advantage acts as a fixed cost which pushes the producer to stick to the same supplier, who can then capture the associated rents. The prospect of capturing these rents induces cooperation

[17] Appendix B.1 does the same for parameter values such that the producer does not always prefer trying a new supplier to staying with a non-cooperating good match. In addition, Appendix A.3 provides comparative static results on how the investment levels depend on the parameters of the model.

[18] Kranton (1996) first demonstrates that in a setting with identical agents and costless switching between partners any equilibrium featuring more cooperation than a one shot interaction cannot be "pair-wise enforceable": any equilibrium with cooperation requires some initial cost of a new relationship from lower initial cooperation, but when two new partners first meet they could credibly agree to skip the initial low level of cooperation and the equilibrium unravels. Both Ghosh and Ray (1996) but also Kranton (1996) build equilibria that overcome this by introducing impatient players who never cooperate. The existence of such players (similar to our bad matches here) serves as an expected cost of establishing a new relationship and enables cooperation.

by a good match supplier in the first place. Crucially, this fixed cost interacts naturally with the incomplete contractibility: there is no cooperation in bad matches, as bad matches have no prospect.[19] Hence bad matches are "even worse" relatively to good matches in the cooperative case than in the contractibility or the Nash cases.

2.7 To switch or not to switch

Having demonstrated that long-term relationships can mitigate the under-investment from contractual incompleteness, we now show that cooperation makes relationships more rigid. In periods when one innovator has a superior technology, producers without good match suppliers try her. Producers already in a good match face the trade-off of accessing the better technology, but at the risk of engaging in a bad match.

Consider first the contractible case. The technological advantage of the innovator lasts for only one period and, if the innovator is a bad match, the producer reverts to his old supplier. Therefore, the producer switches to the innovator if and only if:

$$1 - b + b\theta > \gamma^{-1}, \text{ or equivalently, } \gamma > \gamma^{con} \equiv (1 - b + b\theta)^{-1}. \tag{12}$$

With probability b the new supplier is a bad match, but her technology is γ times more productive. The Nash case is identical except with investment levels at n. Hence, the producer switches to the innovator if and only if $\gamma > \gamma^{Nash} = \gamma^{con}$.

We now turn to the cooperative equilibrium built above. A producer previously in a good match switches to the innovator, if and only his expected value with the innovator is higher than that with his old supplier. These expected values depend on whether the outdated good match supplier punishes a producer who switches to the innovator. S3 and S4 stipulate that the outdated supplier punishes the producer for switching if the innovator turns out to be a good match. Otherwise, in the following period, there would be two good matches (the innovator and the previous good match) willing to cooperate with the producer with the same technology. This means that neither could capture any value which precludes any cooperation in the current period. However, S3 and S4 also stipulate that if the producer switches, and the innovator turns out to be a bad match, then the previous supplier forgives the producer and resumes cooperation in the

[19] If θ is low enough, it is in fact impossible to build an equilibrium which features cooperation in all bad matches. For higher θ one can build mixed strategy equilibria where some, but not all, bad matches feature some level of cooperation. Allowing for such would alter little in our general analysis, but would complicate both exposition and notation.

following period (that innovator is no longer viewed as a threat).[20] As a result, the decision to switch depends only on profits in the first period because in the next period the producer will be with a cooperating good match supplier in either case.

The innovator is a good match with probability $1-b$, in which case she invests x^* and a bad match with probability b, in which case she invests n, while the old good match supplier invests y^* and her technology is γ times less productive. Hence, producers previously in a good match switch to the innovator if and only if $(1-b)\,\Pi(x^*)+b\theta\Pi(n) > \gamma^{-1}\Pi(y^*)$ (Appendix A.1 provides a formal proof). This can be written as:

$$1-b+b\theta\left(\Pi(n)/\Pi(x^*)\right) > \gamma^{-1}\left(\Pi(y^*)/\Pi(x^*)\right). \tag{13}$$

Cooperation occurs in good matches so $x^* > n$, moreover, as explained above $y^* \geq x^*$. Therefore $\Pi(n)/\Pi(x^*) < 1$ and $\Pi(y^*)/\Pi(x^*) \geq 1$ so that (12) is more easily satisfied than (13), which gives us:

Proposition 2. *(i) The parameter set for which innovators capture the whole market in the cooperative case is a subset of the parameter set for which innovators capture the whole market in the contractible or the Nash cases. (ii) In particular, the minimum technological leap required for an innovator to capture the whole market in the cooperative case (γ^{coop}) is higher than that in the contractible or Nash cases: $\gamma^{coop} > \gamma^{con} = \gamma^{Nash}$.*

This proposition is the result of two effects: First, a *worse bad matches effect* as a bad match is more costly relative to a good match in the cooperative case. Indeed, bad matches have an inherently lower productivity level, but they also involve less investment as both parties realize that the relationship will come to an end in the following period. This effect is captured by the term $\Pi(n)/\Pi(x^*)$ in (13).[21] It makes switching to the innovator a riskier activity when the producer is engaged in a relationship.

The second effect is an *encouragement effect*, which comes through the term $\Pi(y^*)/\Pi(x^*)$ in (13): The opportunity to obtain the frontier technology in the following period encour-

[20] This assumption is the natural starting point because, as explained below, it makes the switching decision jointly efficient and it simplifies exposition. We consider the opposite case—where a supplier always punishes a producer if he switches supplier, no matter what happens with the new supplier—in Appendix A.4 and demonstrate that under quite general conditions the qualitative results are the same and that the inability to revert back adds another source of rigidity from cooperation.

[21] Recall that for θ sufficiently small, cooperation in bad matches is necessarily impossible. For θ not small enough, the fact that there is no cooperation in bad matches directly results from our assumption on the equilibrium play. However, even if a pair were to deviate and start cooperating, the level of normalized investment would be lower than in good matches, and so even this "cooperative" bad match would be relatively worse, than in the contractible or Nash cases.

ages an outdated supplier to provide a larger effort, which partly erodes the technological advantage of the innovator. The encouragement effect is especially strong when imitation happens after only one period, but exists as long as the supplier has a positive probability of getting access to the frontier (see also section 6.2).

Proposition 2 delivers the first important message of the paper: in a context of weak contractibility, cooperation makes it more difficult to break up existing relationships. Because of the existence of bad matches, for γ sufficiently low, innovations are not adopted by suppliers in good matches, but the threshold for adoption is higher in the cooperative case than in the contractible or Nash cases. Importantly, so far, there are no welfare cost from this "rigidity" of relationships and welfare in the cooperative equilibrium is necessarily higher than in the Nash case. Indeed, because of the "forgiveness condition", the decision to switch or not is jointly efficient for the outdated supplier, the innovator and the producer.[22] There are only welfare costs once the innovation rate is endogenized.

Furthermore, when $\gamma \in (\gamma^{con}, \gamma^{coop})$, Proposition 2 directly predicts that technological differences across firms should be more important in countries with poor contractibility institutions and high level of cooperation/trust than in countries with good institutions or poor institutions but very low level of cooperation/trust. This is line with a large literature, started with Hsieh and Klenow (2009), which argues that productivity differences are larger in developing than in developed countries.

3 Endogenous innovation

Subsection 2.7 showed that cooperation creates rigidity in long-term relationships. We now turn to the issue of how this rigidity can be the source of dynamic inefficiencies by endogenizing the rate of innovation. We show that it is reduced with noncontractibility and may be further reduced by cooperation.

3.1 Rate of innovation

To endogenize innovation we choose a simple setting, but since the crucial element is the impact of relational contract on the value of an innovation, it should be clear that our

[22] In fact, from a welfare point of view, at a given rate of innovation, producers switch to the innovator "too much". As final good producers are monopolists (the level of normalized investment that maximizes welfare is higher than m) bad matches are even more detrimental to welfare than to profits, and switching to the innovator inevitably involves more bad matches.

results hold more generally. Every period, one supplier gets a new idea which turns into a useful innovation with probability δ^I if the potential innovator invests $A\psi\left(\delta^I\right)$ (where A is the frontier technological level before innovation occurs). The function ψ is convex with $\psi(0)=0$, $\psi'(0)=0$ and $\lim_{\delta^I\to 1}\psi'\left(\delta^I\right)=\infty$. Since the probability that the potential innovator has already made a successful innovation is infinitesimal, the previous period market share of the potential innovator is infinitesimal and for all purposes the potential innovator is an entrant. Here, we compare the rates of innovation in the three different cases: *contractible*, *Nash* and *cooperative*.

Thanks to Bertrand competition the innovator captures the entire surplus of a relationship with her over the second best option of the producer.[23] Because imitation occurs after one period and a supplier forgives a producer who switches to the innovator if the innovator turns out to be a bad match, this surplus corresponds to the difference in profits between the two options in the first period (if it is positive). We denote by $V_{I,K}^{s,t}$ the value captured by the innovator (normalized by the frontier productivity level) from a relationship with a producer, who knows a good match supplier ($t=g$), or does not ($t=b$), for the contractible ($K=con$), the Nash ($K=Nash$) and the cooperative cases ($K=coop$). In the contractible case, depending on whether the producer is in a relationship with a good match or not, the value captured by the innovator is given by:

$$V_{I,con}^{s,g}=\left(1-b+b\theta-\gamma^{-1}\right)^{+}\Pi(m) \text{ or } V_{I,con}^{s,b}=(1-b+b\theta)\left(1-\gamma^{-1}\right)\Pi(m). \tag{14}$$

The situation of producers previously in a good match has been analyzed in (12). The reasoning is similar for the other producers: joint expected profits are the same with the innovator and any other supplier except in the first period where they are γ times higher with the innovator. Similarly, for the Nash case, we get:

$$V_{I,Nash}^{s,g}=\left(1-b+b\theta-\gamma^{-1}\right)^{+}\Pi(n) \text{ and } V_{I,Nash}^{s,b}=(1-b+b\theta)\left(1-\gamma^{-1}\right)\Pi(n). \tag{15}$$

Finally, in the cooperative case, we get:

$$V_{I,coop}^{s,g}=\left((1-b)\Pi(x^*)+b\theta\Pi(n)-\gamma^{-1}\Pi(y^*)\right)^{+}, \tag{16}$$

$$V_{I,coop}^{s,b}=(1-b)\left(\Pi(x^*)-\gamma^{-1}\Pi(y^*)\right)+b\theta\left(1-\gamma^{-1}\right)\Pi(n). \tag{17}$$

[23] More generally, with ex-ante Nash Bargaining, the innovator would capture only part of the difference, but as long as she captures a constant positive part, the results of this subsection carry through.

The case of producers previously in good matches follows from the derivation of (13): if the producer switches his expected profits are $(1-b)\,\Pi(x^*)+b\theta\Pi(n)$, if he stays with his old (good) supplier, she will deliver effort y^*, but will use a technology which is γ times below the frontier, generating profits $\gamma^{-1}\Pi(y^*)$. The innovator captures the difference if it is positive, which gives (16). For a producer who does not know a good supplier, the alternative to starting a relationship with the innovator is to try a new (outdated) supplier. Such a supplier would bring expected profits $\gamma^{-1}((1-b)\,\Pi(y^*)+b\theta\Pi(n))$ as she is a good match with probability $1-b$ but uses a technology γ times below the frontier. The innovator captures the difference in profits, namely (17).

In equilibrium, the steady-state fraction of firms previously not in a good match is constant given by $\omega=\delta^D/\left(1-\left(1-\delta^D\right)b\right)$.[24] Hence, assuming that the steady state has been reached, the innovator solves the problem:

$$\max_{\widehat{\delta}}\gamma\widehat{\delta}\left[\omega V^{s,b}_{I,K}\left(\delta^I\right)+(1-\omega)\,V^{s,g}_{I,K}\left(\delta^I\right)\right]-\psi\left(\widehat{\delta}\right), \tag{18}$$

for $K\in\{con,\ Nash,\ coop\}$. We denote by $Z_K=\omega V^{s,b}_{I,K}\left(\delta^I\right)+(1-\omega)\,V^{s,g}_{I,K}\left(\delta^I\right)$, the expected total value of an innovator. The first order condition $\psi'\left(\delta^I\right)=\gamma Z_K$ uniquely defines the equilibrium rate of innovation in the contractible case (δ^{con}), and in the Nash case (δ^{Nash}). In the cooperative case, the value of the innovator depends on the equilibrium rate of innovation, so any fixed point of the first order condition would be a solution to the problem. We consider the highest one and denote it δ^{coop} (alternatively we could assume that ψ is sufficiently convex to rule out multiple equilibria). A higher expected value Z_K leads to a higher rate of innovation.

From (16) and (17) the reward from innovation in the cooperative case is:

$$Z_{coop}=\Pi(x^*)\left[\begin{array}{c}\omega\left((1-b)\left(1-\gamma^{-1}\frac{\Pi(y^*)}{\Pi(x^*)}\right)+b\theta\left(1-\gamma^{-1}\right)\frac{\Pi(n)}{\Pi(x^*)}\right)\\+(1-\omega)\left((1-b)+b\theta\frac{\Pi(n)}{\Pi(x^*)}-\gamma^{-1}\frac{\Pi(y^*)}{\Pi(x^*)}\right)^+\end{array}\right], \tag{19}$$

implying that Z_{coop} is an increasing function of $\Pi(x^*)$ and $\Pi(n)/\Pi(x^*)$ and decreasing in $\Pi(y^*)/\Pi(x^*)$. In the Nash and contractible cases the ratios are replaced by 1 and $\Pi(x^*)$ by $\Pi(n)$ and $\Pi(m)$ respectively. The comparison between the innovation rates in the three cases results then from three effects. The *worse bad match* effect reduces

[24] ω is the share of firms that know a good match supplier willing to cooperate with them. It does not depend on the rate of innovation, because when an innovation occurs, producers do not lose the possibility to cooperate with their old supplier.

the expected gain from innovation in the cooperative case as the lower productivity of a bad match will be further amplified by the lack of cooperation (this is reflected in $\Pi(n)/\Pi(x^*) < 1$ in (19)). The *encouragement* effect ($\Pi(y^*)/\Pi(x^*) \geq 1$ in (19)) induces more cooperation from the existing supplier in the cooperative case which reduces the gain from switching and therefore the value of the innovator. And the *scale effect*: a higher level of investment by frontier good matches increases profitability should the innovator turn out to be a good match which increases the incentive to innovate ($\Pi(n) < \Pi(x^*) \leq \Pi(m)$). Comparing the contractible case to the cooperative one, all effects go in the same direction and $\delta^{coop} < \delta^{con}$ unambiguously. Comparing the cooperative and Nash cases, the worse bad match and the encouragement effects push towards $\delta^{Nash} > \delta^{coop}$, but the scale effect pushes in the other direction, giving an ambiguous result.

To go further, we investigate in turn what happens for different innovation sizes. First, for sufficiently small innovation sizes, no producer in a good match would try the innovator (that is $\gamma \leq \gamma^{Nash} = (1-b+b\theta)^{-1}$). In this case we use (15) and (17), and the difference in expected value is given by:

$$Z_{Nash} - Z_{coop} = \omega(1-b)\left(\left(1-\gamma^{-1}\right)\Pi(n) - \left(\Pi(x^*) - \gamma^{-1}\Pi(y^*)\right)\right). \tag{20}$$

As shown in Appendix B.4, $\Pi(x^*) - \gamma^{-1}\Pi(y^*) > (1-\gamma^{-1})\Pi(n)$ implying that the scale effect ($\Pi(x^*) > \Pi(n)$) always dominates the encouragement effect ($\Pi(y^*)/\Pi(x^*) \geq 1$). The innovator captures more from producers not in good matches in the cooperative case than in the Nash case ($V^{s,b}_{I,coop} > V^{s,b}_{I,Nash}$) and we must have $\delta^{Nash} < \delta^{coop}$.

For intermediate values of $\gamma \in (\gamma^{con}, \gamma^{coop})$,[25] innovation breaks relationships in the Nash case but not in the cooperative case. Using (15) and (17), we get that

$$Z_{Nash} - Z_{coop} = (1-\omega)\left(1-b+b\theta-\gamma^{-1}\right)\Pi(n) + \omega(1-b)\left(\left(1-\gamma^{-1}\right)\Pi(n) - \left(\Pi(x^*) - \gamma^{-1}\Pi(y^*)\right)\right). \tag{21}$$

In that case the excess rigidity of relationships in the cooperative case creates an *extensive margin* by which an innovator has a lower market size in the cooperative case than in the Nash case. If the death rate of producers δ^D is low, most producers will have found a good match (ω is small), and the market captured by an innovator in the cooperative case

[25] Although we denote by γ^{coop} the size of innovation necessary for switching in the cooperative case in both this section and the preceding, they are mathematically different objects. In the preceding section, γ^{coop} was a function of the exogenous rate of innovation δ^I. In this section, δ^I is a choice variable so γ^{coop} is no longer a function of δ^I. Not making this explicit in the text should not lead to confusion. Further since γ^{Nash} is independent of the innovation rate δ^I, Proposition 2 still applies.

is much smaller than in the Nash case. Cooperation reduces innovation: $\delta^{Nash} > \delta^{coop}$.[26]

Finally, if γ is large enough, $\gamma > \gamma^{coop}$, innovation breaks relationships in all cases, so that this extensive margin disappears. Using (15), (16) and (17), we obtain:

$$Z_{Nash} - Z_{coop} = \frac{(1-b)}{1-\left(1-\delta^D\right)b}\left[\left(1-b\left(1-\delta^D\right)-\gamma^{-1}\right)\Pi(n) - \left(\left(1-b\left(1-\delta^D\right)\right)\Pi(x^*) - \gamma^{-1}\Pi(y^*)\right)\right]. \tag{22}$$

For innovation sizes sufficiently close to γ^{coop}, the innovator still captures little from producers previously in a good match relationship so that innovation is lower in the cooperative case $\delta^{Nash} > \delta^{coop}$ ($\gamma < \left(1-b\left(1-\delta^D\right)\right)^{-1}$ is a sufficient condition). On the other hand, for γ sufficiently large, the outdated supplier is at too large a disadvantage regardless of her effort level, the scale effect dominates and $\delta^{coop} > \delta^{Nash}$.

In particular we can derive the following proposition (proof in Appendix B.4), which combines the three cases but uses stricter assumptions in order to provide sufficient conditions that do not depend on endogenous variables such as x^*, y^* and γ^{coop}.

Proposition 3. *a) The rate of innovation is the highest in the contractible case: $\delta^{con} > \delta^{Nash}, \delta^{coop}$. b) If innovations are small enough ($\gamma \leq \gamma^{Nash}$) or if they are large enough, then the innovation rate in the cooperative case is higher than in the Nash one $\delta^{Nash} < \delta^{coop}$. c) Assume that the death rate of producers is low enough $\delta^D < \theta\frac{\Pi(n)}{\Pi(m)}$, then for an intermediate range of innovation sizes, $\gamma \in \left(\frac{1-\delta^D\frac{\Pi(m)}{\Pi(n)}}{1-\delta^D\frac{\Pi(m)}{\Pi(n)}-b(1-\theta)(1-\delta^D)}, \frac{1}{1-b(1-\delta^D)}\right)$ the innovation rate is higher in the Nash case than in the contractible one: $\delta^{Nash} > \delta^{coop}$.*

A case of special interest is when the cooperative equilibrium can achieve the first best level of efforts in good matches (that is the static inefficiencies are fully overcome). We then obtain the following Remark, which stipulates that the condition of Part c) of the previous proposition is now both sufficient and necessary.

Remark 1. Assume that the cooperative equilibrium ensures the first best level of investment in good matches ($y^* = x^* = m$), and that $\delta^D < \theta\frac{\Pi(n)}{\Pi(m)}$, then the innovation rate is higher in the Nash than the contractible case, $\delta^{Nash} > \delta^{coop}$, if and only if $\gamma \in \left(\frac{1-\delta^D\frac{\Pi(m)}{\Pi(n)}}{1-\delta^D\frac{\Pi(m)}{\Pi(n)}-b(1-\theta)(1-\delta^D)}, \frac{1}{1-b(1-\delta^D)}\right)$. The higher is the level of productivity of bad matches, θ, the smaller is the ratio $\Pi(m)/\Pi(n)$ (that is the smaller is the scope for static inefficiencies) and the smaller is the death rate of producer δ^D, the more likely it is that $\delta^{Nash} > \delta^{coop}$.

[26] Specifically, using the expression for ω, we obtain that $\delta^{Nash} > \delta^{coop}$ if and only if $\delta^D\left(\frac{\Pi(x^*)}{\Pi(n)} - b + b\theta - \gamma^{-1}\frac{\Pi(y^*)}{\Pi(n)}\right) < 1 - b + b\theta - \gamma^{-1}$.

Intuitively, for low θ, production in bad matches is already low regardless of whether cooperation occurs or not. The "worse bad match" effect is dominated by the scale effect and cooperation increases the innovation rate. On the other hand, a small death rate of producers δ^D increases the share of producers already in a good match. Since it is those producers that an innovator may fail to capture in the cooperative case, it becomes more likely that cooperation reduces innovation.

Our model predicts that, for intermediate size of innovations, relationships should last longer in countries with poor contractual enforcement but high level of cooperation relative to countries with either high level of contractual enforcement or low level of trust. Indeed, if $\gamma < \gamma^{cont} = \gamma^{Nash}$, relationships are never broken (unless the producer dies). If $\gamma \in (\gamma^{con}, \gamma^{coop})$, relationships never break up in the cooperative case but do so in the contractible and Nash cases. While if $\gamma > \gamma^{coop}$, relationships break up with innovation in all cases, but as long as $\gamma < 1/(1 - b(1 - \delta^D))$, innovations are the least frequent in the cooperative case.

3.2 Welfare

As innovation is already too low from a welfare perspective because of standard innovation-externalities of imitation and building-on-the-shoulders-of-giants, a lower rate of innovation can easily translate into lower welfare. Relative to the Nash equilibrium, cooperation enhances investment and reduces the static inefficiencies. However, it may also reduce the innovation rate, aggravating the dynamic inefficiency in the economy. When the discount rate ρ is sufficiently low, dynamic inefficiencies matter more for welfare than static ones, so that cooperation reduces welfare when it reduces innovation. We obtain:

Corollary 1. *Welfare is always lower with incomplete contractibility than with complete contractibility. Welfare may be higher or lower in the cooperative case than in the Nash case, but when the discount rate ρ is sufficiently low, the death rate of producers satisfies $\delta^D < \theta \frac{\Pi(n)}{\Pi(m)}$, and for an intermediate range of innovation sizes, $\gamma \in \left(\frac{1-\delta^D \frac{\Pi(m)}{\Pi(n)}}{1-\delta^D \frac{\Pi(m)}{\Pi(n)} - b(1-\theta)(1-\delta^D)}, \frac{1}{1-b(1-\delta^D)} \right)$, cooperation reduces welfare.*

The fact that the rate of innovation is inefficient to start with is essential for this result. Relationships make the profitability of a new innovation smaller for the innovator, but that loss in itself cannot outweigh the benefit of higher investment that comes from the relationship. It is only because innovation is already too low (such that a further reduction lowers welfare for society as a whole) that relationships can decrease overall welfare. Consider alternatively a setup in which an innovation is temporary such that the innovator returns to the old technology after one period and no imitation is possible (such that both imitation and 'building on the shoulders of giants' have been precluded). In such a case, private and social benefits of an innovation are equal.[27] All our results except corollary 1 would still hold.

3.3 Cooperation and expanding varieties

The main proposition of this paper is that cooperation can be a poor substitute for full contractibility as it might reduce new innovations. Countries with higher levels of cooperation vary widely, from mature developed economies like Japan, to rapidly growing economies like India. In the following we show that the existence of relational contracts is more likely to reduce growth for more mature economies.

Extend the model such that the mass of final good producers, N_t, is increasing: $N_{t+1} = N_t (1 + g_N)$. This could represent catch-up growth, horizontal innovation, population growth or periods of increasing outsourcing (interpreting the new final good producers as foreign firms who decide to start acquiring their inputs from the country of study). Innovation costs scale by the number of products (they are given by $\psi\left(\delta^I\right) AN$), so that the innovation problem is independent of the number of products. This ensures that if the share of firms who know a good match is at the steady-state level, the innovation rate is constant and the cooperative equilibrium keeps the same structure as before with constant x^* and y^*. We can then show (proof in Appendix B.5):

Remark 2. Assume that the cooperative equilibrium ensures the first best level of investment in good matches ($y^* = x^* = m$), then the lower is the growth rate of product g_N, the larger is the set of γ for which $\delta^{Nash} > \delta^{coop}$.

Intuitively, growth in the number of products creates a mass of new producers who are not yet in a good match relationship. Cooperation raises the profits that an innovator can make from supplying this type of producers (recall that $V^{s,b}_{I,coop} > V^{s,b}_{I,Nash}$), so a higher growth rate g_N makes it more likely that cooperation increases the innovation rate.

[27] Technically, to ensure efficient innovation it would still be necessary to implement a subsidy to the production of the final good in order to get rid of the existing monopoly distortion.

3.4 The type of switching costs

We now analyze the generality of Propositions 2 and 3 by discussing alternative setups. What drives our result is that if a supplier turns out to be a bad match not only is productivity lower, but so is cooperation. Therefore bad matches become relatively worse which makes switching riskier. More generally, to generate cooperation in an equilibrium where parties can change partners at will, there must be a cost of switching from one partner to another (here, the risk of finding a bad match). In many set-ups this cost interacts with incomplete contractibility to generate a lower level of cooperation at the beginning of a relationship. For instance, if we assume instead that the type of a match is only revealed after the first investment has occurred, then cooperation in the first period of a relationship would lie between the Nash level and the level in a good match. Similarly, in models where suppliers differ in their discount rate, or in models with relationship-specific human capital, the (expected) level of cooperation in a new relationship will be lower than in an established one.[28] In all these settings, relationships would be more rigid in the cooperative equilibrium than in a "Nash" equilibrium where cooperation does not take place. This excess rigidity in return can reduce the incentive to innovate, particularly when it restricts significantly the market of a potential innovator.

Nevertheless, the result that cooperation creates rigidities is not straightforward. Consider an alternative set-up without good and bad matches but with a fixed cost of switching suppliers fA. Then, provided that the fixed cost is sufficiently large, the first best investment level can be achieved in the cooperative equilibrium and the producer switches to the innovator as soon as $(\gamma - 1)\,\Pi(m) \geq f$ in both the contractible and cooperative cases, but he switches if $(\gamma - 1)\,\Pi(n) \geq f$ in the Nash case, that is for higher innovation sizes γ. In contrast with our set-up, the relative cost of switching does not increase with cooperation, and the innovation rate is always higher with cooperation.

Finally, note that even in the current set-up society could do better if suppliers were willing to collude. This could be either by refusing to cooperate with a producer who has deviated on any supplier or by outdated suppliers agreeing not to cooperate with potential producers in periods when innovation has taken place so as to encourage

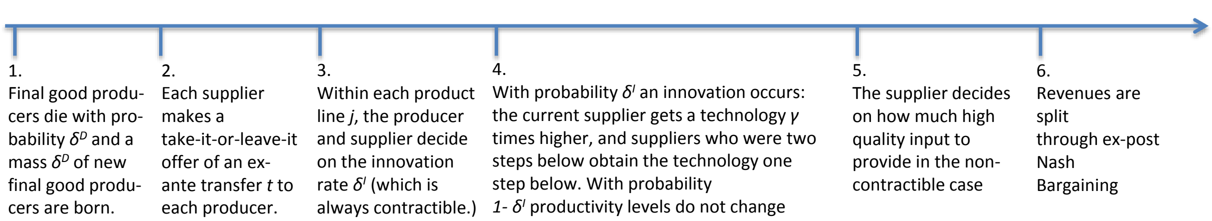
[28]This is consistent with Johnson et al. (2002) who show that the belief in the efficiency of the court matters for the level of trust between firms at the beginning of a new relationship, but much less later. Similarly, Brown, Falk and Fehr (2004), show in an experimental setting that low effort was punished by the termination of the relationship but that effort was high from the beginning in successful relationships. And Macchiavello and Morjaria (2015) found that the value of a relationship increases with its age in the Kenyan rose market export, which, they argue, provides support for their model which, as ours, features heterogeneity on the supplier side.

producers to try out the innovator. This, however, does not fit the description of a competitive industry, and is difficult to generalize in a set-up with imperfect information (for instance if suppliers do not know whether a producer knows a good match or not, whether an innovation has occurred or not).

4 Cooperation and relationship-specific innovation

We now focus on within-relationship innovation by letting the technology (denoted A_{jk}) of a supplier k be specific to the producer j with whom she is working, so that the frontier technology (denoted A_j) is producer-specific. There are no longer good or bad matches. As before, every period, a mass δ^D of producers die and are replaced by new producers. When a new producer is born, all suppliers obtain a technology level equal to the average technology in the economy to work with that producer. If the producer survives, suppliers keep the technology they had at the end of the previous period.

As before, suppliers make take-or-leave-it offers to producers and each producer chooses one supplier. A supplier can innovate with probability δ^I by spending $\psi(\delta^I)A_{jt}$ units of the final good, where ψ is increasing and convex. An innovation increases technology for line j (and only for that line) by a factor γ such that in the following period the successful innovator has a productivity advantage over the other suppliers. In such a case all other suppliers get access to the technology just below: hence suppliers can only be at most one step below the frontier for each line j. We assume throughout that the innovation rate δ is contractible—so as to focus on the consequences of incomplete contractibility in input provision. Whether an innovation occurs or not is revealed before the supplier makes her investment, which simplifies the exposition, but does not contain any element of substance. Within each period we have a timeline as follows.



4.1 Contractible and Nash cases.

We can solve for the equilibrium for each product line independently and start with the contractible and Nash cases. In both cases, when a producer is born he starts working with a supplier and remains indifferent across suppliers until one successfully innovates and becomes an *augmented* supplier. From then on, the producer picks this supplier.

Therefore, either all suppliers are identical, or one has access to a technology which is one step higher than the others. We normalize value functions and profits by the chosen supplier technology at the beginning of the period before innovation occurs. When an augmented supplier exists the joint value is:

$$V_1^T = -\psi\left(\delta_1^I\right) + \left(1 - \delta_1^I + \delta_1^I \gamma\right)\left(\Pi\left(z\right) + \frac{1-\delta^D}{1+\rho} V_1^T\right), \tag{23}$$

where δ_1^I is the equilibrium innovation rate when the producer has access to an augmented supplier, $z = m$ in the contractible case and $z = n$ in the Nash one. After the innovation cost has been paid, innovation fails with probability $1 - \delta_1^I$, in which case the producer and the supplier obtain the profit $\Pi(z)$ and the continuation value V_1^T. Alternatively, if an innovation occurs, the situation is identical except that the technology used by the supplier is γ times more productive.

The equilibrium innovation rate δ_1^I must therefore maximize V_1^T. Taking the first order condition and solving for V_1^T (using (23)) one obtains that the (unique) rate δ_1^I must obey:

$$\psi'\left(\delta_1^I\right) = (\gamma - 1)\left(\Pi\left(z\right) + \frac{1-\delta^D}{1+\rho} V_1^T\right) = (\gamma - 1)\frac{\Pi\left(z\right)(1+\rho) - \left(1-\delta^D\right)\psi\left(\delta_1^I\right)}{1+\rho - \left(1 - \delta_1^I + \delta_1^I\gamma\right)\left(1-\delta^D\right)}. \tag{24}$$

In particular, the scale effect implies that the innovation rate is lower in the Nash case than in the contractible case: $\delta_1^{I,Nash} < \delta_1^{I,cont}$

Alternatively, there is no augmented supplier in which case the joint value is:

$$V_0^T = -\psi\left(\delta_0^I\right) + \left(1 - \delta_0^I\right)\left(\Pi\left(z\right) + \frac{1-\delta^D}{1+\rho} V_0^T\right) + \delta_0^I \gamma \left(\Pi\left(z\right) + \frac{1-\delta^D}{1+\rho} V_1^T\right), \tag{25}$$

with $z = m, n$ and δ_0^I denoting the innovation rate. If no innovation occurs, then the supplier and the producer share the profits $\Pi(z)$, and in the following period the producer will be in the same situation with homogeneous suppliers who are in Bertrand competition, so that the producer will capture the full value of the relationship. On the other hand, if an innovation occurs the producer will stay with the innovating supplier. Using that δ_0^I must maximize V_0^T in (25), we obtain:

$$\psi'\left(\delta_0^I\right) = (\gamma - 1)\,\Pi\,(z) + \frac{1-\delta^D}{1+\rho}\left(\gamma V_1^T - V_0^T\right). \tag{26}$$

It is straightforward to check that $\delta_0^I = \delta_1^I = \delta^I$ and $V_0^T = V_1^T$ is a solution to the problem (see Appendix B.7 for a proof that it is the unique solution), so that the innovation rate is constant. Intuitively, whether a producer knows an augmented supplier or deals with a set of homogeneous supplier has no impact on the joint value of the relationship (beyond the technology level and the innovation rate). Yet, since innovation maximizes the joint value of the relationship, the problem is fully symmetric and the innovation rates must be the same. Moreover, we have $\delta^{I,Nash} < \delta^{I,cont}$.

4.2 Cooperative equilibrium

We build a cooperative equilibrium similarly to section 2 such that suppliers are willing to cooperate as much as possible with a producer they have never worked with before (i.e. the equilibrium must satisfy a bilateral rationality constraint). Such equilibria can only exist if the present supplier is different from other suppliers, since rents in the following period are required to reward cooperation. Here, the relationship-specific innovation takes the role of the good/bad matches from above and ensures some level of cooperation. In Appendix B.7, we demonstrate that there exists an equilibrium where on equilibrium path an augmented supplier would cooperate at a constant level $x^* \in (n, m]$, while a supplier with whom no innovation has occurred would play the Nash level of investment n. In such an equilibrium, a producer chooses a supplier when he is born. He is indifferent about switching suppliers until a supplier successfully innovates and until that happens, the suppliers invest n. Once a supplier has successfully innovated, the producer sticks with that supplier (the innovator) forever and she invests x^*.

Below, we take this structure of the equilibrium as given and derive the innovation rate on the equilibrium path. First, consider a producer who knows an augmented supplier with whom no deviation has occurred. Then, the joint value obeys (23) but with $z = x^*$. As a result, the innovation rate is given by (24). Denoting the solution for the cooperative case as $\delta_1^{I,coop}$, we get that $\delta^{I,Nash} < \delta_1^{I,coop} \leq \delta^{I,cont}$, with equality if and only if $x^* = m$, since $x^* \in (n, m]$. This directly results from the scale effect.

Second, let us focus on a producer who has never matched with a supplier who successfully innovated. We can write the joint value of their relationship as:

$$V_0^T = -\psi\left(\delta_0^I\right) + \left(1-\delta_0^I\right)\left(\Pi\left(n\right) + \frac{1-\delta^D}{1+\rho}V_0^T\right) + \delta_0^I\gamma\left(\Pi\left(x^*\right) + \frac{1-\delta^D}{1+\rho}V_1^T\right). \tag{27}$$

If no innovation occurs, the supplier does not cooperate (since she will be identical to all other suppliers next period) and the continuation value is V_0^T. If an innovation occurs, the technology improves by a factor γ but the supplier also starts cooperating (and the continuation value is V_1^T). The equilibrium innovation rate maximizes V_0^T since it is contractible. Therefore, we have:

$$\psi'\left(\delta_0^I\right) = \left(\gamma - \frac{\Pi\left(n\right)}{\Pi\left(x^*\right)}\right)\Pi\left(x^*\right) + \frac{1-\delta^D}{1+\rho}\left(\gamma - \frac{V_0^T}{V_1^T}\right)V_1^T. \tag{28}$$

Since $x^* > n$, $\Pi\left(n\right) < \Pi\left(x^*\right)$ and $V_0^T < V_1^T$. Comparing this equation with (24), with $z = x^*$, we obtain that $\delta_0^{I,coop} > \delta_1^{I,coop}$. Innovation is higher with a supplier that has not yet innovated, because in addition to pushing the technological frontier, innovation in that case also allows for starting cooperation. The innovation rate is then higher than in the Nash case, both because of the scale effect and the cooperation effect ($\delta_0^{I,coop} > \delta^{I,Nash}$), while it might be higher or lower than in the contractible case as the scale and cooperation effects push in different directions ($\delta_0^{I,coop} \lesseqgtr \delta^{I,cont}$). In particular, if the level of cooperation is low (x^* is close to n), the scale effect dominates and there is more innovation in the contractible than in the cooperative case. If the level of cooperation is high, then innovation is higher in the cooperative case than in the contractible one (in particular if the first best is achieved, $x^* = m$, then $\delta_0^{I,coop} > \delta_1^{I,coop} = \delta^{I,cont}$).

Therefore, with relationship-specific innovations, cooperation in a setting of poor contractibility strengthens innovation, up to a point that the innovation rate may even be larger than in the contractible case. The growth rate of the economy depends on the innovation rates and, in the cooperative case, on the share of firms who know an innovator and their average productivity. We obtain (proof in Appendix B.7):

Proposition 4. *The growth rate is higher in both the contractible and cooperative cases than in the Nash one. The growth rate is higher in the contractible case than in the cooperative one if cooperation is low (x^* close to n), and lower if cooperation is high (x^* close to m).*

5 Relationships, Japan and the United States: A reversal of role models

The central message of our paper is that although the existence of relational contracts can overcome contractual incompleteness, it will simultaneously affect the type of innovation undertaken. In particular, strong relationships, compared to the Nash case, will encourage relationship-specific innovations, but might discourage more general innovations that would require the break-up of such relationships. The positive effects of relationships have long been recognized. Dore (1983) first discusses the Japanese economy as a whole and argues that relational contracts within the "keiretsu" system overcome opportune behavior and allow for risk sharing. He suggests that the origins of relational contracts could be found in cultural differences. Blinder and Krueger (1996) compare US and Japanese labor markets and suggest that lower labor turn-over allows firms in Japan to invest more in training. Helper (1990) and Helper and Henderson (2014) argue that the Japanese auto-industry is a lot more productive than the American. They emphasize that the ongoing tight relationships allow for better sharing of information and fewer hold-up problems. We think of this as a higher provision of the non-contractible input, which is consistent with the 'cooperative' equilibrium. In addition, Toyota's suppliers are encouraged by the promise of continued cooperation to devote resources to innovation specifically designed for Toyota, in line with the results of section 4. This contrasts with their description of the three big automakers in the United States where the lack of trust meant that relationships had to be arm's length, contracts were met to the letter and no more, and relationship-specific investment or innovation were limited (an alternative was for the automakers to vertically integrate). Consequently, we think of the United States as being more closely represented by the 'Nash' equilibrium.[29] Similarly, Bolton, Malmrose and Ouchi (1994) argue that relational contracts in the Japanese semi-conductor industry allows for more participation by suppliers in Japan than in the United States.

Since the 1990s and following the poor economic performance of Japan, the literature has focused more on the disadvantages of the keiretsu system.[30] In a case study

[29] Here our paper bears some similarities with Acemoglu and Pischke (1998) who also develop a model with multiple equilibria and associate Germany with one and the US with the other. Interestingly, their model shows that relationships (more specifically the informational advantage that an employer has over an employee) can encourage the investment in *general* human capital, whereas we show that relationships can discourage general innovations and encourage relationship-specific innovations.

[30] Interestingly, a common quote in the economic sociology literature is from Dore (1983) considers

of a Japanese chemical company and a Japanese steel company, Collinson and Wilson (2006) argue that the keiretsu system led these companies to develop numerous but barely profitable incremental innovations tailored to the needs of their customers to the detriment of broad and flexible innovations. As mentioned in the introduction Dujarric and Hagiu (2009) argue that although Japanese prowess in efficient manufacturing is beyond question, the existence of very strong relationships leads suppliers to focus their innovation primarily on the needs of existing business partners and not new opportunities to increase market share. In addition to the software industry discussed in the introduction, they study the Japanese cellphone industry. There too, carrier providers formed closed relationships with their handset manufacturers, which produced carrier-specific phones. The industry was in fact quite innovative and several features (camera, 3G networks, payment systems, ...) were introduced in Japan before the rest of the world. Yet, handset manufacturers never managed to export their products as they were focused on developing incremental innovations specific to their carrier.[31] And they missed the radical, general, innovation of the smartphone (which included a global open platform for applications and a touchscreen). In 2008, Apple's retail share was 2%, in 2015 it reached 44% (Euromonitor International, 2015).[32]

As a result, while Japan was the role model for business relationships in the management literature in the 80s and the early 90s, the subsequent realization that Japan rarely introduces new technologies to the world market and lags behind in major innovations, has led to a reversal in the management literature which again focuses on the

the Japanese textile industry and discusses the case of the entry of a new and more efficient supplier. The response from the producer to his old supplier is given as : "Look how X has got [*sic*] his price down. We hope you can do the same because we really would have to reconsider our position if the price difference goes on for months. If you need bank finance to get the new type of vat [bucket for dyeing] we can probably help by guaranteeing the loan." This is intended as a positive feature of Japanese business relationships, whereas our paper shows the negative effects on incentives for outside innovators.

[31] Japanese cellphones were referred to as "Galapagos" phones as their kind only existed in Japan. Carrier-specific online platforms also developed and the associated Japanese firms did not manage to compete abroad (see Kushida, 2011). Kushida (2011) gives an illustration of the relationship-specific innovations and the lock-in effect: After the government implemented "number portability" which allowed consumers to keep their numbers after changing carriers, "carriers' responses were, however, to accelerate their development of proprietary features to create new lock-in effects. As the date for number portability approached, carriers engaged in a massive push towards electronic money, music players, thumbprint scans, ever high resolution cameras, and digital television broadcast receivers. As it turned out, widely used elements such as email addresses with carrier-specific domain names, data from various applications, song downloads, games, and other content were widely used, but not "portable."

[32] To be sure, innovation is not the only problem facing Japan. Hoshi and Kashyap (2004) focus on the financial sector and criticizes the willingness of the Japanese government to keep 'zombie' banks alive and with them insolvent borrowers. In Appendix B.10, we consider a model in which relationships also allow unproductive firms to remain in operation, though without a financial sector.

	Count	Mean	Median	Std. Dev	Min	Max
Generality	337,913	0.40	0.48	0.28	0	0.93
Differentiated	337,799	0.81	0.88	0.18	0.16	1

Table 1: Summary Statistics

U.S. system (Pudelko and Mendenhall, 2009). Perhaps it is no surprise that the limits of the Japanese system became apparent as they approached the world technological frontier, in line with Remark 2. The common argument behind these case studies is that the focus on incremental instead of broad innovations help explain why Japanese companies failed at becoming world leaders in certain sectors. As argued in the introduction, the associated welfare losses may be amplified by the fact that the spillovers of broad innovations are larger than that of incremental ones.

Although, Japan is a canonical example of business relationships, the tension between the dynamism of new suppliers and the reliability of old suppliers is more general. Uzzi (1996, 1997) collects quantitative and qualitative data on a set of high-end apparel producers in NYC. He shows that relying on a set of reliable suppliers is essential to overcome problems of contractual incompleteness and facilitate the transmission of information, but that only relying on existing relationships risks stifling innovation and adaptation to new trends. Both sets of suppliers are hence necessary.

To further support our discussion of Japan and the United States, we use the patent data from European Patent Office (details in Appendix C). We use patents originating in the United States and Japan, filed between 1978 and 2009. We focus on the measure of generality from Hall, Jaffe and Trajtenberg (2001) (defined in the introduction) which is the closest empirical parallel to our notion of the broad appeal of innovation: in this framework low generality corresponds to relationship-specific innovations and high generality corresponds to the broader more general innovations of Section 2. The generality measure is available for 337,913 Japanese or U.S. patents (44.5% are US patents). Table 1 shows summary statistics on the two variables of interest. In column (I) in table 2 we simply regress the generality of a patent on a dummy for the United States being the country of origin. This is an analogue of Figure 1, which shows that US patents are on average more general than Japanese. In the regression the 'generality' measure has been standardized to have mean 0 and standard deviation 1, so the estimate suggests that a US. patent is 5.5 percent of a standard deviation more general than a Japanese patent. In column (II) we introduce fixed effects for the two-digit NACE code as well as the year of filing for the patents. This isolates the difference for Japanese and US patents within

	(I)	(II)	(III)
	Generality	Generality	Generality
US.	0.055***	0.027***	0.007
	(13.47)	(6.69)	(1.18)
US x Differentiated			0.012**
			(2.09)
Fixed Effects	None	NACE, Year	NACE, Year
Observations	337,913	337,913	337,799

Standardized beta coefficients; t statistics in parentheses. Std. errors clustered at NACE x country level for (III)
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.001$

Table 2: Regression Results for 2 digit NACE

a NACE code and reduces the estimate by around half.[33] Our theory further predicts that this effect should be more pronounced in sectors that are more differentiated. To test this we associate each NACE code of the patents with a corresponding standardized "Rauch" measure of the extent to which the product is differentiated which slightly reduces the number of observations (details in Appendix C).[34] We include fixed effects for the NACE codes as well as time and year fixed effects in column (III). The coefficient of interest is the interaction term between the dummy for the United States and the measure of differentiated products, which is both substantially positive and significant (standard errors are clustered at NACE x country level). The difference in generality between a U.S and a Japanese patent increases by 1.2% of a standard deviation when moving from one sector to a sector which is one standard deviation more differentiated.

We conclude that the predictions of our model are consistent with the literature on the Japanese and U.S. patterns of innovations and are met by the empirical analysis of Japanese and U.S. patents. In Appendix C, we use a larger set of countries to perform an analogous analysis and find results consistent with our theory.

[33]Controlling for the NACE controls for the fact that the United States patents more heavily in industries with a higher average generality. This is the most direct test of our model. However, one could argue that the size of each industry is endogenous to equilibrium play, in which case the result in Column (I) would be a better test.

[34]PATSTAT attributes 2 to 4 digit NACE code to each patent depending on the sector, and the Rauch's data are available in the SITC classification. We choose to do the regressions at the 2 digit NACE level to have a uniform aggregation level across industries and because the conversion from SITC to NACE introduces some 'noise' at the 4 digit level. Nevertheless, table C.2 in the Appendix shows that our results are similar when we use the lowest level of disaggregation available.

6 Extensions

In this section, we first combine our general innovation model and our relationship-specific innovation model. Second, we extend our analysis of the general innovation model to allow for slow diffusion of innovation. Appendix A.5 makes the point that relational contracts can create macroeconomic inefficiencies in other contexts than innovation by looking at a model which features an information externality.

6.1 Combining the two models

We combine the models of sections 2-3 and section 4 into a single model (the details are in Appendix B.8). With exogenous probability, in some periods, suppliers can engage in relationship-specific innovations, while in others a potential supplier gets the opportunity to undertake a general innovation which pushes the frontier in each line by the same factor. Relationship-specific innovations can only occur in good matches. We denote by δ^A and γ^A the innovation rate and size for general innovations. We refer to a relationship between a producer and a good match supplier where the last innovation to occur was a relationship-specific innovation by the supplier as an "augmented" good match. Other good matches are referred to as "regular". δ_2^B denotes the relationship-specific innovation rate for the augmented good matches and δ_1^B for the regular good matches (x_2^*denotes the investment level in augmented good matches and x_1^* regular good matches when the supplier has access to the frontier technology). γ^B is the size of relationship-specific innovations. Although the two innovation processes interact with one another, the spirit of our analysis still applies: the cooperative equilibrium often favors relationship-specific innovations relative to the Nash equilibrium but may lead to a lower rate of general innovation.

Two subtleties complicate the analysis. First, the relationship-specific innovation rate is determined by how a relationship specific innovation changes the producer-supplier joint value, which depends on the effective discount rate. Since a higher rate of general innovation reduces the effective discount rate, it also increases the rate of relationship specific innovation. This effect makes the cooperative equilibrium look worse: if the general innovation rate is sufficiently lower in the cooperative equilibrium than in the Nash one, then the relationship specific innovation rate might also be lower in the cooperative case than in the Nash case.[35] Second, the higher is the rate of relationship

[35]Note that as in section 4, we assumed that relationship-specific innovations are contractible. Were

specific innovation, the larger is the average technology gap between lines where the producer is in a good match and those where he is not (since there is no relationship specific innovation in a bad match).[36] At the same time, for a given line, the general innovator's rents are proportional to the technology used with a coefficient which is lower in lines where producers are in good matches. Since the innovation cost is scaled by the average technology in the economy, this creates a force which pushes towards less general innovation in the equilibrium where more relationship specific innovation occurs.[37] To summarize, we can show the following proposition:

Proposition 5. *Assume that innovation costs are sufficiently small that* $\psi(\delta_2^{B,coop}) \leq \Pi(x_1^*)(\nu/(1-\nu) + 1 - \delta_2^{B,coop} + \delta_2^{B,coop}\gamma^B)$*. a) Assume that the death rate of producers is low enough* $\delta^D < \theta\frac{\Pi(n)}{\Pi(m)}$*, then for an intermediate range of innovation sizes,* $\gamma^A \in \left(\frac{1-\delta^D\frac{\Pi(m)}{\Pi(n)}}{1-\delta^D\frac{\Pi(m)}{\Pi(n)}-b(1-\theta)(1-\delta^D)}, \frac{1}{1-b(1-\delta^D)}\right)$ *the general innovation rate is higher in the Nash case than in the contractible one:* $\delta^{A,Nash} > \delta^{A,coop}$*. b) If the general innovation rate is weakly higher in the cooperative than in the Nash case* ($\delta^{A,Nash} \leq \delta^{A,coop}$) *then the rate of relationship specific innovation is also lower in the Nash case:* $\delta^{B,Nash} < \delta_2^{B,coop}, \delta_1^{B,coop}$.

Part a) of the Proposition is equivalent to Part c) of Proposition 3: the sufficient conditions under which the innovation rate in the baseline model is higher in the Nash than in the cooperative equilibrium are also sufficient conditions to ensure that the general innovation rate in the combined model is higher in the Nash than in the cooperative equilibrium. Yet, the range of parameters for which this is true is expanded in the combined model because of the second effect described above (for instance the "only if" part of Remark 1 does not hold any more). Part b) corresponds to Proposition 4, but it introduces the assumption $\delta^{A,Nash} \leq \delta^{A,coop}$ as a caveat because, as argued above, a low general innovation rate can end up hurting relationship-specific innovation.[38]

this not the case, then there would be an additional force pushing for a higher level of relationship-specific innovation in cooperative equilibrium than in the Nash one.

[36] This is by assumption but even if we were to allow for relationship specific innovation in bad matches the rate would be lower.

[37] In other words, the share of producers in a bad match ω in equation (18) is replaced by a share weighted according to the average technology level in the lines of producers in bad matches relative to the average frontier technology in the economy. The average technology level in the lines of producers in a bad match is lower because they do not benefit from relationship specific innovations.

[38] The assumption $\psi(\delta_2^{B,coop}) \leq \Pi(x_1^*)(\nu/(1-\nu) + 1 - \delta_2^{B,coop} + \delta_2^{B,coop}\gamma^B)$ is necessary because of an additional interaction between general and relationship-specific innovation: a higher rate of general innovation discourages relationship-specific innovation as the relationship-specific innovation diffuses in this case. This assumption ensures that this effect is always dominated by the discount rate effect, so that a higher rate of general innovation increases relationship-specific innovation.

As in the models of sections 3 and 4, the scale effect proportionately increases the value of any form of innovation in the cooperative equilibrium relative to the Nash equilibrium; while the other effects push toward less general and more relationship specific innovation in the cooperative than in the Nash equilibrium. Although the interaction between the two innovation processes prevents us from showing analytically that the ratio of general to relationship specific innovation is always higher in the cooperative equilibrium than in the Nash one; we can show such a result in a partial equilibrium setting: for given common future exogenous innovation rates then the relative incentive to innovate today in general innovations is higher in the Nash than in the cooperative case under mild conditions. More specifically, we show:

Remark 3. Assume that the innovation rates are exogenous and common to both equilibria, except at time 0 where they are endogenous. Assume that the normalized cost functions at time 0 are given by $\psi^A\left(\delta^A\right) = \frac{\widetilde{\psi}^A}{\psi}\left(\delta^A\right)^{\psi}$ and $\psi^B\left(\delta^B\right) = \frac{\widetilde{\psi}^B}{\psi}\left(\delta^B\right)^{\psi}$. In other periods, exogenous innovation is either free or the cost function is the same as in time 0. Then the ratio of general innovation (δ_0^A) to relationship specific innovations ($\delta_{2,0}^B$) at time 0 in augmented good match is higher in the Nash than in the cooperative case:

$$\delta_0^{A,Nash}/\delta_{2,0}^{B,Nash} > \delta_0^{A,coop}/\delta_{2,0}^{B,coop}.$$

Further provided that either the exogenous rates satisfy $\gamma^B\psi^B\left(\delta_2^B\right) \geq \psi^B\left(\delta_1^B\right)$ or that exogenous innovation is free, we get that the ratio of general innovation (δ_0^A) to relationship specific innovations ($\delta_{1,0}^B$) at time 0 in regular good match is also higher in the Nash than in the cooperative case:

$$\delta_0^{A,Nash}/\delta_{1,0}^{B,Nash} > \delta_0^{A,coop}/\delta_{1,0}^{B,coop}.$$

6.2 Slow diffusion of innovations

Here we generalize the results of the general innovation model to slower diffusion of technology. At the beginning of every period, an outdated supplier gets access to the frontier technology with probability $\Delta \in (0,1]$ if there is no innovation and catches up with the previous frontier technology if further innovation occurs. We consider a cooperative equilibrium with the same structure as in Proposition 1. In particular, an outdated supplier forgives the producer if the producer tries a frontier supplier who turns out to be a bad match (similar results would hold without such “forgiveness”, see

Appendix A.4). For simplicity, we focus our analysis on the case where after a deviation the producer would rather try a new supplier than stay with a non-cooperating good match. Finally, a producer can only keep track of one good match supplier: as soon as he meets another good match supplier, he forgets the identity of the previous good match he knew. This assumption simplifies the exposition in the contractible and Nash cases.[39]

We now describe the incentive constraints that a good match supplier faces in the cooperative case, letting the normalized value functions be V_1^s if she has access to the frontier technology and W_1^s if she has not.[40] Consider first the case in which she has access to the frontier technology at time t. Then, if she cooperates, in the following period she will enjoy $V_1^s A_t$ if there is no innovation and $W_1^s A_{t+1}$ if an innovation occurs. If she does not produce the required quantity her continuation value is 0 as the producer never comes back to a supplier after a deviation. Therefore the reward from cooperating at time t is given by $\frac{1-\delta^D}{1+\rho}((1-\delta^I)V_1^s + \delta^I \gamma W_1^s)A_t$.[41] The problem is the same as in section 2 and there is a unique level of normalized investment undertaken by a frontier good match supplier, x^*, which must satisfy the IC constraint (10).

Consider now the case of an outdated good match at time t, with level of investment y^*. In period $t+1$, this good match supplier will become a good match supplier with the frontier technology with probability Δ, otherwise she stays a good match supplier with an outdated technology. Therefore, in the cooperative equilibrium, the IC constraint for an outdated good match is given by:

$$\gamma^{-1}\varphi\left(y^*\right) \leq \frac{1-\delta^D}{1+\rho}\left(\left(1-\delta^I\right)\Delta V_1^s + \left(\left(1-\delta^I\right)(1-\Delta)+\delta^I\gamma\right)W_1^s\right). \tag{29}$$

[39] Otherwise, one would have to keep track of the number of good match suppliers that a producer knows. A producer who knows more good matches is more likely to benefit from diffusion in the future, which affects his decision to try the innovator or not. Since we focus on the case where he chooses never to work again with a non-cooperating good match in the cooperative case, this assumption only matters for the Nash and contractible cases. Moreover, making this assumption in section 2 would not affect our results, so that the model of this section is a generalization of that of section 2.

[40] As before W_1^s is positive even if the producer chooses to work with a new frontier supplier instead of an outdated good match as cooperation can resume if the new supplier turns out to be a bad match and the outdated good match benefits from imitation.

[41] Importantly this also applies to the innovator. Consider a period t where an innovation occurs, then cooperation by the innovator depends on the outside option of the producer at time $t+1$ (as this determines the value that a cooperating good match can capture). Similarly, the incentive to cooperate for any good match frontier producer at time $t+1$ depends on the producer's outside option at time $t+2$. But, at time $t+1$, a mass of firms will already have imitated the innovator, so that the producer's outside option is the same 1 or 2 periods after an innovation. Hence the problem faced by the innovator at t is identical to that faced by any cooperating frontier good match supplier.

As before, the encouragement effect pushes towards a higher level of cooperation in outdated relationships than in frontier relationships (the term γ^{-1} on the LHS of (29) pushes for $y^* \geq x^*$). Yet, for $\Delta < 1$, the RHS in (29) is also lower than the RHS in (10) since $V_1^s > W_1^s$, which pushes towards a lower level of cooperation in outdated relationships ($y^* \leq x^*$). This occurs because starting a new relationship with a frontier supplier is a more interesting outside option for a producer who is working with an outdated supplier than for one who is already working with a frontier supplier. We refer to this effect as the "outside option" effect. Overall the relationship between x^* and y^* is ambiguous and the arrival of an innovation may weaken cooperation in established relationship. Nevertheless, in Appendix B.9, we show that $\Delta \geq \left(1+\rho-b\left(1-\delta^D\right)\right) / \left(\gamma\left(1+\rho\right)-b\left(1-\delta^D\right)\right)$ is a sufficient condition to ensure that $y^* \geq x^*$.

Furthermore, in Appendix B.9, we show that producers switch to the innovator in the cooperative case if and only if

$$1-b+b\theta\frac{\Pi\left(n\right)}{\Pi\left(x^*\right)}+\left(1-\Delta\right)K\left(1-\gamma^{-1}\frac{\Pi\left(y^*\right)}{\Pi\left(x^*\right)}\right)>\gamma^{-1}\frac{\Pi\left(y^*\right)}{\Pi\left(x^*\right)}. \tag{30}$$

with $K \equiv \frac{(1-b)\left(1-\delta^D\right)\left(1-\delta^I\right)}{1+\rho-(1-\delta^D)(1-\delta^I)(1-\Delta)} > 0$. This expression is the same as (13) except for the last term on the LHS. That term captures the loss experienced by a producer who stays with an outdated good match supplier (generating profits $\gamma^{-1}\Pi\left(y^*\right)$) relative to switching to a frontier good match supplier (with profits $\Pi\left(x^*\right)$) in all periods until either the technology diffuses (which happens with probability Δ), or another innovation occurs (which happens with probability δ^I). Everything else equal, slow diffusion of innovation (a low Δ) encourages producers to switch to the innovator.

In the contractible and Nash cases, the producer switches suppliers when:

$$1-b+b\theta+\left(1-\Delta\right)K\left(1-\gamma^{-1}\right)>\gamma^{-1}. \tag{31}$$

Comparing these two expressions reveals that, as before, the ease with which a switch occurs in the cooperative compared with the contractible and Nash cases depend on the different investment levels with a frontier good match (x^*), an outdated good match (y^*) or a bad match (n). As before, the "worse bad match effect" ($x^* > n$) makes relationships more rigid in the cooperative case. In addition, if the encouragement effect dominates the outside option effect, the investment of outdated suppliers is greater than that of frontier suppliers in the cooperative equilibrium ($y^* > x^*$), which also increases the rigidity of relationships in that case. On the other hand, it is now possible that relationships could be less rigid in the cooperative case than in the contractible or cooperative case if the outside option effect is strong enough (and $y^* < x^*$).

Endogenizing the innovation rate in this set-up can be done as in section 3. As before, the reward to innovation in the cooperative case depends positively on $\Pi(x^*)$ and $\Pi(n)/\Pi(x^*)$ and negatively on $\Pi(y^*)/\Pi(x^*)$, so that the comparison of the innovation rate across the three cases depends on four effects. The scale effect pushes towards more innovation in the contractible than in the cooperative case, and towards more innovation in the cooperative than in the Nash case. The worse bad match effect pushes towards more innovation in the contractible and Nash cases than in the cooperative case. And if the encouragement effect dominates the outside option effect ($y^* > x^*$), we obtain an additional effect pushing towards less innovation in the cooperative case than in the two other cases (having on the other hand $y^* < x^*$ would push in the other direction). The following Proposition summarizes our results.

Proposition 6. *Consider parameters such that* $\Delta > \frac{1+\rho-b\left(1-\delta^D\right)}{\gamma(1+\rho)-b(1-\delta^D)}$ *and assume that* ψ *is sufficiently convex so that the equilibrium is unique, we then obtain: i) The level of investment in outdated good matches is weakly higher than in frontier matches,* $y^* \geq x^*$*. ii) For a given innovation rate, the parameter space under which relationships break in the cooperative case is a subset of the parameter set under which they break in the contractible or Nash case. iii) The innovation rate in the contractible case is larger than in the cooperative case* $\delta^{cont} > \delta^{coop}$*. iv) The innovation rate in the cooperative case may be higher or lower than in the Nash case, but if* δ^D *is small enough and parameters are such that relationships break in the Nash but not the cooperative case, then* $\delta^{coop} > \delta^{Nash}$*.*

Therefore our earlier results are generalized to this case but only if innovations diffuse sufficiently rapidly. How fast innovations diffuse depend on technological and institutional characteristics, for instance weak intellectual property rights may favor rapid technological diffusion. More generally, a slow diffusion of innovation seems to benefit the innovation rate more in the cooperative case than in the two other cases because of the outside option effect. This is illustrated in Figure 2 which shows how the three innovation rates depend on the speed of diffusion for a low value of the probability of finding a bad match $b = 0.3$ and a higher value $b = 0.6$.[42] In both cases, for fast diffusion, the innovation rate in the cooperative case is lower than both in the Nash and contractible

[42] The other parameters are the same and given by $\theta = 0.5$, $\gamma = 1.5$, $\rho = 0.05$, $\delta^D = 0.04$, $\sigma = 3$, $\beta = 0.5$ and $\psi(\delta) = \delta^2/5$.

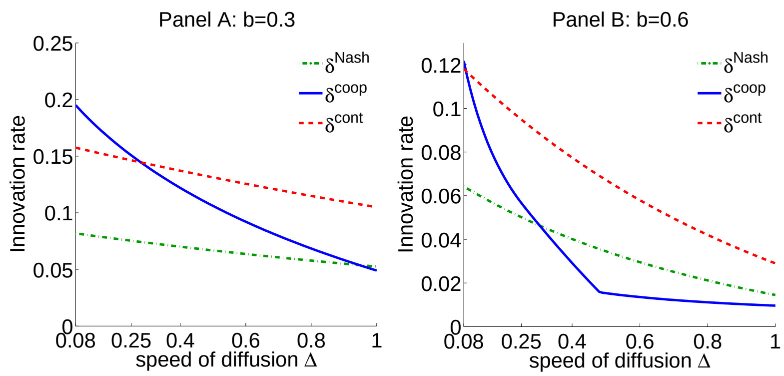


Figure 2: Innovation rate and speed of diffusion

cases. On the other hand, for slow diffusion, the innovation rate in the cooperative case is even higher than in the contractible case. The innovation rates are lower when innovation diffuses faster as fast innovation improves the outside option of producers and therefore limits the reward that an innovator can capture. A lower share of bad matches, b, reduces the importance of the worse bad match effect, which allows for a higher innovation rate in the cooperative case relative to the two other cases. Overall, our results suggest that IPR are a complement to contractual complexity and that weak contractibility is particularly damaging in sectors of weak contractibility.

7 Conclusion

In this paper, we show that the development of relational contracts shifts technological change away from broad to relationship-specific innovations. In a nutshell, our argument goes as follows: Cooperative long-term relationships, can overcome the classic underinvestment associated with the lack of contractibility. However, it is only in relationships which are a good fit—where parties understand that they are going to keep working together for a long time—that cooperation is sustainable in the first place. Consequently, switching to a new supplier becomes a riskier activity because if the new supplier is a bad fit, cooperation will not take place. More rigid relationships, in turn, slow down the process of creative destruction. On the other hand, the complementarity between cooperative behavior and relationship-specific innovations boosts the latter in a cooperative equilibrium. We relate this to the recent economic experiences of Japan and the United States. While Japan was highly praised in the 1980s and early 1990s for the level of cooperation that firms demonstrated in the keiretsu system, Japan has been less successful than the United States in introducing new technologies to the global market and the keiretsu system is now criticized for the rigidities that it has created.

An interesting extension to our analysis would be to include foreign outsourcing as issues of incomplete contractibility and long-term relationships may be even more salient when a firm is dealing with a supplier in a different country, as the firm may be less familiar with the local judicial system.

References

Acemoglu, D., Aghion, P., and Zilibotti, F. (2003). Distance to frontier, selection, and economic growth. *Journal of European Economic Association*, 4(1):37–74.

Acemoglu, D., Antràs, P., and Helpman, E. (2007). Contracts and technology adoption. *American Economic Review*, 97(3):916–943.

Acemoglu, D. and Pischke, J.-S. (1998). Why do firms train? theory and evidence. *The Quarterly Journal of Economics*, 113:79.

Aghion, P., Akcigit, U., and Howitt, P. (2015). Lessons from schumpeterian growth theory. *The American Economic Review: Papers & Proceedings*.

Aghion, P. and Bolton, P. (1987). Contracts as a barrier to entry. *The American Economic Review*, 77(3):388–401.

Akcigit, U. and Kerr, B. (2016). Growth through heterogeneous innovations. mimeo.

Allen, F., Chakrabarti, R., De, S., Qian, J., and Meijun, Q. (2008). *Emerging Giants: China and India in the World Economy*, chapter Law, Institutions and Finance in China and India.

Allen, F., Chakrabarti, R., De, S., Qian, J., and Qian, M. (2006). Financing firms in india. *World Bank Policy Research Working Paper no. 3975*.

Allen, F., Qian, J., and Qian, M. (2005). Law, finance, and economic growth in china. *Journal of Financial Economics*, 77(1):57–116.

Banerjee, A. and Duflo, E. (2000). Reputation effects and the limits of contracting: A study of the indian software industry. *The Quarterly Journal of Economics*, 115(3):989–1017.

Bessen, J. (2008). The value of u.s. patents by owner and patent characteristics. *Research Policy*, 37:932–945.

Blinder, A. and Krueger, A. B. (1996). Labor turnover in the usa and japan: A tale of two countries. *Pacific Economic Review*, 1(1):27–57.

Board, S. (2011). Relational contracts and the value of loyalty. *The American Economic Review*, 101:3349–3367.

Boehm, J. (2013). The impact of contract enforcement costs on outsourcing and aggregate productivity.

Bolton, M., Malmrose, R., and Ouchi, W. (1994). The organization of innovation in the united states and japan: neoclassical and relational contracting. *Journal of Management Studies*, 31(5):653–679.

Bonfiglioli, A. and Gancia, G. A. (2014). Growth, selection and appropriate contracts. *Review of Economic Dynamics*, 17 (1):21–38.

Bresnahan, T. and Trajtenberg, M. (1995). General purpose technologies 'engines of growth'? 65:83.

Brown, M., Falk, A., and Fehr, E. (2004). Relational contracts and the nature of market interactions. *Econometrica*, 72:747–780.

Calzolari, G. and Spagnolo, G. (2009). Relational contracts and competitive screening.

Colinson, S. and Wilson, D. (2006). Inertia in japanese organizations: Knowledge management routines and failure to innovate. *Organization Studies*, 27(9):1359–1387.

Cowan, K. and Neut, A. (2007). Intermediate goods, institutions and output per worker. *Central Bank of Chile Working Papers*, 420.

Dore, R. (1983). Goodwill and the spirit of market capitalism. *The British Journal of Sociology*, 34(4):459–482.

Dujarric, R. and Hagiu, A. (2009). Capitalizing on innovation: The case of japan. Harvard Business School Working Paper 09-114.

Euromonitor International (2015). Mobile phones in japan.

Fauli-Oller, R., Sandonis, J., and Santamaria, J. (2011). Downstream mergers and upstream investment. *The Manchester School*, 79(4):884–898.

Francois, P. and Roberts, J. (2003). Contracting productivity growth. *The Review of Economic Studies*, 70:59–85.

Ghosh, P. and Ray, D. (1996). Cooperation in community interaction without information flows. *The Review of Economic Studies*, 63:491–519.

Grossman, S. J. and Hart, O. (1986). The costs and benefits of ownership: A theory of vertical and lateral integration. *The Journal of Political Economy*, 94(4):691.

Hall, B., Jaffe, A., and Trajtenberg, M. (2001). The nber patent citation data file: Lessons, insights and methodological tools. NBER Working Paper 8498.

Hall, B. and Trajtenberg, M. (2004). Uncovering gpts with patent data. NBER wp 10901.

Helper, S. (1991). Strategy and irreversibility in supplier relations: The case of the u.s. automobile industry. *The Business History Review*, 65(4):781–824.

Helper, S. and Henderson, R. (2014). Management practices, relational contracts, and the decline of general motors. NBER working paper 19867.

Helpman, E. (1998). *General Purpose Technologies and Economic Growth*. MIT Press.

Hoshi, T. and Kashyap, A. (2004). Japan's financial crisis and economic stagnation. *The Journal of Economic Perspectives*, 18(1):3–26.

Hsieh, C.-T. and Klenow, P. (2009). Misallocation and manufacturing tfp in china and india. *Quarterly Journal of Economics*, 124 (2):771–807.

Inderst, R. and Shaffer, G. (2007). Retail mergers, buyer power and product variety. *The Economic Journal*, 117:45–67.

Inderst, R. and Wey, C. (2011). Countervailing power and dynamic efficiency. *Journal of the European Economic Association*, 9(4):702–720.

Jaffe, A., Trajtenberg, M., and Henderson, R. (1993). Geographics localizlocal of knowledge spillovers as eevidence by patent citations. 108(3):577–598.

Johnson, S., McMillan, J., and Woodruff, C. (2002). Courts and relational contracting. *Journal of Law, Economics and Organization*, 18(1)(5):221–277.

Jovanovic, B. and Rousseau, P. (2005). General purpose technologies. Chapter 18.

Kranton, R. E. (1996). Reciprocal exchange: A self-sustaining system. *The American Economic Review*, 86(4):830–851.

Kushida, K. (2011). Leading without followers: How politics and market dynamics trapped innovations in japan's domestic galapagos telecommunications sector. *Journal of Industry, Competition and Trade*, 11:279–307.

Macaulay, S. (1963). Non-contractual relations in business: A preliminary study. *American Sociological Review*, 28(1):55.

Macchiavello, R. and Morjaria, A. (2014). Competition and relational contracts: Evidence from rwanda's coffee mills.

Macchiavello, R. and Morjaria, A. (2015). The value of relationships: evidence from a supply shock to kenyan rose exports. *American Economic Review*, 105(9):2911–2945.

MacLeod, W. B. and Malcomson, J. M. (1989). Implicit contracts, incentive compatibility, and involuntary unemployment. *Econometrica*, 57(2):447–480.

Nunn, N. (2007). Relationship-specificity, incomplete contracts, and the pattern of trade. *Quarterly Journal of Economics*, 122(2):569–600.

OECD (2015). Oecd, indicators database and regpat database.

Pudelko, M. and Mendenhall, M. (2009). The contingent nature of best practices in national competitiveness: The case of american and japanese innovation processes. *European Management Journal*, 27:456–466.

Rauch, J. E. (1999). Networks versus markets in international trade. *Journal of International Economics*, 48(1):7–35.

Thompson, P. and Fox-Kean, M. (2005). Patent citations and the grography of knowledge spillovers: A reassessment. 95(1):450–460.

Trajtenberg, M., Henderson, R., and Jaffe, A. (1997). Univeristy versus corporate ppatent: A window on the basicness of invention. 5:19–50.

Uzzi, B. (1996). The sources and consequences of embeddedness for the economic performance of organizations: the network effect. *American Sociological Review*, 61(4):674–698.

Uzzi, B. (1997). Social structure and competition in interfirm networks: the paradox of embeddedness. *Administrative Science Quarterly*, 42(1):35–67.

A Main Appendix

A.1 Proof of Proposition 1

This appendix proves Proposition 1 in the case where a producer prefers working with a new supplier over a non-cooperating good match in a period without innovation and with a new outdated supplier over a non-cooperating good match in a period with innovation. Since the incentive compatibility constraints of the supplier are satisfied, since the agents revert to the one shot Nash strategy after a deviation, and since ex-ante transfers are determined through Bertrand competition, it is direct that if the levels $x^*$and y^* exist then the strategies described in Proposition 1 lead to a SPNE. Proving the existence of $x^*$and y^* requires first showing that in all possible scenarii, on path or off path, there are only two possible forms for the IC constraint of the supplier depending on whether she has access to the frontier technology or not. Second, we need to show that these IC constraints admit a solution with $x^*, y^* > n$.

The proof proceeds in 4 steps: first we derive the condition under which a producer in a good match tries out the innovator—equation (13) in the text. Second, we derive the general form of the IC constraint. Third, we derive detailed expressions for the two possible IC constraints in function of $x^*$and y^*—in this appendix we do it only when a producer prefers working with a new supplier over a non-cooperating good match in a period without innovation and with a new outdated supplier over a non-cooperating good match in a period with innovation, the other cases are included in Appendix B.1. Fourth, we show that there exist $x^*, y^* > n$, satisfying the IC constraint under all possible cases—in the same special case here and in general in Appendix B.1.

A.1.1 Step 1. Condition under which a producer in a good match switches to the innovator (equation (13))

We consider a producer who knows a good match supplier with whom no deviation has occurred and we study whether the producer would want to switch to the innovator or not.[43] We use the notations V_i^z and W_i^z, with $i \in \{0,1\}$ and $z \in \{s,p,T\}$ defined in the text. Furthermore, in periods with an innovation and for a relationship with the innovator, we denote by $V_I^{z,t}$ the value of the producer ($z = p$), or the supplier/innovator

[43]This analysis always applies on path. Off path it also applies except when the producer already knows a non-cooperating good match and the value of a relationship with the innovator is lower than the value of staying with this non-cooperating good match. That case is treated in Appendix B.1.

$(z = s)$, knowing that previously the producer was in a good match who did not deviate $(t = g)$, or in a bad match $(t = b)$. As a supplier forgives a producer who switches to the innovator if the innovator turns out to be a bad match, the continuation value of a good match supplier who is not chosen by the producer in a period with innovation does not fall to 0 as the producer may come back to her if the innovator turns out to be a bad match. We denote the expected value of such an (outdated) supplier by V_A^s.

The innovator and the old supplier enter in Bertrand competition, the old supplier would be willing to offer a transfer that would guarantee herself at least V_A^s in order to keep the producer, hence SPNE requires that:

$$W_1^s \geq V_A^s. \tag{A.1}$$

Moreover Bertrand Competition ensures that the supplier with whom the relationship is the highest captures the entire benefit of the relationship over the second best one, hence the value of the producer whether he switches supplier or not is the same:

$$V_I^{p,g} = W_1^p. \tag{A.2}$$

The producer ends up switching if the highest amount that the innovator can offer is higher than the highest amount that the old supplier can offer, that is if if the total value of the producer and the innovator $(V_I^{T,g})$ is higher than the surplus value of the old relationship $(W_1^T - V_A^s)$.[44]

$$V_I^{T,g} > W_1^T - V_A^s. \tag{A.3}$$

The total value of a relationship with the innovator is given by:

$$\begin{aligned} V_I^{T,g} = {} & (1-b)\,\Pi\left(x^*\right) + (1-b)\,\frac{1-\delta^D}{1+\rho}\left(\left(1-\delta^I\right)V_1^T + \delta^I\gamma W_1^T\right) \\ & + b\theta\Pi\left(n\right) + b\frac{1-\delta^D}{1+\rho}\left(\left(1-\delta^I\right)V_1^p + \delta^I\gamma W_1^p\right). \end{aligned} \tag{A.4}$$

With probability $1-b$ the relationship turns out to be good delivering profits $\Pi\left(x^*\right)$ in the first period and with continuation value V_1^T if no innovation occurs and W_1^T if innovation occurs. With probability b, the relationship turns out to be a bad match, the

[44]Technically this is derived under the condition that the value a good match old supplier is willing to offer is (weakly) higher than the value another outdated supplier would be willing to offer, when the innovator is actually the best choice (otherwise it is obvious since an innovator necessarily offers more than a new outdated supplier). We show in step 3 that this is necessarily true.

continuation value for the supplier is then zero, and the producer goes back to his old good match supplier, so that his value is V_1^p if no innovation occurs and W_1^p otherwise.

This leaves us with the expected value to the supplier from the possibility that the producer returns, V_A^s as the only missing element. If the producer switches, the current profits enjoyed by the old supplier are zero, but with probability b, the innovator will turn out to be a bad match, in which case cooperation will resume, the old supplier will get V_1^s if no innovation occurs and W_1^s otherwise, hence:

$$V_A^s = \frac{1-\delta^D}{1+\rho} b \left(\left(1-\delta^I\right) V_1^s + \delta^I \gamma W_1^s \right). \tag{A.5}$$

Now combining (5), (A.4) and (A.5) one gets:

$$V_I^{T,g} - \left(W_1^T - V_A^s\right) = (1-b)\,\Pi\left(x^*\right) + b\theta\Pi\left(n\right) - \frac{1}{\gamma}\Pi\left(y^*\right), \tag{A.6}$$

which show that a good match producer switches to the innovator provided that equation (13) holds.

A.1.2 Step 2. The general form of the incentive constraint

As argued in the text, the gain a supplier would get by deviating from the agreed level of investment is given by $\varphi(x)A_k$ with φ defined in (9). Should a deviation occurred, the continuation value of the supplier may not always be 0 as in the case studied in the text. Therefore, in general the incentive constraints obey:

$$\varphi\left(x^*\right) \leq \frac{1-\delta^D}{1+\rho} I \text{ and } \gamma^{-1}\varphi\left(y^*\right) \leq \frac{1-\delta^D}{1+\rho} I, \tag{A.7}$$

where we define the effect of cooperation on the continuation value of the supplier:

$$I \equiv \left(1-\delta^I\right) V_1^s + \delta^I \gamma W_1^s - \left(\left(1-\delta^I\right) V_N^s + \delta^I \gamma W_N^s \right). \tag{A.8}$$

V_N^s and W_N^s are the value the supplier would get if she becomes a non-cooperating good match (and investment would then be given by the Nash level), in periods where, respectively, there is not and there is innovation. If the supplier cooperates, her value in the following period is given by V_1^s if there is no innovation and W_1^s otherwise. The factor γ^{-1} on the LHS of the second IC constraint comes from the fact that the technology of the outdated supplier is only $\gamma^{-1}A$.

Combining (A.1), (A.2) and (A.6), we get:[45]

$$W_1^s = V_A^s + \left(\frac{1}{\gamma}\Pi\left(y^*\right) - \left(\left(1-b\right)\Pi\left(x^*\right) + b\theta\Pi\left(n\right)\right)\right)^+, \tag{A.9}$$

where $X^+ \equiv \max\{X, 0\}$.

Using equation (A.5) and (A.9) we get:

$$\begin{aligned}&\left(1-\delta^I\right)V_1^s + \delta^I\gamma W_1^s \qquad\qquad \text{(A.10)}\\ =&\frac{1+\rho}{1+\rho-b\left(1-\delta^D\right)\delta^I\gamma}\left(\left(1-\delta^I\right)V_1^s + \delta^I\gamma\left(\frac{1}{\gamma}\Pi\left(y^*\right) - \left(\left(1-b\right)\Pi\left(x^*\right) + b\theta\Pi\left(n\right)\right)\right)^+\right).\end{aligned}$$

Finally note that V_1^T must satisfy (4) which combined with (5) leads to:

$$V_1^T = \frac{\left(1+\rho-\left(1-\delta^D\right)\delta^I\gamma\right)\Pi\left(x^*\right) + \left(1-\delta^D\right)\delta^I\Pi\left(y^*\right)}{1+\rho-\left(1-\delta^D\right)\left(1-\delta^I+\delta^I\gamma\right)}. \tag{A.11}$$

If the producer does not already know a non-cooperating good match, we necessarily get through Bertrand competition:

$$V_1^p = V_0^{T,n} \text{ and } V_1^s = V_1^T - V_0^{T,n}, \tag{A.12}$$

where $V_0^{T,n}$ is the value of starting a new relationship when the producer knows a non-cooperating good match (this is the general expression, see footnote 16). Indeed, the outside option for the producer is to start a new relationship, but should he do so, he would now know a non-cooperating good match, namely the good match he was previously working with. If the producer knows a non-cooperating good match, then his second best option will either be to resume a relationship with the non-cooperating good match or to start a new relationship, now knowing two non-cooperating good match suppliers, so that we get, through Bertrand Competition:

$$V_1^{s,n} = V_1^{T,n} - \max\left(V_N^T, V_0^{T,n}\right), \tag{A.13}$$

where V_N^T denotes the joint value of a relationship with the non-cooperating good match.

As mentioned in the text, depending on parameters, there is a number of different

[45]That still requires that switching to the innovator is a better option than switching to a potential non-cooperating good match when the producer knows one (see footnote 43).

cases to consider. In order to save space we will consider only the case where in case of a deviation the producer always seeks out a new producer. The other cases are considered in Appendix B.1. The results of the paper hold in all cases.

A.1.3 Step 3 in a special case: When a deviation always leads the producer to try out a different supplier

Assume that in periods without innovation, the producer would always rather try out a new supplier than a non-cooperating good match, and, in periods with innovation, the producer would prefer both the innovator or an outdated new supplier to an (outdated) non-cooperating good match. That is, we assume:

$$V_N^T < V_0^T \text{ and } W_N^T < W_0^T. \tag{A.14}$$

and we need not index V_0^T and W_0^T by n as whether a producer knows a non-cooperating good match or not is now irrelevant. As the producer will never return to a non-cooperating good match, the continuation value of a non-cooperating good match (with that producer) is 0: $\left(1-\delta^I\right)V_N^s + \delta^I\gamma W_N^s = 0$. In (A.8), we can therefore focus on $\left(1-\delta^I\right)V_1^s + \delta^I\gamma W_1^s$ which is given by equation (A.10). (A.12) implies that in this case (8) holds. Therefore the incentive to cooperate is directly related to the value a good match supplier captures in periods without innovation (V_1^s). In addition, whenever the profits generated by an outdated good match supplier exceed the expected profits with the innovator, the difference contributes to the value of the outdated supplier and therefore to her incentive to cooperate.

(6) and (7) imply that the joint value W_0^T obeys:

$$W_0^T = V_0^T - (1-b)\left(\Pi(x^*) - \gamma^{-1}\Pi\left(y^*\right)\right) - b\theta\left(1-\gamma^{-1}\right)\Pi\left(n\right). \tag{A.15}$$

This equation, together with (4) and (8) determine V_1^s as a function of x^*, y^* and n:

$$V_1^s = \frac{b\left(\left(1+\rho-b\left(1-\delta^D\right)\delta^I\gamma\right)\left(\Pi\left(x^*\right)-\theta\Pi\left(n\right)\right)+b\left(1-\delta^D\right)\delta^I\left(\Pi\left(y^*\right)-\theta\Pi\left(n\right)\right)\right)}{1+\rho-b\left(1-\delta^D\right)\left(1-\delta^I+\delta^I\gamma\right)}. \tag{A.16}$$

Therefore V_1^s corresponds to the appropriately discounted and weighted sum between the difference in profits between a good match and bad match in periods without innovation $\left(\Pi\left(x^*\right)-\theta\Pi\left(n\right)\right)$ and in periods with innovation $\left(\gamma^{-1}\left(\Pi\left(y^*\right)-\theta\Pi\left(n\right)\right)\right)$. The factor b in front of the fraction reflects that a new supplier is a bad match with probability b. Even for x^*, y^* arbitrarily close to n, V_1^s is positive as a good match supplier can capture the rents associated with having revealed her type.

Combining (A.16) and (A.10), we find

$$I = \frac{1+\rho}{1+\rho-b\left(1-\delta^D\right)\delta^I\gamma}\left(\begin{array}{c}\left(1-\delta^I\right)\frac{b\left(\left(1+\rho-b\left(1-\delta^D\right)\delta^I\gamma\right)\left(\Pi(x^*)-\theta\Pi(n)\right)+b\left(1-\delta^D\right)\delta^I\left(\Pi(y^*)-\theta\Pi(n)\right)\right)}{1+\rho-b\left(1-\delta^D\right)\left(1-\delta^I+\delta^I\gamma\right)} \\ +\delta^I\gamma\left(\frac{1}{\gamma}\Pi\left(y^*\right)-\left(\left(1-b\right)\Pi\left(x^*\right)+b\theta\Pi\left(n\right)\right)\right)^+\end{array}\right). \tag{A.17}$$

This establishes the IC constraints together with (10) and (11) in the main text and determines the equilibrium investment levels x^* and y^*.

Further, we had to check that in a period with innovation, when the producer switches to the innovator, staying with a previous outdated good match supplier is still a better outside option than trying a new outdated supplier (this is not obvious since the good match supplier only offers $W_1^T - V_A^s$ to the producer). That is we need to check that $W_1^T - V_A^s > W_0^T$. Combining (7) with (5), (A.5) and using (A.12) we obtain:

$$W_1^T - V_A^s - W_0^T = \frac{b}{\gamma}\left(\Pi\left(y^*\right)-\theta\Pi\left(n\right)\right)+\frac{1-\delta^D}{1+\rho}b\delta^I\gamma\left(W_1^T - V_A^s - W_0^T\right),$$

which shows that $W_1^T - V_A^s - W_0^T > 0$.

A.1.4 Step 4: Existence of a solution for x^*, y^* in the same special case

Here we show that should the economy be in the case described above, then there is a solution $x^*, y^* > n$ to the problem. To do that we simply need to show that the IC constraints do not bind for (x, y) just above n. Because n minimizes φ, we have

$$\varphi\left(x\right) = o\left(x-n\right) \text{ and } \gamma^{-1}\varphi\left(y\right) = o\left(y-n\right).$$

Therefore, we simply have to check that I is positive at the first order in $(x-n)$ and $(y-n)$ when x and y are greater than n. Using (A.17), we get:

$$\begin{aligned} I = {} & \frac{1+\rho}{1+\rho-b\left(1-\delta^D\right)\delta^I\gamma}\left(1-\delta^I\right)\frac{b\left(1+\rho-b\left(1-\delta^D\right)\delta^I\left(\gamma-1\right)\right)\left(1-\theta\right)}{1+\rho-b\left(1-\delta^D\right)\left(1-\delta^I+\delta^I\gamma\right)}\Pi\left(n\right)+ \\ & + \frac{1+\rho}{1+\rho-b\left(1-\delta^D\right)\delta^I\gamma}\delta^I\gamma\left(\frac{1}{\gamma}-\left(1-b+b\theta\right)\right)^+\Pi\left(n\right), \end{aligned}$$

which is positive at first order in $(x - n)$, $(y - n)$. This proves existence provided that the conditions to be in this case are met (see Appendix B.1 for the rest of the proof).

A.2 Cooperative equilibrium characterization

In this appendix we provide a set of conditions on the equilibrium strategies that imply that the agents must play according to Proposition 1. We denote by $H_t^n(j,k)$ the set of histories of the game after t repetitions just after phase 5 has occurred (just after the type has been revealed) when producer j and supplier k are matched for the first time and supplier k has turned out to be a good match. We define a *symmetry and information* condition:

Condition 1. ***Symmetry and Information (SI)*** *i) For any history belonging to* $\bigcup_k H_t^n(j,k)$ *where the supplier* k *has access to the frontier technology, the path of normalized investment undertaken in the following histories by the new supplier* k *are the same, and the decision of the producer to continue the relationship with the supplier* k *or not is the same; similarly for any history belonging to* $\bigcup_k H_t^n(j,k)$ *where the supplier* k *does not have access to the frontier technology; ii) the strategies played with one producer are independent of the history of the game played with other producers; iii) if a supplier has been chosen by the producer, her normalized investment is independent of the ex-ante transfer paid by the supplier.*

Part i) is a symmetry condition. Provided that the supplier has access to the frontier technology, every new good match relationship is identical in terms of the level of normalized investment and of the producer' decision to retain the supplier or not (both on and off the equilibrium path). In particular, if a producer starts a relationship with the innovator and the innovator turns out to be a good match, the outcome is symmetric to the case where the producer started his first relationship. We cannot however require that the strategies are identical, because, in general, the ex-ante transfer exchanged depends on whether the producer knows a good match supplier or not. This condition rules out equilibria where there is never cooperation with the innovator even if she is a good match—without this condition it would be possible to build equilibria where the path of investment levels is systematically lower with a new supplier than with the first supplier. Part ii) allows us to keep the strategies with other producers independent, so, for instance, producers cannot coordinate on punishing a supplier. Part iii) is necessary to ensure that the supplier gets the full value of the relationship when the first best is achieved. Otherwise it is possible to build equilibria where part of the surplus of a relationship would go to the producer, despite Bertrand competition. It should be clear that conditions i) and ii) avoid equilibria where players could coordinate their actions on histories that should have no direct impact on their interactions. Such restrictions would necessarily operate in an alternative environment where we directly restricted the information available to the players. Condition iii) does not affect our results but simplifies the exposition.

As described in the text, we define a *forgiveness* condition which ensures that a supplier does not punish a producer who switched to the innovator if the innovator turns out to be a bad match.

Condition 2. ***Forgiveness***. *The strategy played by a good match supplier at time t, is the same when the producer has worked with the supplier at time $t-1$ and when the producer has worked with an innovator but the innovator turned out to be a bad match.*

Denoting respectively by $V^{p,j}(\sigma)$ and $V^{s,k}(\sigma)$ the values of producer j and supplier k, when the profile of strategy is σ, we formally define the bilateral rationality condition as follows.

Condition 3. ***Bilateral rationality***. *At any history $h_t \in H_t^n(j,k)$, $\sigma|h_t$ is such that there is no $\sigma' = \left(\sigma'_j|h_t, \sigma'_k|h_t, \sigma_{-k}|h_t\right)$ (where σ_{-k} denotes the profile of the other suppliers) where $\sigma'_j|h'_t = \sigma_j|h'_j$ for all histories $h'_t \in H_t^n(j,k')$ $(k \neq k')$, σ'* satisfies condition 2*, and neither player j nor player k have an incentive to deviate from σ', such that $V^{p,j}(\sigma') + V^{s,k}(\sigma') > V^{p,j}(\sigma) + V^{s,k}(\sigma)$.*

Bilateral rationality here means that a new pair chooses strategies that maximize their joint value under the condition that the strategy of the producer with a new good match is given (the producer is expected to renegotiate his strategies once he has found a new good match), strategies are enforceable (neither the producer nor the supplier have an incentive to deviate), and the forgiveness condition is not violated. This condition rules out "collusive" behavior by suppliers: in a good match, suppliers are willing to cooperate as much as possible right away.[46] Finally, we impose:

Condition 4. ***No investment in bad matches***. *Normalized investment levels in bad matches are given by the Nash investment level, n.*

[46]This condition should not be confused with a "renegotiation-proof" condition. If one of the players deviates from the prescribed strategies a punishment phase is allowed even if it yields lower profits.

If the productivity level θ is sufficiently low, this condition is automatically met as a producer would continue to search for a new supplier regardless of whether cooperation in bad matches is possible or not. We then obtain the following Proposition:

Proposition 7. *In any symmetric SPNE satisfying conditions 1-4, agents' strategies are given as in Proposition 1.*

Proof. It is direct to check that the strategies of Proposition 1 obey conditions 1-4. Appendix B.2 shows that conditions 1-4 imply the strategies of Proposition 1. □

A.3 Level of cooperation

In this section we study how the levels of investment in the cooperative equilibrium depend on the model's parameters. We restrict attention to the case where the innovation rate is exogenous. We obtain the following proposition and remark, which are proved in Appendix B.3.

Proposition 8. *(i) The investment levels (x^*, y^*) weakly increase with the number of bad matches, b, and decrease with the relative productivity of bad matches, θ, the discount rate, ρ, and the probability of death δ^D; (ii) when the innovator captures the entire market, the investment levels (x^*, y^*) increase in the size of innovations γ.*

Remark 4. *When a producer would always rather try a new supplier than work with a non-cooperative good match supplier, and the innovator captures the entire market, the investment levels (x^*, y^*) decrease with the rate of innovation δ^I provided that innovations are not too large ($\gamma b \left(1 - \delta^D\right) \left(2 - \delta^I\right) < 1 + \rho$ is a sufficient condition).*

How much suppliers cooperate depends on how bad the alternative option is. Therefore if the probability of a bad match, b is higher, or if they are more severe (low θ), a good relationship will have more value, and the potential for cooperation is higher. A higher value of the future (lower ρ and δ^D) have the same effect. This follows directly from (A.16) and (A.10) in the specific case where a producer does not work again with a good match supplier who has stopped cooperating. Furthermore, we get that when the innovator captures the entire market ($\gamma > \gamma^{coop}$), large innovations favor cooperation. The reason is that larger innovations lead to a higher growth rate, which increases the expected value a supplier can capture by cooperating, favoring more investment in good matches. If the innovator does not capture the entire market then larger innovations also reduce the value a good match supplier can capture in periods with innovation.

Finally, the effect of the rate of innovation is in general ambiguous, even when the innovator captures the entire market. More frequent innovations will have three effects on investment levels: (i) a positive effect through a higher growth rate, (ii) a negative effect through a higher probability of ending the relationship, and (iii) a further negative effect which reflects that the benefit of being in a good match over a random match is higher in periods without innovation (and this benefit is precisely what drives the incentive to cooperate). For sufficiently small innovations, effect (ii) dominates effect (i), so that more frequent innovations lower the level of cooperation. We can compare this result to Francois and Roberts (2003), who show that an increase in innovation can push firms towards providing short-term contract arrangements instead of implicit guarantees of lifetime employment to their workers. In our model, the same idea is captured by the possible decrease in cooperation following an increase in the innovation rate.

A.4 Alternative equilibrium where suppliers systematically punish producers who switch to the innovator

In this appendix, we describe an *alternative cooperative equilibrium* where the supplier always refuses to reengage in cooperation if the producer switches to the innovator. That is the strategy of the supplier described in Proposition 1 is modified such that a cooperating good match becomes a non-cooperating good match as soon as a producer switches to the innovator (regardless of the innovator's type). For the sake of simplicity, we focus on parameters value for which a producer would rather switch supplier than stay with a non cooperative good match. We also assume that when innovators decide on how much to invest, they are unaware of when the last innovation occurred.[47] We prove the following proposition (where the innovation rate in the alternative cooperative case refers to the highest equilibrium level) in Appendix B.6.

Proposition 9. *(i) The parameter set for which innovators capture the whole market in the alternative cooperative case is strictly smaller than the parameter set for which innovators capture the whole market in the contractible or the Nash cases; in particular, the minimum technological leap required for an innovator to capture the whole market in the alternative cooperative case (γ^{coop2}) is higher than that in the contractible or Nash cases ($\gamma^{con}, \gamma^{Nash}$): $\gamma^{coop2} > \gamma^{con} = \gamma^{Nash}$. (ii) For ρ small enough*

[47]Otherwise there would not be a steady-state because the share of producers who are not in an ongoing good match relationship depends on when the last innovation occurred if innovations are large enough.

($\rho < (\gamma/\delta^{coop2} - 1)\left(1 - b\left(1 - \delta^D\right)\right) + b\left(1 - \delta^D\right)\delta^{coop2}(\gamma - 1)$ is a sufficient condition), the innovation rate in the alternative cooperative case is lower than in the contractible case. (iii) The innovation rate in the alternative cooperative case may be lower or higher than in the Nash case, it is lower if $\gamma \in (\gamma^{cont}, \gamma^{coop2})$ and δ^D is sufficiently small.

This proposition stipulates that our results carry through in this alternative equilibrium. This is not surprising and in some sense the results are reinforced. Indeed, if a producer switches to the innovator, and the innovator turns out to be a bad match, the producer would have to suffer additional losses in the periods following innovation as he would have to keep looking for a good match, since the previous one would have stopped cooperating. This *loss of cooperation effect* pushes towards more rigid relationships in the cooperative case than in the contractible or Nash cases. In Appendix B.6, we show that producers would switch to the innovator if and only if

$$(1-b) + b\theta\frac{\Pi(n)}{\Pi(x^*)} - b\frac{1-\delta^D}{1+\rho}\frac{\left(1-\delta^I\right)\left(V_1^T - V_0^T\right) + \delta^I\gamma\left(W_1^T - W_0^T\right)}{\Pi(x^*)} > \gamma^{-1}\frac{\Pi(y^*)}{\Pi(x^*)}, \tag{A.18}$$

The third term in (A.18) (which is absent in (13)) reflects the loss of cooperation effect. It is equal to the loss in expected profits which occurs if the innovator turns out to be a bad match and the producer has to look for a new supplier in the subsequent periods, scaled by the profits in a good match at the frontier ($\Pi(x^*)$). This loss corresponds to the difference in the joint value of a relationship with a good match compared to a new relationship, namely $V_1^T - V_0^T$ in periods without innovation and $W_1^T - W_0^T$ in periods with an innovation. Therefore, Proposition (2) carries through.

The scale effect still pushes towards more innovation in the cooperative case than in the Nash case, but towards less innovation than in the contractible case. The encouragement effect, the worse bad match effect and now the loss of cooperation effect, by making relationships more rigid, push towards less innovation in the cooperative case than in both the Nash and contractible cases. There is however a counteracting *general equilibrium effect*: when innovations are sufficiently large to break up existing relationships ($\gamma > \gamma^{coop2}$), there will be more producers not in an ongoing good match relationship in the cooperative than in both the contractible and Nash cases.[48] As an innovator captures more value from producers who are not in an ongoing good match relationship,

[48] In the cooperative case, the share of producers previously not in a good match is given by $\frac{\delta^D + b\delta^I\left(1-\delta^D\right)}{1-b(1-\delta^D)(1-\delta^I)}$ when $\gamma > \gamma^{coop}$ but by $\frac{\delta^D}{1-b(1-\delta^D)}$ when $\gamma < \gamma^{coop}$ or in the Nash or contractible cases.

this force pushes towards more innovation in the cooperative than in the Nash but also contractible cases. As a lower discount rate strengthens the loss-of-cooperation effect, the *general equilibrium effect* is dominated for a sufficiently low discount rate ρ, which explains Part ii) of Proposition 9.[49] As before if $\gamma \in \left(\gamma^{Nash}, \gamma^{coop2}\right)$, the innovator breaks relationships in the Nash case but not in the cooperative case, this implies that if the death rate of producers δ^D is sufficiently small, the innovator gets a much smaller market so that the innovation rate is lower in the cooperative case than in the Nash (Part iii) of Proposition 9).[50]

Loss of good matches in the contractible and Nash cases. Alternatively, it may be that even in the contractible or Nash cases, a producer cannot resume working with a supplier after the relationship was halted, either because the two parties suffer a utility loss, or because the producer forgets the identity of a good match once he has stopped working with her. Under this scenario, switching to an innovator involves losing a good match supplier also for the contractible and Nash cases. Nevertheless, our results carry through: the parameter space for which a switch occurs is smaller in the cooperative case than in the contractible or Nash cases; the innovation rate is lower in the cooperative case than in the contractible case; and it is also lower than in the Nash case for an intermediate range of innovation sizes provided that the death rate of producers is low enough.[51]

A.5 Rigidity in relationships and information externalities

Though we have used an endogenous growth model, the point that relationships can be detrimental to welfare can be made in other contexts. Instead of the externalities associated with the endogenous growth model (imitation and standing-on-the-shoulders-

[49] The condition in the proposition 9 will be satisfied for reasonable parameter values since $\gamma > \gamma^{coop2} > (1-b+b\theta)^{-1}$ is necessary for the general equilibrium effect to exist, and δ^{coop2} is small.

[50] The proposition focused on the cooperative equilibrium with the highest innovation rate. Yet, in the cooperative case, the expected share of producers who are not with a good match supplier (ω) increases with the innovation rate for $\gamma > \gamma^{coop2}$, so that there is significant room for multiple equilibria. For instance, there could be an equilibrium where innovation is scarce, so that most producers have found a good match supplier and cooperation is widespread, and another equilibrium, where innovation is frequent and cooperation is rare.

[51] We obtain that $\gamma^{Nash} = \gamma^{cont} = \left[1-b+b\theta - \frac{\left(1-\delta^D\right)b^2(1-\theta)}{1+\rho-b(1-\delta^D)(1-\delta^I+\delta^I\gamma)}\right]^{-1}$. Hence the worse bad match and the encouragement effects make the loss of a good match supplier relatively more costly in the cooperative than in the Nash or contractible cases. In addition, for a given rate of innovation, the share of producers who do not know a good match at the beginning of a period in steady-state is the same in all cases, so that the general equilibrium effect described above ceases to play a role.

of-giants) we consider here an information externality: firms are more likely to choose a supplier who is already active. Therefore a producer who decides to keep a supplier who has suffered a negative productivity shock because of their ongoing relationship exerts a negative externality on other producers. This externality is needed for relationships to reduce welfare.

As before, a producer needs to pick a supplier to produce and the match can be either good or bad. For simplicity we set $\delta^D = 0$ such that all producers are infinitely-lived and therefore know a good match supplier. Contrary to section 2, we now assume that there is no growth in productivity. Instead a supplier's productivity A_k is drawn each period and takes three values with equal probability: 1, γ and γ^2. Productivity draws are independent. The reason for three values will become apparent below. We formalize the information externality as follows. Suppliers cannot make take-it or leave-it offers to all producers. Instead producers must choose between a limited set of suppliers in a staggered fashion. At the beginning of the period a share λ of producers can costlessly choose one additional potential supplier. They do not yet have any information on the productivity shocks of suppliers and will choose one at random. The potential supplier and the previous good match then make take-it or leave-it offers to the producer who decides with whom to work.[52] The remaining $1-\lambda$ producers observe these choices—but not any productivity shocks—before choosing their potential supplier. They also receive take-it or leave-it offers from the potential new supplier and the previous good match supplier before choosing a supplier. Since the choice of the first λ producers on whether to continue operation with a supplier contains information on the productivity shock of this supplier, we label it an "information" externality. More generally, this is meant to capture that for a variety of reasons—search costs, reputation benefits etc.—firms are more likely to choose business partners already in operation.

In the cooperative case, we consider an equilibrium which is similar to that described in section 2. In particular, there is no cooperation in bad matches and there are 3 levels of cooperation in good matches (x_0, x_1 and x_2 depending on the technology level A_k), but for simplicity we consider parameters such that $x_0 = x_1 = x_2 = m$ in the text. Further, cooperation between a good match supplier and a producer ceases if either the supplier deviated on her investment level, the producer switched to a supplier with a weakly worse technology or the producer switched to a supplier with a better technology and that supplier turned out to be a good match. As a result, in all cases (cooperative, Nash

[52]At this stage the productivity of each supplier becomes known by the alternative supplier, the previous supplier and the producer.

and contractible) a producer chooses to switch supplier if and only if current expected profits are higher with the new supplier than with the previous good match.

A producer keeps his good match supplier if she has a higher productivity than the alternative supplier. In the Nash or contractible case, he switches to a supplier with a technology that is γ times more productive than the existing one if and only if $\gamma(1-b+b\theta) > 1$ (as in (12)). Similarly, he switches to an alternative supplier with a technology γ^2 times more productive if and only if $\gamma^2(1-b+b\theta) > 1$. And for reasons analogous to (13), in the cooperative case, he switches to a supplier with a technology γ times more productive if $\gamma((1-b)\Pi(m) + b\theta\Pi(n)) > \Pi(m)$ and to one with a technology γ^2 times more productive if $\gamma^2((1-b)\Pi(m) + b\theta\Pi(n)) > \Pi(m)$. As before cooperation creates rigidity for intermediate values of γ: if $\gamma \in (\gamma^{Nash}, \gamma^{coop})$, a producer switches supplier if her previous one does not have the higher technology in the Nash but not in the cooperative case.

Assume that $\gamma \in (\gamma^{Nash}, \gamma^{coop})$ but $\gamma^2 > \gamma^{coop}$: a producer switches supplier if and only if that supplier has a technology at least 1 step ahead in the Nash case but 2 steps ahead in the cooperative case. The first round of producers choose their alternative supplier at random who are therefore equally likely to have productivities 1, γ and γ^2.

Now, consider the remaining $(1-\lambda)$ producers. If they choose among suppliers randomly they have an equal probability of meeting a supplier with probability 1, γ and γ^2. However, as derived in Appendix B.10 if they choose their potential suppliers among those that are already in production, the distribution of productivity will be 1 with probability 1/9, γ with probability 1/3 and γ^2 with probability 5/9 in the Nash or contractible cases. Since their alternative supplier has already been judged a better option by another producer, their odds are better than for the first group.

By comparison, in the cooperative case, the alternative supplier's productivity in the second round is distributed as follows: 1 with probability 2/9, γ with probability 1/3 and γ^2 with probability 4/9. As the first round producers prefer to stick to a supplier who has a technology one step below that of their alternative supplier, the average productivity of suppliers who secure a market during the first round in the cooperative case is worse than in the Nash or contractible cases. This information externality reduces the appeal of the cooperative equilibrium relative to the Nash case. More specifically, we demonstrate in Appendix B.10 the following proposition.

Proposition 10. *i) If all producers are in the first group ($\lambda = 1$), welfare is always higher in the cooperative than in the Nash case. ii) Otherwise, welfare may be lower in* *the cooperative case than in the Nash case; in particular this happens when* λ *is close to* 0*, cooperation achieves the first best in good matches,* $\gamma^{Nash} < \gamma < \gamma^{coop} < \gamma^2$ *and the level of investment in the Nash case is sufficiently high. iii) Welfare is the highest in the contractible case.*

Absent the information externality, cooperation necessarily increases welfare despite the additional rigidity. This is because producers choose the supplier who maximizes their expected profits, which maximizes aggregate profits. In addition, producers are less likely to switch in the cooperative case, which increases the average expected level of investment from suppliers. Since from a welfare stand-point, investment is too low because of the standard monopoly distortion, it must be the case that welfare is higher in the cooperative than in the Nash case.

On the other hand, the interaction between the information externality and the excess rigidity of relationships in the cooperative case reduces welfare in the cooperative case. Although, it requires somewhat specific parameter combinations, this effect can be sufficiently strong to make cooperation welfare reducing.[53] This is a general lesson of the paper: as long as producers choose their suppliers efficiently from the point of view of the expected profits in their line, an externality (here the information externality, earlier the imitation and standing on the shoulders of giants externality) is necessary to make rigid relationships potentially welfare reducing.

[53]This does not rest on the monopoly distortion, as cooperation also reduces aggregate profits. Moreover, one can increase the parameter space for which cooperation is welfare reducing, for instance by adding a share of one-period lived producers who get to pick their supplier after the first round of long-lived producers do so.

B Online Appendix

B.1 Cooperative equilibrium characterization: complements

In this appendix we complete the proof of Proposition 7 in Appendix A.1 by going through steps 3 and 4 in all possible cases and showing that they cover the full range of possibilities.

B.1.1 Step 3: deriving the IC constraint in all cases.

Whether a producer would rather stick to a non-cooperating good match (a good match playing the Nash level because a deviation has occurred) or keep looking for a new supplier will affect the IC constraint. We derive it in all the possible cases:

- case 1, when the producer will choose the non-cooperating good match in any circumstances,

- case 2, when the producer will choose the innovator over the non-cooperating good match, but stick to the non-cooperating good match otherwise,

- case 3, when the producer will choose the non-cooperating good match in period without innovation, but in period with innovation the non-cooperating good match is worse than even an outdated supplier,

- case 4, when the producer will choose a new match in periods without innovation, but in period with innovation, the non-cooperating good match is better than trying an outdated supplier,

- case 5, when the non-cooperating good match is never one of the two best options (which is the special case studied in Appendix A.1).

Moreover, in cases 1, 2 and 3, the non-cooperating good match could be chosen one step away from the equilibrium path, we then need to check that whether the producer knows only one non-cooperating good match or more matters.

Case 1.1 We consider the case where the non-cooperating good match is always better than starting a new relationship. We consider a producer who only knows one non-cooperating good match (and no other good match), we derive the conditions under which this case applies, and the incentive constraint of a producer who would be in a good match relationship and would not know any non-cooperating good match. We still denote by V_N^T the joint value of the producer and the non-cooperating good match in periods without innovation and we define W_N^T as the corresponding value in periods with innovation. We then get:

$$V_N^T = \Pi(n) + \frac{1-\delta^D}{1+\rho}\left(\left(1-\delta^I\right)V_N^T + \delta^I \gamma W_N^T\right), \tag{B.1}$$

$$W_N^T = \gamma^{-1}\Pi(n) + \frac{1-\delta^D}{1+\rho}\left(\left(1-\delta^I\right)V_N^T + \delta^I \gamma W_N^T\right). \tag{B.2}$$

Now recall that $V_0^{T,n}$ denotes the joint value when a producer starts a new relationship (n indicates that the producer knows a non-cooperating good match) in periods without innovation, we denote by $V_I^{T,n}$ the same value in periods with innovation, and we get:

$$V_I^{T,n} = V_0^{T,n} = (1-b)V_1^T + b\theta\Pi(n) + b\frac{1-\delta^D}{1+\rho}\left(\left(1-\delta^I\right)V_N^p + \delta^I \gamma W_N^p\right), \tag{B.3}$$

with probability $(1-b)$ the new supplier is a good match and the joint value becomes V_1^T, with probability b the new supplier is a bad match, in which case the producer should revert back to the non-cooperating good match in the following period. Bertrand competition ensures that:

$$V_N^p = V_0^{T,n} \text{ and } V_N^s = V_N^T - V_0^{T,n}, \tag{B.4}$$

$$W_N^p = V_I^{T,n} \text{ and } W_N^s = W_N^T - V_I^{T,n}. \tag{B.5}$$

The condition to be in that case is that in periods with innovation the producer would rather stick to the non-cooperating good match than choose the innovator, note that if the producer chooses the innovator, the value of the non-cooperating good match is not null, instead it is given by:

$$V_{AN}^s = \frac{1-\delta^D}{1+\rho}b\left(\left(1-\delta^I\right)V_N^s + \delta^I \gamma W_N^s\right), \tag{B.6}$$

as with probability b the innovator will be a bad match and the producer would revert back to the non-cooperating good match in the following period. The condition to be in that case can then be expressed as:

$$W_N^T \geq V_0^T + V_{AN}^s. \tag{B.7}$$

Combining (B.1) and (B.2), we get:

$$W_N^T = V_N^T - \left(1 - \gamma^{-1}\right) \Pi(n), \tag{B.8}$$

so that:

$$V_N^T = \frac{1 + \rho - \left(1 - \delta^D\right) \delta^I (\gamma - 1)}{1 + \rho - (1 - \delta^D)(1 - \delta^I + \delta^I \gamma)} \Pi(n). \tag{B.9}$$

Combining (B.3), (B.6) and (B.8), we can rewrite (B.7) as:

$$V_N^T - \left(1 - \gamma^{-1}\right) \Pi(n) \geq (1 - b) V_1^T + b\theta\Pi(n) + b\frac{1 - \delta^D}{1 + \rho} \left(\left(1 - \delta^I\right) V_N^T + \delta^I \gamma W_N^T\right),$$

which using (B.9) translates into:

$$\begin{aligned} &\frac{\left(\gamma^{-1}(1 + \rho) - b\left(1 - \delta^D\right) + \left(1 - \delta^D\right)\left(1 - \delta^I\right)(1 - \gamma^{-1})\right)\Pi(n)}{(1 + \rho)(1 + \rho - (1 - \delta^D)(1 - \delta^I + \delta^I \gamma))} \\ &\geq \left((1 - b) V_1^T + b\theta\Pi(n)\right). \end{aligned} \tag{B.10}$$

Now we want to express the IC constraint of a producer in a good match who does not know any non-cooperating good match. To do so, we first need to compute the expected value of a non-cooperating good match. Combining (B.3), (B.4) and (B.5) we get:

$$V_I^{T,n} = V_0^{T,n} = (1 - b) V_1^T + b\theta\Pi(n) + b\frac{1 - \delta^D}{1 + \rho} \left(1 - \delta^I + \delta^I \gamma\right) V_0^{T,n},$$

so that:

$$V_I^{T,n} = V_0^{T,n} = \frac{1 + \rho}{(1 + \rho - b(1 - \delta^D)(1 - \delta^I + \delta^I \gamma))} \left((1 - b) V_1^T + b\theta\Pi(n)\right), \tag{B.11}$$

which combined with (B.9) gives:

$$\begin{aligned} V_N^s = &\frac{1 + \rho - \left(1 - \delta^D\right) \delta^I (\gamma - 1)}{1 + \rho - (1 - \delta^D)(1 - \delta^I + \delta^I \gamma)} \Pi(n) \\ &- \frac{1 + \rho}{(1 + \rho - b(1 - \delta^D)(1 - \delta^I + \delta^I \gamma))} \left((1 - b) V_1^{T\prime} + b\theta\Pi(n)\right), \end{aligned}$$

$$W_N^s = \frac{\gamma^{-1}(1+\rho) + (1-\gamma^{-1})\left(1-\delta^D\right)\left(1-\delta^I\right)}{1+\rho-(1-\delta^D)(1-\delta^I+\delta^I\gamma)}\Pi(n)$$
$$-\frac{1+\rho}{(1+\rho-b(1-\delta^D)(1-\delta^I+\delta^I\gamma))}\left((1-b)V_1^{T\prime}+b\theta\Pi(n)\right).$$

Therefore we can write:

$$\begin{aligned}&\left(1-\delta^I\right)V_N^s+\delta^I\gamma W_N^s\\&=\frac{1+\rho}{1+\rho-(1-\delta^D)(1-\delta^I+\delta^I\gamma)}\Pi(n)\\&-\frac{(1+\rho)\left(1-\delta^I+\delta^I\gamma\right)}{(1+\rho-b(1-\delta^D)(1-\delta^I+\delta^I\gamma))}\left((1-b)V_1^T+b\theta\Pi(n)\right)\end{aligned}\tag{B.12}$$

Using (A.12) and (B.11) we get:

$$V_1^s = V_1^T - \frac{1+\rho}{(1+\rho-b(1-\delta^D)(1-\delta^I+\delta^I\gamma))}\left((1-b)V_1^T+b\theta\Pi(n)\right).\tag{B.13}$$

Combining (A.8), (A.10), (B.12), (B.13) and (A.11), and knowing that if the non-cooperating good match is better than the innovator, then a good match supplier is also necessarily better than the innovator, we get:

$$I = \frac{(1+\rho)\left(\left(1-\delta^I\right)\Pi(x^*)+\delta^I\Pi(y^*)-\Pi(n)\right)}{1+\rho-(1-\delta^D)(1+\delta^I(\gamma-1))}.$$

Case 1.2 We consider now the same situation except that the producer already knows at least two non-cooperating good match suppliers. Note that Bertrand competition implies that the producer can then capture the entire value of the relationship so that:

$$V_N^s = W_N^s = V_{AN}^s = 0,\ V_N^p = V_N^T \text{ and } W_N^p = W_N^T\tag{B.14}$$

(B.1), (B.2) and therefore (B.9) still hold. However (B.3) combined with (B.14) now gives:

$$V_0^{T,n} = V_I^{T,n} = (1-b)V_1^T + b\theta\Pi(n) + b\frac{1-\delta^D}{1+\rho}\left(\left(1-\delta^I\right)V_N^T+\delta^I\gamma W_N^T\right),\tag{B.15}$$

and the condition to be in that case is now

$$W_N^T \geq V_0^{T,n},$$

instead of (B.7) (as the value of the non-cooperating good match is always null). This condition then leads to:

$$\frac{\left(\gamma^{-1}(1+\rho)-b\left(1-\delta^{D}\right)+\left(1-\delta^{D}\right)\left(1-\delta^{I}\right)\left(1-\gamma^{-1}\right)\right)\Pi(n)}{(1+\rho)\left(1+\rho-\left(1-\delta^{D}\right)\left(1-\delta^{I}+\delta^{I}\gamma\right)\right)} \geq \left((1-b)V_{1}^{T}+b\theta\Pi(n)\right).$$

which is the same condition as in case 1.

(A.13) gives:

$$V_1^s = V_1^T - V_N^T.$$

Now note that we are precisely in the case where the analysis leading to (A.9) may not apply. If the producer would rather switch to the non-cooperating good match than the innovator, we get that:

$$W_1^p = W_N^T \text{ and } W_1^s = W_1^T - W_N^T,$$

so that we get:

$$I = \frac{(1+\rho)\left(\left(1-\delta^{I}\right)\Pi(x^*)+\delta^{I}\Pi(y^*)-\Pi(n)\right)}{1+\rho-\left(1-\delta^{D}\right)\left(1+\delta^{I}(\gamma-1)\right)}.$$

The incentive constraint for a supplier with a producer who knows a non-cooperating good match is the same as in case 1, which is a necessary requirement for the existence of the equilibrium (because condition 1 requires that the profile of investment with a good match supplier is always the same even if the good match supplier is not the first good match supplier).

We need however to check that switching to the innovator when one is in a good match remains worse than switching to the non-cooperating good match (this is conceptually not equivalent as saying that a producer with a non-cooperating good match would not switch to the innovator, indeed for a producer in a good match, switching to the innovator does not necessarily lead to punishment in the following period, whereas switching to the non-cooperating good match does so[54]). We prove this by contradiction, assume that a producer in a good match would rather deviate by switching to the innovator than to

[54]We will not have to worry about this in the following cases because there the non-cooperating good match will be worse than the innovator for a producer not in a good match (by assumption), which therefore implies that the non-cooperating good match will be worse than the innovator also for a producer in a good match.

the non-cooperating good match. We would then get:

$$W_1^p = V_I^{T,g} = (1-b)\, V_1^T + b\theta\Pi\,(n) + \frac{b\left(1-\delta^D\right)}{1+\rho}\left(\left(1-\delta^I\right) V_N^T + \delta^I\gamma V_I^{T,g}\right),$$

with the condition $V_I^{T,g} > W_N^T$, however this condition leads to the reverse of condition (B.10).

Case 2.1 We now consider again a producer who knows only a single non-cooperating good match. We assume that sticking to the non-cooperating good match is preferred to trying out a new supplier in periods without innovation, or an outdated supplier in periods with innovation, but remains worse than switching to the innovator in periods with innovation. In other words, we assume:

$$V_N^T \geq V_0^{T,n} \text{ and } W_0^{T,n} \leq W_N^T - V_{AN}^s \leq V_I^{T,n}. \tag{B.16}$$

Bertrand competition leads to (B.4) and to:

$$W_N^s = V_{AN}^s \text{ and } W_N^p = V_I^{p,n} = W_N^T - V_{AN}^s, \tag{B.17}$$

the value of the non-cooperating good match in periods with innovation is not null because if the innovator turns out to be a bad match, the producer would come back to the non-cooperating good match in the following periods.

(B.1), (B.2) (and therefore (B.8) and (B.9)), (B.3) and (B.6) still hold. (B.6), (B.17) and (B.4) give:

$$V_{AN}^s = \frac{b\left(1-\delta^D\right)\left(1-\delta^I\right)}{1+\rho-\left(1-\delta^D\right) b\delta^I\gamma}\left(V_N^T - V_0^{T,n}\right). \tag{B.18}$$

Moreover, we get (using (B.8), (B.17) and (B.18)):

$$W_N^p = V_I^{p,n} = \frac{1+\rho-b\left(1-\delta^D\right)\left(1-\delta^I+\delta^I\gamma\right)}{1+\rho-\left(1-\delta^D\right) b\delta^I\gamma} V_N^T - \left(1-\gamma^{-1}\right)\Pi\,(n) + \frac{b\left(1-\delta^D\right)\left(1-\delta^I\right)}{1+\rho-\left(1-\delta^D\right) b\delta^I\gamma} V_0^{T,n}.$$

Plugging this into (B.3) leads to:

$$
\begin{aligned}
V_0^{T,n} = {} & \frac{\left(1+\rho-\left(1-\delta^D\right)b\delta^I\gamma\right)(1-b)V_1^T}{1+\rho-b\left(1-\delta^D\right)\left(1-\delta^I+\delta^I\gamma\right)} + b\frac{1-\delta^D}{1+\rho}\delta^I\gamma V_N^T \\
& + \frac{1+\rho-\left(1-\delta^D\right)b\delta^I\gamma}{1+\rho-b\left(1-\delta^D\right)\left(1-\delta^I+\delta^I\gamma\right)}b\theta\Pi(n) \\
& - \frac{1+\rho-\left(1-\delta^D\right)b\delta^I\gamma}{1+\rho-b\left(1-\delta^D\right)\left(1-\delta^I+\delta^I\gamma\right)}b\frac{1-\delta^D}{1+\rho}\delta^I\gamma\left(1-\gamma^{-1}\right)\Pi(n).
\end{aligned}
$$

Using (B.9), we further get:

$$
\begin{aligned}
V_0^{T,n} = {} & \frac{\left(1+\rho-\left(1-\delta^D\right)b\delta^I\gamma\right)\left((1-b)V_1^T+b\theta\Pi(n)\right)}{1+\rho-b\left(1-\delta^D\right)\left(1-\delta^I+\delta^I\gamma\right)} \\
& + b\frac{1-\delta^D}{1+\rho}\delta^I\gamma \\
& \times\left(\frac{1+\rho-\left(1-\delta^D\right)\delta^I\gamma\left(1-\gamma^{-1}\right)}{1+\rho-\left(1-\delta^D\right)\left(1-\delta^I+\delta^I\gamma\right)} - \frac{1+\rho-\left(1-\delta^D\right)b\delta^I\gamma}{1+\rho-b\left(1-\delta^D\right)\left(1-\delta^I+\delta^I\gamma\right)}\left(1-\gamma^{-1}\right)\right)\Pi(n).
\end{aligned}
\tag{B.19}
$$

We can then express the conditions of (B.16) as:

$$
\begin{aligned}
& \frac{\gamma^{-1}(1+\rho)-b\left(1-\delta^D\right)+\left(1-\delta^D\right)\left(1-\delta^I\right)\left(1-\gamma^{-1}\right)}{1+\rho-\left(1-\delta^D\right)\left(1-\delta^I+\delta^I\gamma\right)}\Pi(n) \\
& \leq (1-b)V_1^T+b\theta\Pi(n) \\
& \leq \frac{\left(1+\rho-b\left(1-\delta^D\right)-\left(1-\delta^D\right)\delta^I(\gamma-1)\right)\Pi(n)}{1+\rho-\left(1-\delta^D\right)\left(1-\delta^I+\delta^I\gamma\right)} \\
& - \left(\left(1-\gamma^{-1}\right)(1-b\theta)\Pi(n)-(1-b)\left(\Pi(x^*)-\gamma^{-1}\Pi(y^*)\right)\right)^+.
\end{aligned}
\tag{B.20}
$$

We now move on to compute the incentive constraint of a producer in a good match. Using (B.17) and (B.6), we get:

$$
\left(1-\delta^I\right)V_N^s+\delta^I\gamma W_N^s = \frac{(1+\rho)\left(1-\delta^I\right)V_N^s}{1+\rho-\left(1-\delta^D\right)b\delta^I\gamma}. \tag{B.21}
$$

In this case (A.10) and (A.12) apply, combining them with (A.11), (B.4) and (B.21) we can express the reward from cooperation I as:

$$I = \frac{(1+\rho)\left(\left(1-\delta^I\right)V_1^s + \delta^I\gamma\left(\frac{1}{\gamma}\Pi\left(y^*\right) - \left(\left(1-b\right)\Pi\left(x^*\right) + b\theta\Pi\left(n\right)\right)\right)^+\right)}{1+\rho-b\left(1-\delta^D\right)\delta^I\gamma} - \frac{(1+\rho)\left(1-\delta^I\right)V_N^s}{1+\rho-\left(1-\delta^D\right)b\delta^I\gamma} \tag{B.22}$$

$$= \frac{(1+\rho)\left(\left(1-\delta^I\right)\left(V_1^T - V_N^T\right) + \delta^I\gamma\left(\frac{1}{\gamma}\Pi\left(y^*\right) - \left(\left(1-b\right)\Pi\left(x^*\right) + b\theta\Pi\left(n\right)\right)\right)^+\right)}{1+\rho-b\left(1-\delta^D\right)\delta^I\gamma}$$

$$= \frac{1+\rho}{1+\rho-b\left(1-\delta^D\right)\delta^I\gamma}\left(\begin{array}{c}\frac{\left(1-\delta^I\right)\left(\left(1+\rho-\left(1-\delta^D\right)\delta^I\gamma\right)\Pi(x^*)+\left(1-\delta^D\right)\delta^I\Pi(y^*)-\left(1+\rho-\left(1-\delta^D\right)\delta^I\gamma\left(1-\gamma^{-1}\right)\right)\Pi(n)\right)}{1+\rho-(1-\delta^D)(1-\delta^I+\delta^I\gamma)} \\ +\delta^I\gamma\left(\frac{1}{\gamma}\Pi\left(y^*\right) - \left(\left(1-b\right)\Pi\left(x^*\right)+b\theta\Pi\left(n\right)\right)\right)^+\end{array}\right).$$

Recall that we needed to check that when a producer in a good match switches to the innovator, sticking to the old supplier remains better than trying out a new outdated supplier, that is we need to check that $W_1^p = W_1^T - V_A^s > W_0^{T,n}$. Using that $V_1^p = V_n^p = V_0^{T,n}$, we get:

$$\begin{aligned}
W_1^p &= \frac{1}{\gamma}\Pi\left(y^*\right) + \frac{1-\delta^D}{1+\rho}\left((1-b)\left(\left(1-\delta^I\right)V_1^T + \delta^I\gamma W_1^T\right) + b\left(\left(1-\delta^I\right)V_0^{T,n} + \delta^I\gamma W_1^p\right)\right) \\
W_0^{T,n} &= \frac{1-b}{\gamma}\Pi\left(y^*\right) + \frac{b\theta\Pi\left(n\right)}{\gamma} \\
&\quad + \frac{1-\delta^D}{1+\rho}\left((1-b)\left(\left(1-\delta^I\right)V_1^T + \delta^I\gamma W_1^T\right) + b\left(\left(1-\delta^I\right)V_0^{T,n} + \delta^I\gamma W_0^{T,n}\right)\right)
\end{aligned}$$

so the inequality is satisfied.

Case 2.2 As for case 1 2, we now consider the same situation as in case 2.1 except that the producer knows two non-cooperating good match suppliers. To ensure the existence of the equilibrium we need that the conditions to be in case 2.2 are the same as the conditions to be in case 2.1, and that the IC constraint that we derive here (the IC constraint for a producer in a good match who knows a non-cooperating good match) is the same as the incentive constraint derived in case 2..1. (B.1), (B.2) (and therefore (B.8) and (B.9)) still hold. Bertrand competition now leads to (B.14), so that (B.3) gives (B.15) as in case 1 2. The conditions to be in that case now writes as

$$V_N^T \geq V_0^{T,n} \text{ and } W_0^{T,n} \leq W_N^T \leq V_I^{T,n}, \tag{B.23}$$

as $V_{AN}^s = 0$. Using equations (B.15), (B.8) and (B.9) we get that these conditions are equivalent to (B.20) as it should.

For the IC constraint, (A.13) gives $V_1^s = V_1^T - V_N^T$, (A.10) still holds, so using (B.14) we directly get that I is given by (B.22) as it should.

Note that we need to check that when the producer does not switch to the innovator, sticking to a good match supplier remains a better option than going for the innovator, that is we need to check that $W_1^p = W_1^T - V_A^s$ remains greater than W_N^T this is direct because:

$$W_1^p = \frac{1}{\gamma}\Pi\left(y^*\right) + \frac{1-\delta^D}{1+\rho}\left((1-b)\left(\left(1-\delta^I\right)V_1^T + \delta^I\gamma W_1^T\right) + b\left(\left(1-\delta^I\right)V_N^T + \delta^I\gamma W_1^p\right)\right)$$
$$W_N^T = \frac{1}{\gamma}\Pi\left(y^*\right) + \frac{1-\delta^D}{1+\rho}\left(\left(1-\delta^I\right)V_N^T + \delta^I\gamma W_N^T\right),$$

and $\left(1-\delta^I\right)V_1^T + \delta^I\gamma W_1^T > \left(1-\delta^I\right)V_N^T + \delta^I\gamma W_1^p$.

Case 3.1 We now consider the case where the producer would rather stick to the non-cooperating good match in periods without innovation, but, in periods with innovations, the non-cooperating good match is worse than even a new outdated supplier. We consider a producer who only knows one non-cooperating good match. The conditions to be in that case can then be expressed as:

$$V_N^T \geq V_0^{T,n} \text{ and } \left(W_N^T - V_{AN}^s\right) \leq W_0^{T,n}.$$

As a consequence, the value of a producer in a period without innovation when he knows a non-cooperating good match is given by

$$V_I^{p,n} = W_0^{T,n}. \tag{B.24}$$

We get that (B.1) must be replaced by:

$$V_N^T = \Pi\left(n\right) + \frac{1-\delta^D}{1+\rho}\left(\left(1-\delta^I\right)V_N^T + \delta^I\gamma\left(W_0^{T,n} + V_{AN}^s\right)\right), \tag{B.25}$$

in a period with innovation, the value of the non-cooperating good match supplier is indeed not null and given by V_{AN}^s, where V_{AN}^s is given by

$$V_{AN}^s = \frac{1-\delta^D}{1+\rho} b\left(\left(1-\delta^I\right) V_N^s + \delta^I \gamma V_{AN}^s\right). \tag{B.26}$$

Note that (B.8) still holds. (B.3) is replaced by:

$$V_0^{T,n} = (1-b) V_1^T + b\theta\Pi(n) + b\frac{1-\delta^D}{1+\rho}\left(\left(1-\delta^I\right) V_N^p + \delta^I \gamma V_I^{p,n}\right), \tag{B.27}$$

while the value of starting a relationship with an outdated supplier is given by:

$$W_0^{T,n} = V_0^{T,n} - (1-b)\left(\Pi(x^*) - \gamma^{-1}\Pi(y)\right) - b\theta\left(1-\gamma^{-1}\right)\Pi(n). \tag{B.28}$$

Bertrand competition still leads to (B.4), which, together with (B.24), (B.27) and (B.28) gives:

$$\begin{aligned} V_0^{T,n} = {} & \frac{(1+\rho)\left((1-b) V_1^T + b\theta\Pi(n)\right)}{1+\rho-b\left(1-\delta^D\right)\left(1-\delta^I+\delta^I\gamma\right)} - \frac{\delta^I\gamma b\left(1-\delta^D\right)(1-b)\left(\Pi(x^*)-\gamma^{-1}\Pi(y^*)\right)}{1+\rho-b\left(1-\delta^D\right)\left(1-\delta^I+\delta^I\gamma\right)} \\ & - \frac{b\left(1-\delta^D\right)\delta^I\gamma b\theta\left(1-\gamma^{-1}\right)\Pi(n)}{1+\rho-b\left(1-\delta^D\right)\left(1-\delta^I+\delta^I\gamma\right)}. \end{aligned} \tag{B.29}$$

Now (B.4) and (B.26) give:

$$V_{AN}^s = \frac{b\left(1-\delta^D\right)\left(1-\delta^I\right)}{1+\rho-\left(1-\delta^D\right) b\delta^I\gamma}\left(V_N^T - V_0^{T,n}\right). \tag{B.30}$$

Combining (B.25), (B.28), (B.29) and (B.30), we get:

$$\begin{aligned} V_N^T = {} & \frac{1+\rho-\left(1-\delta^D\right) b\delta^I\gamma}{1+\rho-\left(1-\delta^I\right)\left(1-\delta^D\right)-\left(1-\delta^D\right) b\delta^I\gamma}\Pi(n) + \frac{\left(1-\delta^D\right)\delta^I\gamma\left((1-b) V_1^T + b\theta\Pi(n)\right)}{1+\rho-\left(1-\delta^I\right)\left(1-\delta^D\right)-\left(1-\delta^D\right) b\delta^I\gamma} \\ & - \frac{\left(1-\delta^D\right)\delta^I\gamma(1-b)\left(\Pi(x^*)-\gamma^{-1}\Pi(y^*)\right)}{1+\rho-\left(1-\delta^I\right)\left(1-\delta^D\right)-\left(1-\delta^D\right) b\delta^I\gamma} - \frac{\left(1-\delta^D\right)\delta^I\gamma b\theta\left(1-\gamma^{-1}\right)\Pi(n)}{1+\rho-\left(1-\delta^I\right)\left(1-\delta^D\right)-\left(1-\delta^D\right) b\delta^I\gamma}. \end{aligned} \tag{B.31}$$

The condition $V_N^T \geq V_0^{T,n}$ and $\left(W_N^T - V_{AN}^s\right) \leq W_0^{T,n}$ then translate into:

$$\begin{aligned}
&\frac{\left(1+\rho-b\left(1-\delta^D\right)-\left(1-\delta^D\right)\delta^I\left(\gamma-1\right)\right)\Pi\left(n\right)}{1+\rho-\left(1-\delta^D\right)\left(1-\delta^I+\delta^I\gamma\right)} \\
&-\left(\left(1-b\theta\right)\left(1-\gamma^{-1}\right)\Pi\left(n\right)-\left(1-b\right)\left(\Pi\left(x^*\right)-\gamma^{-1}\Pi\left(y^*\right)\right)\right) \\
&\leq \left(1-b\right)V_1^T+b\theta\Pi\left(n\right) \\
&\leq \frac{1+\rho-b\left(1-\delta^D\right)-\left(1-\delta^D\right)\delta^I\left(\gamma-1\right)}{1+\rho-\left(1-\delta^D\right)\left(1-\delta^I+\delta^I\gamma\right)}\Pi\left(n\right) \\
&+\left(\frac{\left(1-b\right)\delta^I\gamma\left(1-\delta^D\right)}{1+\rho-\left(1-\delta^D\right)\left(1-\delta^I+\delta^I\gamma\right)}\right)\left(\left(1-b\theta\right)\left(1-\gamma^{-1}\right)\Pi\left(n\right)-\left(1-b\right)\left(\Pi\left(x^*\right)-\gamma^{-1}\Pi\left(y^*\right)\right)\right),
\end{aligned} \tag{B.32}$$

this case exists only when $\left(1-b\theta\right)\left(1-\gamma^{-1}\right)\Pi\left(n\right)-\left(1-b\right)\left(\Pi\left(x^*\right)-\gamma^{-1}\Pi\left(y^*\right)\right) \geq 0$.

We now move to express the IC constraint. First note that (B.4) and (B.30) lead to

$$\left(1-\delta^I\right)V_N^s+\delta^I\gamma V_{AN}^s=\frac{\left(1+\rho\right)\left(1-\delta^I\right)}{1+\rho-\left(1-\delta^D\right)b\delta^I\gamma}\left(V_N^T-V_0^{T,n}\right).$$

Combining this with (A.10), (A.12), (A.11) and (B.31), we get:

$$I=\frac{1+\rho}{1+\rho-b\left(1-\delta^D\right)\delta^I\gamma}\left(\begin{array}{c}\frac{\left(1-\delta^I\right)\left(\left(1+\rho-b\left(1-\delta^D\right)\delta^I\gamma\right)\left(\Pi\left(x^*\right)-\Pi\left(n\right)\right)+b\left(1-\delta^D\right)\delta^I\left(\Pi\left(y^*\right)-\theta\Pi\left(n\right)\right)\right)}{\left(1+\rho-\left(1-\delta^I\right)\left(1-\delta^D\right)-\left(1-\delta^D\right)b\delta^I\gamma\right)} \\ +\delta^I\gamma\left(\frac{1}{\gamma}\Pi\left(y^*\right)-\left(\left(1-b\right)\Pi\left(x^*\right)+b\theta\Pi\left(n\right)\right)\right)\end{array}\right). \tag{B.33}$$

Checking that trying a new outdated supplier is worse than staying with a good match supplier for a producer when innovation occurs proceeds as in case 2.

Case 3.2 As before we redo this case assuming that there are several non-cooperating good match suppliers. The condition now writes as:

$$V_N^T > V_0^{T,n} \text{ and } W_N^T < W_0^{T,n}. \tag{B.34}$$

Bertrand competition leads to

$$V_N^s = W_N^s = V_{AN}^s = 0,\ V_N^p = V_N^T \text{ and } V_I^{p,n} = W_0^{T,n}. \tag{B.35}$$

(B.25) is replaced by

$$V_N^T = \Pi(n) + \frac{1-\delta^D}{1+\rho}\left(\left(1-\delta^I\right)V_N^T + \delta^I\gamma W_0^{T,n}\right), \tag{B.36}$$

and (B.27) by:

$$V_0^{T,n} = (1-b)V_1^T + b\theta\Pi(n) + b\frac{1-\delta^D}{1+\rho}\left(\left(1-\delta^I\right)V_N^T + \delta^I\gamma W_0^{T,n}\right), \tag{B.37}$$

while (B.8) and (B.28) still hold. Using (B.28) and (B.37) we can now write:

$$\begin{aligned} V_0^{T,n} = & \frac{(1+\rho)\left((1-b)V_1^T + b\theta\Pi(n)\right)}{1+\rho-b(1-\delta^D)\delta^I\gamma} + \frac{b\left(1-\delta^D\right)}{1+\rho-b(1-\delta^D)\delta^I\gamma}\left(1-\delta^I\right)V_N^T \\ & - \frac{b\left(1-\delta^D\right)\delta^I\gamma\left((1-b)\left(\Pi(x^*)-\gamma^{-1}\Pi(y^*)\right)+b\theta\left(1-\gamma^{-1}\right)\Pi(n)\right)}{1+\rho-b(1-\delta^D)\delta^I\gamma}, \end{aligned} \tag{B.38}$$

which combined with (B.8) and (B.36) gives:

$$\begin{aligned} V_N^T = & \frac{1+\rho-b\left(1-\delta^D\right)\delta^I\gamma}{1+\rho-(1-\delta^D)(1-\delta^I)-b(1-\delta^D)\delta^I\gamma}\Pi(n) \\ & + \frac{\delta^I\gamma\left(1-\delta^D\right)\left((1-b)V_1^T+b\theta\Pi(n)\right)}{1+\rho-(1-\delta^D)(1-\delta^I)-b(1-\delta^D)\delta^I\gamma} \\ & - \frac{\delta^I\gamma\left(1-\delta^D\right)\left((1-b)\left(\Pi(x^*)-\gamma^{-1}\Pi(y^*)\right)+b\theta\left(1-\gamma^{-1}\right)\Pi(n)\right)}{1+\rho-(1-\delta^D)(1-\delta^I)-b(1-\delta^D)\delta^I\gamma}, \end{aligned} \tag{B.39}$$

which plugged back in (B.38) leads to:

$$\begin{aligned} V_0^{T,n} = & \frac{\left(1+\rho-\left(1-\delta^D\right)\left(1-\delta^I\right)\right)\left((1-b)V_1^T+b\theta\Pi(n)\right)}{1+\rho-(1-\delta^D)(1-\delta^I)-b(1-\delta^D)\delta^I\gamma} \\ & + \frac{b\left(1-\delta^D\right)\left(1-\delta^I\right)}{1+\rho-(1-\delta^D)(1-\delta^I)-b(1-\delta^D)\delta^I\gamma}\Pi(n) \\ & - \frac{b\left(1-\delta^D\right)\delta^I\gamma\left((1-b)\left(\Pi(x^*)-\gamma^{-1}\Pi(y^*)\right)+b\theta\left(1-\gamma^{-1}\right)\Pi(n)\right)}{1+\rho-(1-\delta^D)(1-\delta^I)-b(1-\delta^D)\delta^I\gamma} \end{aligned}$$

Using these last expressions, we can rewrite (B.34) as (B.32) (which is necessary to get the equilibrium in the first place).

Finally (A.13) gives $V_1^s = V_1^T - V_N^T$, (A.10) still holds, so using (A.11), (B.35) and (B.39), we can express I exactly as in (B.33).

Case 4. We now consider the case where a producer not in a good match would rather look for a new supplier than stick to a non-cooperating good match in periods without innovations, while in periods with innovation he tries out the innovator but the non-cooperating good match represents a better alternative than trying out an outdated supplier. Note that no matter what, the non-cooperating good match actually never works with the producer, his value is then always null and it does not matter whether the producer knows only one non-cooperating good match or more. The conditions to be in that case can then be expressed as:

$$V_N^T < V_0^{T,n} \text{ and } W_0^{T,n} < W_N^T. \tag{B.40}$$

Bertrand competition implies that

$$V_I^{p,n} = W_N^T.$$

We then get

$$V_N^T = \Pi(n) + \frac{1-\delta^D}{1+\rho}\left(\left(1-\delta^I\right)V_0^{T,n} + \delta^I\gamma W_N^T\right), \tag{B.41}$$

and

$$V_0^{T,n} = (1-b)V_1^T + b\theta\Pi(n) + b\frac{1-\delta^D}{1+\rho}\left(\left(1-\delta^I\right)V_0^{T,n} + \delta^I\gamma W_N^T\right), \tag{B.42}$$

while (B.8) and (B.28) still hold.

Using (B.8), (B.41) and (B.42), we can write $V_0^{T,n}$ as:

$$V_0^{T,n} = \frac{\left(1+\rho-\left(1-\delta^D\right)\delta^I\gamma\right)\left((1-b)V_1^T + b\theta\Pi(n)\right) + b\delta^I\left(1-\delta^D\right)\Pi(n)}{1+\rho-\left(1-\delta^D\right)\delta^I\gamma - b\left(1-\delta^D\right)\left(1-\delta^I\right)}, \tag{B.43}$$

plugging this back in (B.41) and using (B.8), we get:

$$
\begin{aligned}
&V_N^T \\
&= \frac{\left(1+\rho-\left(1-\delta^D\right)\delta^I\gamma\left(1-\gamma^{-1}\right)\right)\left(1+\rho-\left(1-\delta^D\right)\delta^I\gamma-b\left(1-\delta^D\right)\left(1-\delta^I\right)\right)\Pi\left(n\right)}{\left(1+\rho-\left(1-\delta^D\right)\delta^I\gamma\right)\left(1+\rho-\left(1-\delta^D\right)\delta^I\gamma-b\left(1-\delta^D\right)\left(1-\delta^I\right)\right)} \\
&= \frac{-b\delta^I\left(1-\delta^D\right)^2\left(1-\delta^I\right)\Pi\left(n\right)}{\left(1+\rho-\left(1-\delta^D\right)\delta^I\gamma\right)\left(1+\rho-\left(1-\delta^D\right)\delta^I\gamma-b\left(1-\delta^D\right)\left(1-\delta^I\right)\right)} \\
&+ \frac{\left(1-\delta^D\right)\left(1-\delta^I\right)}{1+\rho-\left(1-\delta^D\right)\left(\delta^I\gamma\right)-b\left(1-\delta^D\right)\left(1-\delta^I\right)}\left(\left(1-b\right)V_1^T+b\theta\Pi\left(n\right)\right).
\end{aligned}
$$

Now combining these two last expressions with (B.8) and (B.28) we can rewrite (B.40) as:

$$
\begin{aligned}
&\frac{\left(1+\rho-b\left(1-\delta^D\right)-\left(1-\delta^D\right)\delta^I\left(\gamma-1\right)\right)\Pi\left(n\right)}{1+\rho-\left(1-\delta^D\right)\left(1-\delta^I+\delta^I\gamma\right)} \\
&< \left(1-b\right)V_1^T+b\theta\Pi\left(n\right) \\
&< \frac{\left(1+\rho-\left(1-\delta^D\right)\delta^I\left(\gamma-1\right)-b\left(1-\delta^D\right)\right)\Pi\left(n\right)}{1+\rho-\left(1-\delta^D\right)\left(1-\delta^I+\delta^I\gamma\right)} \\
&+ \frac{1+\rho-b\left(1-\delta^D\right)\left(1-\delta^I\right)-\left(1-\delta^D\right)\delta^I\gamma}{1+\rho-\left(1-\delta^D\right)\left(1-\delta^I+\delta^I\gamma\right)}\left(\left(1-b\right)\left(\Pi\left(x^*\right)-\gamma^{-1}\Pi\left(y^*\right)\right)-\left(1-b\theta\right)\left(1-\gamma^{-1}\right)\theta\Pi\left(n\right)\right)
\end{aligned}
\tag{B.44}
$$

note that this case requires that $\left(1-b\right)\left(\Pi\left(x^*\right)-\gamma^{-1}\Pi\left(y^*\right)\right) > \left(1-b\theta\right)\left(1-\gamma^{-1}\right)\theta\Pi\left(n\right).$

Finally to express the incentive constraint, first note that (A.12) holds so combining (A.11) and (B.43), we get:

$$
V_1^s = \frac{b\left(\left(1+\rho-\left(1-\delta^D\right)\delta^I\gamma\right)\left(\Pi\left(x^*\right)-\theta\Pi\left(n\right)\right)+\left(1-\delta^D\right)\delta^I\left(\Pi\left(y^*\right)-\Pi\left(n\right)\right)\right)}{1+\rho-\left(1-\delta^D\right)\left(\delta^I\gamma\right)-b\left(1-\delta^D\right)\left(1-\delta^I\right)},
$$

now, as (A.10) holds, we get:

$$
I = \frac{1+\rho}{1+\rho-b\left(1-\delta^D\right)\delta^I\gamma}\left(\begin{array}{c}\left(1-\delta^I\right)b\frac{\left(1+\rho-\left(1-\delta^D\right)\delta^I\gamma\right)\left(\Pi\left(x^*\right)-\theta\Pi\left(n\right)\right)+\left(1-\delta^D\right)\delta^I\left(\Pi\left(y^*\right)-\Pi\left(n\right)\right)}{1+\rho-\left(1-\delta^D\right)\left(\delta^I\gamma\right)-b\left(1-\delta^D\right)\left(1-\delta^I\right)} \\ +\delta^I\gamma\left(\frac{1}{\gamma}\Pi\left(y^*\right)-\left(\left(1-b\right)\Pi\left(x^*\right)+b\theta\Pi\left(n\right)\right)\right)^+\end{array}\right).
$$

Further, note that when a producer in a good match switches to the innovator, we do get that staying with the good match supplier is indeed the second best option and not switching to the non-cooperating good match, that is $W_1^p = W_1^T - V_A^s > W_N^T$.

Case 5. We treated that case in Appendix A.1, except that we did not derive the conditions to be in it. Case 5 occurs when $V_N^T < V_0^T$ and $W_N^T < W_0^T$.

In a period without innovation, the joint value of a relationship with the non-cooperating good match now obeys

$$V_N^T = \Pi(n) + \frac{1-\delta^D}{1+\rho}\left(\left(1-\delta^I\right)V_0^T + \delta^I\gamma W_0^T\right). \tag{B.45}$$

Indeed, after one period, the producer will look for a new supplier in a period without innovation, and in a period with innovation a new outdated supplier will be his second best option (after the innovator). Similarly in a period with innovation, the joint value of a relationship with a non-cooperating good match (necessarily outdated) is given by:

$$W_N^T = \frac{1}{\gamma}\Pi(n) + \frac{1-\delta^D}{1+\rho}\left(\left(1-\delta^I\right)V_0^T + \delta^I\gamma W_0^T\right). \tag{B.46}$$

Combining (B.45) and (B.46) with (6) and (A.15), we obtain that:

$$W_0^T - W_N^T = V_0^T - V_N^T - \left((1-b)\left(\Pi(x^*) - \gamma^{-1}\Pi(y^*)\right) - (1-b\theta)\left(1-\gamma^{-1}\right)\Pi(n)\right) \tag{B.47}$$

Therefore if $(1-b)(\Pi(x^*) - \gamma^{-1}\Pi(y^*)) > (1-b\theta)(1-\gamma^{-1})\Pi(n)$, then $W_N^T < W_0^T$ is the stricter constraint and otherwise $V_N^T < V_0^T$ is the stricter one.

Combining (B.45), (6) and (A.15), we further get:

$$\begin{aligned} & V_0^T - V_N^T \\ = \; & \frac{\left(1+\rho-\left(1-\delta^D\right)\left(1-\delta^I+\delta^I\gamma\right)\right)\left((1-b)V_1^T + b\theta\Pi(n)\right)}{1+\rho-b\left(1-\delta^D\right)\left(1-\delta^I+\delta^I\gamma\right)} \\ & -\frac{\left(1+\rho-b\left(1-\delta^D\right)-\left(1-\delta^D\right)\delta^I(\gamma-1)\right)\Pi(n)}{1+\rho-b\left(1-\delta^D\right)\left(1-\delta^I+\delta^I\gamma\right)} \\ & -\frac{(1-b)\left(1-\delta^D\right)\delta^I\gamma\left[(1-b\theta)\left(1-\gamma^{-1}\right)\Pi(n)-(1-b)\left(\Pi(x^*)-\frac{1}{\gamma}\Pi(y^*)\right)\right]}{1+\rho-b\left(1-\delta^D\right)\left(1-\delta^I+\delta^I\gamma\right)} \end{aligned} \tag{B.48}$$

Therefore $V_0^T > V_N^T$ is equivalent to

$$\begin{aligned}
& (1-b) V_1^T + b\theta\Pi(n) && \text{(B.49)} \\
> \; & \frac{1+\rho-b\left(1-\delta^D\right)-\left(1-\delta^D\right)\delta^I(\gamma-1)}{1+\rho-(1-\delta^D)(1-\delta^I+\delta^I\gamma)}\Pi(n) \\
& + \frac{(1-b)\left(1-\delta^D\right)\delta^I\gamma\left[(1-b\theta)(1-\gamma^{-1})\Pi(n)-(1-b)\left(\Pi(x^*)-\frac{1}{\gamma}\Pi(y^*)\right)\right]}{1+\rho-(1-\delta^D)(1-\delta^I+\delta^I\gamma)}
\end{aligned}$$

Combining (B.47) with (B.48), we obtain that $W_0^T > W_N^T$ is equivalent to:

$$\begin{aligned}
& (1-b) V_1^T + b\theta\Pi(n) && \text{(B.50)} \\
> \; & \frac{1+\rho-b\left(1-\delta^D\right)-\left(1-\delta^D\right)\delta^I(\gamma-1)}{1+\rho-(1-\delta^D)(1-\delta^I+\delta^I\gamma)}\Pi(n) \\
& + \frac{\left(1+\rho-b\left(1-\delta^D\right)\left(1-\delta^I\right)-\left(1-\delta^D\right)\delta^I\gamma\right)}{1+\rho-(1-\delta^D)(1-\delta^I+\delta^I\gamma)} \\
& \times \left((1-b)\left(\Pi(x^*)-\gamma^{-1}\Pi(y^*)\right)-(1-b\theta)\left(1-\gamma^{-1}\right)\Pi(n)\right).
\end{aligned}$$

Recalling that $W_N^T < W_0^T$ is the stricter constraint if and only if $(1-b)\left(\Pi(x^*)-\gamma^{-1}\Pi(y^*)\right) > (1-b\theta)\left(1-\gamma^{-1}\right)\Pi(n)$, we can combine (B.49) and (B.50), to get that the equilibrium is in case 5 if and only if:

$$\begin{aligned}
& (1-b) V_1^T + b\theta\Pi(n) && \text{(B.51)} \\
> \; & \frac{1+\rho-b\left(1-\delta^D\right)-\left(1-\delta^D\right)\delta^I(\gamma-1)}{1+\rho-(1-\delta^D)(1-\delta^I+\delta^I\gamma)}\Pi(n) \\
& + \frac{(1-b)\left(1-\delta^D\right)\delta^I\gamma}{1+\rho-(1-\delta^D)(1-\delta^I+\delta^I\gamma)}\left((1-b\theta)\left(1-\gamma^{-1}\right)\theta\Pi(n)-(1-b)\left(\Pi(x^*)-\gamma^{-1}\Pi(y^*)\right)\right)^+ \\
& + \frac{1+\rho-\left(1-\delta^D\right)\left(b\left(1-\delta^I\right)+\delta^I\gamma\right)}{1+\rho-(1-\delta^D)(1-\delta^I+\delta^I\gamma)}\left((1-b)\left(\Pi(x^*)-\gamma^{-1}\Pi(y^*)\right)-(1-b\theta)\left(1-\gamma^{-1}\right)\theta\Pi(n)\right)^+.
\end{aligned}$$

Summary Overall conditions (B.10), (B.20), (B.32), (B.44) and (B.51) span all the possibilities. Moreover, a producer who knows a non-cooperating good match and one who does not always face the same incentive constraint, so that the profile of investment levels played by a new supplier can indeed be the same no matter whether the producer knows other good matches or not. The IC constraint of a good match supplier only takes two forms depending on whether the supplier has access to the frontier technology or not.

B.1.2 Part 4

Therefore to prove the existence of the equilibrium, the last step is to show that there always exists a solution to x^* and y^* such that IC constraint binds or the first best is achieved. As argued in Appendix A.1, we simply need to show the IC constraints do not bind for (x, y) just above n, and since n minimizes φ, it is enough to show that I is positive at the first order in $(x-n)$ and $(y-n)$ when x and y are greater than n.

Note that as $((1-b)(\Pi(x)-\gamma^{-1}\Pi(y))-(1-b\theta)(1-\gamma^{-1})\Pi(n)) = -(1-\theta)(1-\gamma^{-1})b\Pi(n)+O(x-n)+O(y-n)$, the only possible cases when x, y are close to n are 1, 2, 3 and 5. We have already checked that I is positive for x and y just above n in case 5. In case 1, we get

$$I = \Pi'(n)\frac{(1+\rho)\left(\left(1-\delta^I\right)(x-n)+\delta^I(y-n)\right)}{1+\rho-(1-\delta^D)\left(1+\delta^I(\gamma-1)\right)} + o(x-n) + o(y-n),$$

which is positive at first order in $(x-n)$, $(y-n)$ when x and y approach n by superior values. In case 2, we get

$$I = \frac{1+\rho}{1+\rho-b\left(1-\delta^D\right)\delta^I\gamma}\left(\begin{array}{c}\Pi'(n)\frac{\left(1-\delta^I\right)\left(1+\rho-\left(1-\delta^D\right)\delta^I\gamma\right)(x-n)+\left(1-\delta^D\right)\delta^I(y-n)}{1+\rho-\left(1-\delta^D\right)\left(1-\delta^I+\delta^I\gamma\right)} \\ +\delta^I\gamma\left(\frac{1}{\gamma}\Pi(y)-((1-b)\Pi(x)+b\theta\Pi(n))\right)^+\end{array}\right) \tag{B.52}$$
$$+ o(x-n) + o(y-n),$$

also positive at first order in $(x-n)$, $(y-n)$ when x and y approach n by superior values. In case 3, we get

$$\begin{aligned}
&I \\
&= \frac{1+\rho}{1+\rho-b\left(1-\delta^D\right)\delta^I\gamma}\frac{\left(1-\delta^I\right)b\left(1-\delta^D\right)\delta^I(1-\theta)\Pi(n)}{1+\rho-\left(1-\delta^I\right)\left(1-\delta^D\right)-\left(1-\delta^D\right)b\delta^I\gamma} \\
&+ \frac{1+\rho}{1+\rho-b\left(1-\delta^D\right)\delta^I\gamma}\delta^I\gamma\left(\frac{1}{\gamma}\Pi(y)-((1-b)\Pi(x)+b\theta\Pi(n))\right)^+ + O(x-n)+O(y-n),
\end{aligned}$$

which is positive at first order in $(x-n)$, $(y-n)$. Therefore, there always exist x^* and y^* solutions to the problem above n. This achieves the proof.

B.2 Proof of Proposition 7

Here we prove Proposition 7.

Step 1: Incentive constraint

The incentive constraint must be of the following form. After a history of h_t when a good match supplier makes her investment decision she can invest n instead of the prescribed $z(h_t)$, which would increase ex-post profits this period by $\varphi(z(h_t)) A_k(h_t)$, where $A_k(h_t)$ is the technology of supplier and

$$\varphi(z) \equiv \beta R(n) - n - (\beta R(z) - z)$$

Denoting by $I \in \{0, 1\}$ either no new innovation ($I = 0$) or a new innovation ($I = 1$) we can express the incentive constraint as:

$$\varphi(z(h_t)) A_k(h_t) \leq \frac{1-\delta^D}{1+\rho} \times \tag{B.53}$$

$$\left(\begin{array}{c} \left(1-\delta^I\right) V^{s,k}\left(h_t \cup \{z(h_t)\} \cup \{I=0\}\right) + \delta^I V^{s,k}\left(h_t \cup \{z(h_t)\} \cup \{I=1\}\right) \\ -\left(\left(1-\delta^I\right) V^{s,k}\left(h_t \cup \{n\} \cup \{I=0\}\right) + \delta^I V^{s,k}\left(h_t \cup \{n\} \cup \{I=1\}\right)\right) \end{array} \right),$$

where $V^{s,k}(h)$ denotes the value of the supplier after history h (The continuation value after a deviation other than n could be different, but the producer has no reason not to punish any deviation in the same way so we focus on the incentive not to play n).

Step 2: Producers in a good match do not switch suppliers in periods without innovation

Let us consider the first time the producer meets a good match supplier. Then this good match supplier has an advantage over any other supplier in the future except for a possible innovator, as a consequence any payoff achievable with another supplier is achievable with the good match supplier in periods without innovation. In order to maximize the joint value of the producer and the supplier, the producer should then stick to the supplier at least as long as no innovation occurs. Moreover if an innovation occurs, and the equilibrium is such that the producer should not switch to the innovator, the argument carries through. Note that the argument applies no matter whether the first good match supplier happens to be outdated or not, moreover, because of condition 1, this must be true for any relationship not only the first time the producer starts a relationship with a good match supplier.

Step 3 Bertrand competition

Condition 1 imposes that strategies once a producer has chosen a supplier are independent of the ex-ante transfer that was made, so that the ex-ante transfer does not affect the joint value of a relationship if the producer and the supplier end up working together. As a consequence the supplier whom the producer ends up working with, must offer an ex-ante transfer low enough that the value of the producer is the same as it would have been if he had chosen another supplier. In return, the supplier with whom the producer ends up being just indifferent to start working with or not must make an ex-ante transfer sufficiently high that he is himself just indifferent between working with the producer or not. A new supplier will then just break-even (as his value is zero if the producer does not choose him); and, the value of a good match supplier when an innovation occurs and for parameters such that the producer switches to the innovator, $V^{s,k}\left(h_t \cup \{z\left(h_t\right)\} \cup \{I=1\}\right)$, is not zero because the producer may go back to the old supplier after having tried the innovator.

Without condition 1 it is possible to build equilibria where the value of the producer is strictly higher than the value he gets with his second best option by conditioning normalized investment levels on the ex-ante transfer offered.

Step 4: The joint value of a producer and a supplier is the value on the path where they never stop working together

This is obvious in the case where the producer does not switch to the innovator. In the case where the producer does switch to the innovator, the take-it or leave-it offer of the innovator implies that the value of the producer should be equal to the value of the producer had he stayed with his old supplier, as specified by step 3. Now the old supplier is willing to offer an ex-ante transfer that leaves the producer with the entire joint value of a relationship with him had they stayed together minus what the supplier still gets when the producer switches to the innovator. Therefore the joint value of the producer and the old supplier is the joint value of the producer and the supplier on the path where the producer and the supplier do not break-up their relationships when there is an innovation. Note that on this path the argument of step 2 applies and the producer and the supplier never break up.

Step 5: The value of the supplier is also determined by the value on the path where the producer and the supplier never stop working together

Combining steps 3 and 4, the value of the supplier in periods without innovation is simply given by the joint value of the producer and the supplier on the path where they relationship never breaks down minus the value of the second best relationship that the producer can get. This reasoning extends to periods with innovations when the producer does not switch. When the producer switches, the value of the supplier is given by the rents he can capture in the future if the innovator turns out to be a bad match, so that the producer comes back to the supplier, and by condition 2, we know that the strategies must then be identical to the strategies if the producer had not switched. Therefore in this case too, the value of the supplier is ultimately determined by the joint value of the producer and the supplier in the future, on the path where they never stop working together.

Step 6: Higher investment levels in the future increase the RHS of the IC constraint

Higher investment levels on the path where the relationship does not break down then necessarily increases the joint value of the relationship of the producer and the supplier (step 4), as a consequence, they also increase the value of the supplier and make the IC constraint looser (step 5).

Note that condition 2 stipulates that if the producer switches and the innovator has turned out to be a bad match, the behavior of the producer and the supplier is identical to what they would have done had they stick together, so that the investment levels in that case are in fact the investment levels in the path where the producer and the supplier never break up. On the contrary, if the innovator turns out to be a good match, we know from step 2 that the producer would then stay with the innovator. As a consequence, the previous claim (the higher the investment on the path where the relationships does not break, the looser the IC constraints in previous periods) implies that the higher the investment in the relationship on the equilibrium path, the looser the IC constraints in the previous period.[55]

Step 7: Investments are at first best or the IC constraint binds

From step 6, it is then direct that investment in a good match relationship should be as high as possible (on the equilibrium path and on the path where the relationship never breaks down), therefore either the first best must be achieved or the IC constraint binds.

Step 8: Punishment strategies

To achieve the highest possible investment level, the RHS of the incentive constraint must in return be the highest possible. Therefore, the value of the supplier in case a

[55]Without condition 2 this would not necessarily be the case, lower investment levels if the producer switches to the innovator and comes back could reduce the incentive for the innovator to switch and therefore increase the joint value of the relationship.

deviation occurs must be as low as possible, this occurs if the supplier plays the Nash level of investment n in any future interaction between the producer and the supplier. The value of the supplier if no deviation occurs must be as high as possible.

Let us first consider the case of periods without innovation. Step 5 already ensures that the suppliers gets the largest possible value out of the relationship, so that the producer is just indifferent between staying in the current relationship or starting a new one (from which he would capture the entire benefit).[56] Now, to still ensure the highest value for the supplier it is then necessary that the value the producer can get with his second best option must be as low as possible. If the producer switches to a new supplier who turns out to be a bad match, then the producer may be willing to come back to the old supplier; as long as the first best is not achieved, the strategy of the old supplier should then be never to cooperate in that case, as cooperation in the future increases the value that the producer could capture by switching. If the producer switches to a new supplier and this supplier turns out to be a good match, then condition 1 specifies what the outcome is (and we come back to that case to discuss what the strategy of the old supplier must be in step 9). If the producer switches to a good match with whom a deviation has occurred, then the strategy of the old supplier should be such that - on that path- the producer does not cooperate again with him (for the same reason). In periods where innovation occurs, the same reasoning applies: Bertrand competition ensures that the supplier already gets the largest possible value of the relationship if the producer does not switch, while, otherwise, the value of the supplier is fixed according to step 5; if the producer deviates to an old good match supplier with whom a deviation has occurred, the strategy of the supplier must be such that he does not cooperate with the producer again in the future (a producer will necessarily prefer to deviate by switching to the innovator than a new outdated good match, so we don't have to specify what happens in that case); however condition 2 stipulates that if the producer switches to the innovator and the innovator turns out to be a good match, the old supplier must forgive the producer.

Step 9: Investment by good match suppliers when a deviation has occurred

Step 8 already specified that if a supplier deviates, his investment in the future must be at the Nash level. It also specified the behavior of the supplier if the producer has

[56] One could then dispense with the assumption that normalized investment levels do not depend on the ex-ante transfers if the first best is not reached: condition 3 would then ensures that the supplier would capture the entire benefit of the relationship.

deviated but only found bad matches. Now let us focus on the case where the producer deviates and finds a new good match, we denote by k the previous supplier and by k' the new supplier. Condition 1 stipulates that the outcome should be identical to the outcome in the first interaction between the producer and a good match supplier. As explained in step 8, the investment level in the relationship with supplier k' is going to directly depend on the outside option of the producer. A better outside option for the producer implies a lower lower value for supplier k', and therefore (unless the first best is achieved), lower investment levels. When the producer met a good match supplier for the first time, however, there was no other good match supplier that the producer knew, so his outside option was a priori worse. To satisfy condition 1, we must then ensure that the value of the producer once he has started a relationship with supplier k' must be as low as possible if he is to switch to a different supplier. If switching to a new match is better than resuming working with the old good match supplier, then what the old good match supplier would do does not matter, however, if resuming working with supplier k is the best option, the value of a relationship with the supplier k must be as low as possible, which is achieved if supplier k plays the Nash level of investment in any possible future interaction.[57]

Therefore, as soon as the producer switches suppliers[58] (except if it is the innovator and the innovator turns out to be a bad match), or the supplier under-invests, investment in any future interaction between the producer and the supplier leads to the Nash level of normalized investment.

Step 10: Excepted strategies of the other players are identical in all periods.

Condition 1 stipulates how new good match suppliers are expected to behave in the future. Condition 4 stipulates that the investment level of bad matches is n, and we derived that the investment level of good match suppliers with whom a deviation has occurred must be at the Nash level n too. Moreover Bertrand competition determines

[57]Note that without condition 1, the value of the supplier k could be increased if when the producer switches to the supplier k', some cooperation were to arise if the producer comes back to the supplier k in the future. This does not contradict condition 3 though, because condition 3 takes as given the strategy of the producer once the producer has started working with a new good match supplier. Note also that even if the old good match supplier keeps playing the Nash level in any future interaction there is no guarantee that it is possible to satisfy condition 1, part 3 of the proof of Proposition 1 showed that there is no contradiction though.

[58]Technically this needs to be true only if working again with a good match supplier with whom a deviation has occurred is a better second option than starting over a relationship, for a producer who is in a new good match relationship. Of course, in the other case, the strategy of the old good match does not matter, so we can assume that he plays the Nash level without affecting the analysis.

the form of the ex-ante transfer that these suppliers are willing to offer: in periods without innovation they should all break even, in periods with innovation the innovator can capture the surplus from his innovation if the producer switches. Therefore when the supplier makes his investment decision all periods are identical in term of the strategies played by the other players, the only difference is that in periods where the supplier has access only to an outdated technology he knows that he will get access to the frontier technology in the next period.

Step 11: Investment levels are constant

The only thing that remains to be proved is that investment levels are the same in good matches in all periods when the supplier has access to the frontier technology and when the supplier has access only to an outdated technology, step 8 has already proved that the situation was symmetric in all periods where the supplier has access to the frontier technology and in all periods where he does not, so if a path of investment is sustainable after one given history, it is also sustainable after any other history (provided that the access to the frontier technology for the supplier is the same).

Let us then consider a history h_{t_0}, where a producer and a good match supplier starts working together for the first time. Let us denote by $x(h_t)$ and $y(h_t)$ the investment levels in all histories h_t belonging to the set of histories H_{t_0} of histories following h_{t_0} on the path where the relationship between the producer and the supplier never break-up their relationship, where $x(h_t)$ is used for histories where the supplier has access to the frontier technology and $y(h_t)$.for when he does not. The joint value of the producer and a supplier at an history $\widetilde{h}_t \in H_{t_0}$ is then simply the discounted sum of the expected values of profits on the path following $\widetilde{h}_t$ where the producer and the supplier never break up their relationship, that is it a discounted sum of the $\Pi(x(h_t))$ and $\Pi(y(h_t))$ for h_t following $\widetilde{h}_t$ in H_{t_0} Therefore, there is a M_1 such that if for all h_t following $\widetilde{h}_t$ in H_{t_0}, $|\Pi(x(h_t)) - \Pi(\widehat{x})| < \nu/M_1$ and $[\Pi(y(h_t)) - \Pi(\widehat{y})] < \nu/M_1$, then the joint value of the relationship is within ν of what the joint value of the relationship would have been if the investments levels where $\widehat{x}$ at all histories where the producer has access to the frontier technology and $\widehat{y}$ in histories where he does not, and by symmetry between the different periods, we can choose the same M_1 for all histories $h_t \in H_{t_0}$.

Now because of step 5, the value of the supplier (in all cases) is determined by the joint value of the producer and the supplier on the history path where the relationship does not break down, and, because of step 10 the strategies of the other players are the same over time, therefore there exists a M_2, such that if the joint value of the producer and the supplier is within v/M_2 of the joint value of the producer and the supplier if investment levels had been $\widehat{x}$ in periods where the supplier has the frontier technology and $\widehat{y}$ in the other periods, the value of the supplier is within ν of the value of the producer and the supplier if investment levels had followed the same alternative (and this M_2 can be the same for all histories $h_t \in H_{t0}$). Finally because the RHS of the IC constraint is just the discounted expectation of the value of the supplier in the next period, there is therefore a M_3 such that if the value of the supplier is (in both the case with innovation and the case without innovation) within ν/M_3 of the value of the supplier if the investment levels were given by $\widehat{x}$ in periods where the supplier has the frontier technology and $\widehat{y}$ otherwise, then the RHS of the IC constraint is within ν of the RHS of the IC constraint under the alternative profile.

Let us define $\overline{x} = \sup\left(x\left(h_t\right)|h_t \in H_{t_0}\right)$ and $\overline{y} = \sup\left(x\left(y_t\right)|h_t \in H_{t_0}\right)$. Then, for any $\varepsilon > 0$, there exists some $\eta > 0$ such that if $x\left(h_t\right) \in \left[\overline{x} - \eta, \overline{x}\right]$ and $y\left(h_t\right) \in \left[\overline{y} - \eta, \overline{y}\right]$, for all histories $h_t \in H_{t_0}$, the RHS of the IC constraint after any history $h_t \in H_{t_0}$ when the profile of normalized investment is given by $x\left(\widetilde{h}_t\right)$ $y\left(\widetilde{h}_t\right)$ where $\widetilde{h}_t$ are the histories following h_t in the set H_{t_0}, is weakly smaller than ε + the RHS of the IC constraint if the profile of normalized investment was given by $\overline{x} - \eta$, $\overline{y} - \eta$ (we just have to choose η such that if $x \in \left[\overline{x} - \eta, \overline{x}\right]$, and $y \in \left[\overline{y} - \eta, \overline{y}\right]$, then $\left|\Pi\left(x\right) - \Pi\left(\overline{x}\right)\right| < \varepsilon/\left(M_1 M_2 M_3\right)$ and $\left|\Pi\left(y\right) - \Pi\left(\overline{y}\right)\right| < \varepsilon/\left(M_1 M_2 M_3\right)$). Let us then define $\varphi_\varepsilon \equiv \varphi - \varepsilon$.

Moreover there exists a history $h_t^1 \in H_{t_0}$, where $x\left(h_t^1\right) \in \left[\overline{x} - \min\left(\eta, \varepsilon\right), \overline{x}\right]$, so that $x\left(h_t^1\right)$ must satisfy the IC constraint at history h_t^1, therefore it necessarily satisfies the IC constraint if the normalized investment in the future were given by $max\left(x\left(\widetilde{h}_t\right), \overline{x} - \min\left(\eta, \varepsilon\right)\right)$ and $max\left(y\left(\widetilde{h}_t\right), \overline{y} - \min\left(\eta, \varepsilon\right)\right)$ instead of the actual path $x\left(\widetilde{h}_t\right)$ $y\left(\widetilde{h}_t\right)$ where $\widetilde{h}_t$ are the histories following h_t^1 in the set H_{t_0}. Note then that, by the definition of η, $x\left(h_t^1\right)$ would satisfy the IC constraint if the incentive to deviate was given by φ_ε instead of ε and the profile of normalized investment levels was given by $\overline{x} - \min\left(\eta, \varepsilon\right)$ and $\overline{y} - \min\left(\eta, \varepsilon\right)$ in any future histories. Similarly we can find a history $h_t^2 \in H_{t_0}$, where $y\left(h_t^2\right) \in \left[\overline{y} - \min\left(\eta, \varepsilon\right), \overline{y}\right]$, and the same property arises.

Now let us consider the profile of normalized investment where for all histories $h_t \in h_{t_0}$, we replace $x\left(h_t\right)$ by $max\left(x\left(h_t\right), \overline{x} - \min\left(\eta, \varepsilon\right)\right)$ and $y\left(h_t\right)$ by $max\left(y\left(h_t\right), \overline{y} - \min\left(\eta, \varepsilon\right)\right)$, then this alternative profile necessarily leads to a strictly higher investment joint value (as long as $x\left(h_t\right)$ is not always equal to $\overline{x}$, and $y\left(h_t\right)$ is not always equal to $\overline{y}$) and for any history where the normalized investment level has not changed, the IC constraint remains satisfied. Let us now consider a history h_t' where the investment level has changed under the alternative profile and the supplier has access to the frontier technology, the profile of future investment levels is within $[\overline{x} - \min(\eta, \varepsilon), \overline{x}]$ and $[\overline{y} - \min(\eta, \varepsilon), \overline{y}]$, and we know that $x(h_t^1)$ (which is weakly larger than $\overline{x} - \min(\eta, \varepsilon)$) satisfies the IC constraint if the profile of future investment is given by $\overline{x} - \min(\eta, \varepsilon)$ and $\overline{y} - \min(\eta, \varepsilon)$ and φ is replaced by φ_ε, therefore the investment level $\overline{x} - \min(\eta, \varepsilon)$ also satisfies the IC constraint at history h_t' under the alternative profile provided that φ is replaced by φ_ε. The same logic applies to periods where the supplier does not have access to the frontier technology.

Thus, the alternative profile leads to a higher joint value and is sustainable up to replacing φ by φ_ε, letting ε goes to 0, we get that a profile with constant investment $\overline{x}$, $\overline{y}$ satisfies the IC constraint and yields a higher joint value. As a consequence, normalized investment must take two values only one for when the supplier has access to the frontier technology and one for when he has access to the outdated technology.

Furthermore the situation is symmetric whether during their first interaction the supplier has access to the frontier technology or not, so if he does not, investment levels are also determined by the same two constants. Finally, condition 1 stipulates that the profile of investment levels needs to be the same in any new good match relationship.

Step 12 Summary

So far we have then shown that in a SPNE equilibrium satisfying conditions 1-4:

1. investment levels in all good matches are given by two constant x^* and y^*, where the former is undertaken when the supplier has access to the best technology and the latter when he does not, as long as no deviation has occurred in the relationship between the producer and the supplier;

2. investment levels are at the first best level if possible and otherwise the IC constraint binds;

3. producers stay with the same supplier until an innovation or a deviation occurs, if an innovation occurs, the producer may or may not switch, but if he switches and the innovator turned out to be a bad match, he goes back to his old supplier;

4. producers are just indifferent between choosing the supplier they are supposed to work with on equilibrium path and choosing the "second best" supplier, the "second best" supplier is just indifferent between being chosen and not being chosen by the producer;

5. if a supplier deviates once, investment is at the Nash level in any further interaction, and - without loss of generality-[59] if the producer deviates (by switching to another supplier except if it is the innovator and the innovator turned out to be a bad match) investment is also at the Nash level in any future interaction.

These conditions correspond to the strategies described in Proposition 1.

B.3 Proof of Proposition 8 and Remark 4

We seek to demonstrate that both x^* and y^* are (weakly) increasing in γ and the conditions under which x^* and y^* are decreasing in δ^I. The proof of the rest of Proposition 8 follows along the same lines and is omitted. We prove the proposition in the same special case as considered in the main text. The proof for the remaining cases is analogous and is omitted. Trivially, the effect is zero when x and y are equal to the first best m. When they are not, equations (10) and (11) deliver:[60]

$$\varphi(x) - f(x,y,\gamma) = 0, \tag{B.54}$$

$$\varphi(y) - \gamma f(x,y,\gamma) = 0, \tag{B.55}$$

where

$$\varphi(x) = (\beta R(n) - n) - (\beta R(x) - x),$$

$$f(x,y,\gamma) = \frac{\left(1-\delta^D\right) b \left(1-\delta^I\right)}{1+\rho - b\left(1-\delta^D\right)\left(1-\delta^I\right) - b\left(1-\delta^D\right)\delta^I\gamma} \times$$

$$\left(\begin{array}{c} \Pi(x) - \theta\Pi(n) \\ + \frac{b\left(1-\delta^D\right)\delta^I}{1+\rho-b(1-\delta^D)\delta^I\gamma}\left(\Pi(y) - \theta\Pi(n)\right) \end{array} \right).$$

We define $g_0(x,y,\gamma) \equiv \varphi(x) - f(x,y,\gamma)$ and $g_1(x,y,\gamma) \equiv \varphi(y) - \gamma f(x,y,\gamma)$ and note that $g_{0,\gamma} < 0$ and $g_{1,\gamma} < 0$, where a subscript denotes a partial derivative.[61] Then there are three cases i) both x and y equal to m if $g_0(m,m) \leq 0$ and $g_1(m,m) \leq 0$, ii) $y = m$ and $x < m$ as a solution to $g_0(x,m) = 0$ if $g_1(x,m) \leq 0$ and iii) x and y are solutions to $g_0(x,y) = 0$ if $g_1(x,y) = 0$.

[59]Stricto sensu, this is not necessary if a good match supplier with whom a deviation has occurred does not count, in the sense that he represents a worse alternative than trying a new supplier.

[60]To be consistent we should use (x^*, y^*) as these are the equilibrium values. This omission will not lead to confusion.

[61]In order to avoid cluttering the notation we will suppress the dependence of g_0 and g_1 on γ

As $\varphi(y)$ is convex and $f(x,y)$ is concave in y (as $\Pi(y)$ is concave in y) on $y\epsilon[n,m]$ it follows that $g_1(x,y)$ is convex in y. Let us define $y = h(x)$ such that $h(x) = m$ if $g_1(x,m) < 0$ and otherwise $g_1(x,h(x)) = 0$ ($h(x)$ is single-valued as $g_1(x,h(x)) = 0$ has a unique solution from $g_1(x,n) < 0$ and convexity of $g_1(x,y)$ in y). Note that $h(x)$ is trivially increasing in x. Further, define $\widetilde{g}(x) = g_0(x,h(x))$ such that either i) $x = m$ and $\widetilde{g}(m) \leq 0$ ii) x is a solution to $\widetilde{g}(x) = 0$. Note that $\widetilde{g}(n) < 0$ and that $\widetilde{g}(x)$ is convex when $h(x)(= y)$ is constant and equal to m.

We first seek to demonstrate that $h(x)'' < 0$ (when not equal to m). Note first, that by concavity of $\Pi(x)$ it follows that $g_{1,x} < 0$ and $g_{1,xx} > 0$ and it has already been argued that g_1 is convex with $g_{1,y} > 0$ and $g_{1,yy} > 0$ (for $y < m$). Differentiate $g_1(x,h(x)) = 0$ twice and get:

$$g_{1,xx} + 2g_{1,xy}h'(x) + g_{1,yy}\left(h'(x)\right)^2 + g_{1,y}h''(x) = 0.$$

By inspection $g_{1,xy} = 0$ and hence $h''(x) < 0$ when $h(x) \neq m$. Along the same lines and using the properties of $h(x)$ we can show that $\widetilde{g}(x)$ is increasing and strictly convex in $x\epsilon(n,m)$. Hence, if $\widetilde{g}(m) \leq 0$ then $x = m$ is optimal. As $\widetilde{g}(x)$ is decreasing in γ this implies that x is decreasing in γ and so is y.

To study the impact of an increase in the innovation rate, δ^I, note that we can rewrite $f(x,y,\gamma)$ as:

$$\frac{\left(1-\delta^D\right)b\left(1-\delta^I\right)}{\left[1+\rho-b\left(1-\delta^D\right)\left(1+\delta^I(\gamma-1)\right)\right]\left(1+\rho-b(1-\delta^D)\delta^I\gamma\right)}\times \tag{B.56}$$

$$\left(\begin{array}{c}(1+\rho)\left(\Pi(x)-\theta\Pi(n)\right)\\ -\delta^I(1-\delta^D)b\left(\gamma\left(\Pi(x)-\theta\Pi(n)\right)-\left(\Pi(y)-\theta\Pi(n)\right)\right)\end{array}\right).$$

Below, we demonstrate that$\gamma\left(\Pi(x)-\theta\Pi(n)\right) > \Pi(y)-\Pi(n)$. Under this condition, a sufficient condition for the expression in equation (B.56) to be decreasing in δ^I (and thereby for x and y to be decreasing in δ^I) is that $\gamma < \frac{1+\rho}{b(1-\delta^D)(2-\delta^I)}$ as written in the proposition. All we need is therefore $\gamma\left(\Pi(x)-\theta\Pi(n)\right) > \Pi(y)-\Pi(n)$ which we now demonstrate.

First, define a function $\kappa(x)$ which is $\kappa(x) = \varphi^{-1}(\gamma\varphi(x))$ if $\gamma\varphi(x) \leq \varphi(m)$ and $\kappa(x) = m$ otherwise. We then define the function:

$$\lambda(x) \equiv \gamma\left(\Pi(x)-\theta\Pi(n)\right)-\left(\Pi(\kappa(x))-\theta\Pi(n)\right). \tag{B.57}$$

Note that as $\kappa(n) = n$ (as $\varphi(n) = 0$), $\lambda(n) > 0$, and

$$\lambda'(x) = \gamma \Pi'(x) - \Pi'(\kappa(x))\kappa'(x). \tag{B.58}$$

Obviously, on parts where κ is constant λ will be increasing. Where κ is not constant, $\kappa'(x) = \gamma\varphi'(x)/\varphi'(\kappa'(x))$. Using this in equation (B.58) returns:

$$\lambda'(x) = \gamma \Pi'(x) \left[1 - \Pi'(\kappa(x))\varphi'(x)/ \left(\Pi'(x)\varphi'(\kappa'(x))\right)\right].$$

With $\kappa(x) \geq x$,$\Pi(x)$ concave and $\varphi(x)$ convex, we get that $\lambda'(x) \geq 0$, hence for any pair $x\epsilon[n, m]$, $\lambda(x) > 0$, in particular $\lambda(x) > 0$ for the equilibrium investment pair (x, y), which completes the proof.

B.4 Proof of Propositions 3 and Remark 1

We derive necessary and sufficient conditions under which $\delta^{Nash} > \delta^{coop}$ in each of the three cases ($\gamma \leq \gamma^{Nash}$, $\gamma \in (\gamma^{Nash}, \gamma^{coop}]$ and $\gamma > \gamma^{coop}$). Then we combine them to prove Propositions 3 and Remark 1.

First case: Assume $\gamma < \gamma^{Nash}$

As part of Appendix B.3 we demonstrated that the function $\lambda(x)$ defined in (B.57) was increasing in x (note that this held regardless of whether after a deviation a producer preferred a non-cooperating good match to a new supplier or not). This directly implies that $\Pi(x) - \gamma^{-1}\Pi(\kappa(x))$, where κ is defined as in Appendix B.3, is also increasing in x. Therefore we must always have

$$\left(1 - \gamma^{-1}\right)\Pi(n) < \Pi(x^*) - \gamma^{-1}\Pi(y^*) \leq \left(1 - \gamma^{-1}\right)\Pi(m).$$

This shows directly that $Z_{Nash} - Z_{coop}$ in (20) is strictly negative, and that if $\gamma \leq \gamma^{Nash}$ then $\delta^{Nash} < \delta^{coop}$ (which proves the first part of Part b)).

Third case: Assume $\gamma > \gamma^{coop}$

Then using (23), we obtain that

$$Z_{Nash} > Z_{coop} \Leftrightarrow \gamma < \frac{1}{1 - b\left(1 - \delta^D\right)} \frac{\Pi(y^*) - \Pi(n)}{\Pi(x^*) - \Pi(n)}. \tag{B.59}$$

Since $\Pi(y^*) \geq \Pi(x^*)$, then $\gamma < \frac{1}{1-b(1-\delta^D)}$ is a sufficient condition to achieve $Z_{Nash} > Z_{coop}$. On the other hand, this equality must be violated for γ large enough (since

$\Pi(x^*) > \Pi(n)$), proving the second part of Part b).

Second case: Assume $\gamma^{Nash} < \gamma \leq \gamma^{coop}$

Then using the definition of ω and (21), we can rewrite:

$$Z_{Nash} - Z_{coop} =$$

$$\frac{1-b}{1-(1-\delta^D)b}\left[\left(1-\delta^D\right)\left(1-b+b\theta-\gamma^{-1}\right)\Pi(n)+\delta^D\left(\left(1-\gamma^{-1}\right)\Pi(n)-\left(\Pi(x^*)-\gamma^{-1}\Pi(y^*)\right)\right)\right].$$

Hence we can rewrite that in this case:

$$Z_{Nash} > Z_{coop} \Leftrightarrow \delta^D < \frac{1-b+b\theta-\gamma^{-1}}{\frac{\Pi(x^*)}{\Pi(n)}-b+b\theta-\gamma^{-1}\frac{\Pi(y^*)}{\Pi(n)}}.$$

To derive this we used that since $\gamma^{-1} < 1-b+b\theta$, and $\frac{\Pi(x^*)}{\Pi(n)}-\gamma^{-1}\frac{\Pi(y^*)}{\Pi(n)} > 1-\gamma^{-1}$, then both the numerator and the denominator are positive. We can also rewrite this equivalence as

$$Z_{Nash} > Z_{coop} \Leftrightarrow \gamma^{-1}\left(1-\delta^D\frac{\Pi(y^*)}{\Pi(n)}\right) < 1-\delta^D\frac{\Pi(x^*)}{\Pi(n)}-b(1-\theta)\left(1-\delta^D\right).$$

Hence we have that if $1-\delta^D\frac{\Pi(x^*)}{\Pi(n)}-b(1-\theta)\left(1-\delta^D\right) > 0$, then $Z_{Nash} > Z_{coop}$ is equivalent to:

$$\gamma > \frac{1-\delta^D\frac{\Pi(y^*)}{\Pi(n)}}{1-\delta^D\frac{\Pi(x^*)}{\Pi(n)}-b(1-\theta)(1-\delta^D)}. \tag{B.60}$$

Proof of Part c).

Therefore we get that $\delta^{coop} < \delta^{Nash}$ if $\gamma > \gamma^{coop}$ and $\gamma < \frac{1}{1-b(1-\delta^D)}$ or if $\gamma^{Nash} < \gamma \leq \gamma^{coop}$ and $\gamma > \frac{1-\delta^D\frac{\Pi(y^*)}{\Pi(n)}}{1-\delta^D\frac{\Pi(x^*)}{\Pi(n)}-b(1-\theta)(1-\delta^D)}$ with $1-\delta^D\frac{\Pi(x^*)}{\Pi(n)}-b(1-\theta)\left(1-\delta^D\right) > 0$. Assume that $\delta^D < \theta\frac{\Pi(n)}{\Pi(m)}$, this ensures that for any x^*, y^* we have $1-\delta^D\frac{\Pi(x^*)}{\Pi(n)}-b(1-\theta)\left(1-\delta^D\right) > 0$ and $1-\delta^D\frac{\Pi(y^*)}{\Pi(n)} > 0$. Moreover we get that

$$\frac{1-\delta^D\frac{\Pi(y^*)}{\Pi(n)}}{1-\delta^D\frac{\Pi(x^*)}{\Pi(n)}-b(1-\theta)(1-\delta^D)} < \frac{1-\delta^D\frac{\Pi(m)}{\Pi(n)}}{1-\delta^D\frac{\Pi(m)}{\Pi(n)}-b(1-\theta)(1-\delta^D)}$$

Hence $\gamma > \frac{1-\delta^D\frac{\Pi(m)}{\Pi(n)}}{1-\delta^D\frac{\Pi(m)}{\Pi(n)}-b(1-\theta)(1-\delta^D)}$ is a stricter condition than $\gamma > \frac{1-\delta^D\frac{\Pi(y^*)}{\Pi(n)}}{1-\delta^D\frac{\Pi(x^*)}{\Pi(n)}-b(1-\theta)(1-\delta^D)}$.

In addition we have:

$$\frac{1-\delta^D\frac{\Pi(m)}{\Pi(n)}}{1-\delta^D\frac{\Pi(m)}{\Pi(n)}-b\left(1-\theta\right)\left(1-\delta^D\right)}=\frac{1}{1-b\left(1-\theta\right)\frac{1-\delta^D}{1-\delta^D\frac{\Pi(m)}{\Pi(n)}}}>\frac{1}{1-b\left(1-\theta\right)}=\gamma^{Nash}$$

since $\Pi\left(m\right)>\Pi\left(n\right)$.

Hence we have that $\delta^{coop}<\delta^{Nash}$ if $\gamma\in\left(\frac{1-\delta^D\frac{\Pi(m)}{\Pi(n)}}{1-\delta^D\frac{\Pi(m)}{\Pi(n)}-b(1-\theta)(1-\delta^D)},\max\left(\gamma^{coop},\frac{1}{1-b(1-\delta^D)}\right)\right)$, which implies that $\delta^{coop}<\delta^{Nash}$ if $\gamma\in\left(\frac{1-\delta^D\frac{\Pi(m)}{\Pi(n)}}{1-\delta^D\frac{\Pi(m)}{\Pi(n)}-b(1-\theta)(1-\delta^D)},\frac{1}{1-b(1-\delta^D)}\right)$. This interval is non-empty as long as $\delta^D<\theta\frac{\Pi(n)}{\Pi(m)}$.

Proof of Remark 1

We now assume that we are in the first best, so that $x^*=y^*=m$ and $\gamma^{coop}=\left(1-b+b\theta\frac{\Pi(n)}{\Pi(m)}\right)^{-1}$, further we still have that $\delta^D<\theta\frac{\Pi(n)}{\Pi(m)}$. Then we get that as long as $\gamma\in\left(\gamma^{Nash},\gamma^{coop}\right)$, then $\delta^{Nash}>\delta^{coop}$ if and only if $\gamma>\frac{1}{1-b(1-\theta)\frac{1-\delta^D}{1-\delta^D\frac{\Pi(m)}{\Pi(n)}}}$ (using (B.60)). Furthermore if $\gamma>\gamma^{coop}$, then $\delta^{Nash}>\delta^{coop}$ if and only if $\gamma<\frac{1}{1-b(1-\delta^D)}$ (using (B.59)). Further with $\delta^D<\theta\frac{\Pi(n)}{\Pi(m)}$, we must have that $\frac{1}{1-b(1-\theta)\frac{1-\delta^D}{1-\delta^D\frac{\Pi(m)}{\Pi(n)}}}<\gamma^{coop}<\frac{1}{1-b(1-\delta^D)}$, so that we obtain:

$$\delta^{Nash}>\delta^{coop}\Longleftrightarrow\frac{1}{1-b\left(1-\delta^D\right)\frac{1-\theta}{1-\delta^D\frac{\Pi(m)}{\Pi(n)}}}<\gamma<\frac{1}{1-b\left(1-\delta^D\right)}.$$

B.5 Proof of Remark 2

Denote by B_t the number of producers who do not know a good match at the start of period. We obtain the law of motion

$$B_{t+1}=\left(1-\delta^D\right)bB_t+\delta^DN_t+N_{t+1}-N_t.$$

Indeed, among the producers that were in the same situation in the last period, only a fraction $1-\delta^D$ survived and of those a fraction b met a bad match. Moreover, the new producers, namely those that correspond to new products plus those that replace producers who died, also count as producers who do not know a good match at the beginning of the period. The share of producers who do not know a good match then obeys:

$$\omega_{t+1} = 1 - \frac{\left(1-\delta^D\right)(1-b\omega_t)}{1+g_N},$$

so that its steady-state value is given by:

$$\omega = \frac{g_N + \delta^D}{1 + g_N - b\left(1-\delta^D\right)}.$$

Using this expression in (21), we get that for $\gamma \in \left(\gamma^{Nash}, \gamma^{coop}\right)$,

$$\begin{aligned} Z_{Nash} - Z_{coop} = & \frac{1-b}{1+g_N - b\left(1-\delta^D\right)} \left(1-\delta^D\right)\left(1-b+b\theta-\gamma^{-1}\right)\Pi(n) \\ & + \frac{g_N+\delta^D}{1+g_N-b\left(1-\delta^D\right)}(1-b)\left(\left(1-\gamma^{-1}\right)\Pi(n) - \left(\Pi(x^*) - \gamma^{-1}\Pi(y^*)\right)\right). \end{aligned}$$

Therefore if $x^* = y^* = m$, we have:

$$\begin{gathered} Z_{Nash} > Z_{coop} \Longleftrightarrow \\ \left(1-\delta^D\right)\left(1-b+b\theta-\gamma^{-1}\right)\Pi(n) - \left(1-\gamma^{-1}\right)\delta^D\left(\Pi(m) - (\Pi(n))\right) > g_N\left(1-\gamma^{-1}\right)\left(\Pi(m) - \Pi(n)\right). \end{gathered}$$

This expressions clearly shows tat the higher is g_N, the more difficult is it to get $\delta^{Nash} > \delta^{coop}$.

Similarly, if $\gamma > \gamma^{coop}$, then using (15), (16), 17) and the new expression of ω, we obtain:

$$\begin{gathered} Z_{Nash} - Z_{coop} = \\ \tfrac{1-b}{1+g_N-b(1-\delta^D)} \left(\begin{array}{c} \left(\gamma^{-1}\Pi(y^*) - \left(1-b\left(1-\delta^D\right)\right)\Pi(x^*)\right) - \left(\gamma^{-1} - 1 + b\left(1-\delta^D\right)\right)\Pi(n) \\ -\left(\left(\Pi(x^*) - \gamma^{-1}\Pi(y^*)\right) - \left(1-\gamma^{-1}\right)\Pi(n)\right) g_N \end{array} \right) \end{gathered}$$

In the special case where $x^* = y^* = m$, this translates into

$$Z_{Nash} > Z_{coop} \Longleftrightarrow \gamma^{-1} - 1 + b\left(1-\delta^D\right) - g_N\left(1-\gamma^{-1}\right) > 0,$$

so that in that case too, a higher growth rate g_N makes it more difficult to get $\delta^{Nash} > \delta^{coop}$.

B.6 Proof of Proposition 9

In this appendix we consider the case where the strategy of suppliers is to punish the producer - by playing the Nash strategy - if we he switches to an innovator that turns out to be a bad match. We derive expression (A.18) in the special case in which the expected value of a new relationship is higher than remaining with a non-cooperating good match, such that if the innovator turns out to be a bad match the producer will seek out a new supplier rather than stick with the old one.

Compare to the situation in Appendix A.1, if the producer switches the old supplier loses all its value, hence $V_A^s = 0$. The producer will now switch if and only if:

$$V_I^{T,g} > W_1^T, \tag{B.61}$$

that is the total value of a new relationship with the innovator is higher than the total value of a relationship with the old supplier instead of (A.3). If the innovator turns out to be a bad match, the producer will try another new supplier in the following period, so the total value of the relationship with the innovator does not depend on whether the producer already knew a good match or not:

$$V_I^{T,g} = V_I^{T,b} = V_0.$$

Equation (A.4) is replaced by:

$$\begin{aligned} V_I^{T,g} = V_0^T &= (1-b)\,\Pi(x^*) + (1-b)\,\frac{1-\delta^D}{1+\rho}\left(\left(1-\delta^I\right)V_1^T + \delta^I\gamma W_1^T\right) \\ &+ b\theta\Pi(n) + b\frac{1-\delta^D}{1+\rho}\left(\left(1-\delta^I\right)V_0^T + \delta^I\gamma W_0^T\right). \end{aligned} \tag{B.62}$$

Using that (5) still holds, we get:

$$V_I^{T,g} - W_1^T = (1-b)\,\Pi(x^*) + b\theta\Pi(n) - \gamma^{-1}\Pi(y^*) - b\frac{1-\delta^D}{1+\rho}\left(\left(1-\delta^I\right)\left(V_1^T - V_0^T\right) + \delta^I\gamma\left(W_1^T - W_0^T\right)\right).$$

We use (5), (B.62) and (A.15) and (A.11) which both still hold to obtain

$$\left(1-\delta^I\right)\left(V_1^T - V_0^T\right) + \delta^I\gamma\left(W_1^T - W_0^T\right) = \frac{b\,(1+\rho)\left[\left(1-\delta^I\right)\left(\Pi(x^*) - \theta\Pi(n)\right) + \delta^I\left(\Pi(y^*) - \theta\Pi(n)\right)\right]}{1+\rho - b\left(1-\delta^D\right)\left(1-\delta^I+\gamma\delta^I\right)}. \tag{B.63}$$

Therefore, a producer in a good match will switch to the innovator if and only if (A.18) holds, which defines a γ^{coop2}. Note, that equation (A.18) differs from equation (13) only in the last term, it then follows that $\gamma^{coop2} > \gamma^{con} = \gamma^{Nash}$.

To show that the incentive to innovate is lower we need the fraction of the firms that are in good matches. In all cases, a producer in a bad match switch. If $\gamma < \gamma^{coop2}$ then only producers in bad matches in the cooperate equilibrium will switch, implying that in steady state (weakly) more producers will be in good matches in the cooperative equilibrium than in the contractible equilibrium. As the extra benefit for the innovator from contractibility is higher for good matches than bad matches, it follows that the incentive to innovate is higher in the contractible case, $\delta^{coop2} < \delta^{con}$.

Now, consider the case where $\gamma > \gamma^{coop2}$, such that good matches remain with the same producer when innovation takes place. Use the fact that in the contractible case a fraction $\tilde{\omega}^c = \frac{\delta^D}{1-b(1-\delta^D)}$ of producers will not be in good relationships, whereas in the noncontractible case a fraction $\tilde{\omega}^{nc} = \frac{\delta^D + b\delta^I\left(1-\delta^D\right)}{1-b(1-\delta^D)(1-\delta^I)}$ will not be in a good relationship. Inserting into the expressions for expected profits in the contractible and noncontractible case, respectively:

$$\left(\tilde{\omega}^c \left(\gamma - 1\right) \left(1 - b + b\theta\right) + \left(1 - \tilde{\omega}^c\right) \left(\left(1 - b + b\theta\right) \gamma - 1\right)\right) \Pi \left(m\right)$$

Cooperative case:

$$\begin{aligned}
&\tilde{\omega}^{nc} \left(\left(1-b\right)\left(\gamma \Pi\left(x^*\right) - \Pi\left(y^*\right)\right) + b\theta\left(\gamma - 1\right)\Pi\left(n\right)\right) \\
&+ \left(1 - \tilde{\omega}^{nc}\right)\left(\gamma\left(1-b\right)\Pi\left(x^*\right) + \gamma b\theta\Pi\left(n\right) - \Pi\left(y^*\right)\right) \\
&- \left(1 - \tilde{\omega}^{nc}\right) \frac{\gamma\left(1-\delta^D\right)b^2\left(\left(1-\delta^I\right)\left(\Pi\left(x^*\right) - \theta\Pi\left(n\right)\right) + \delta^I\left(\Pi\left(y^*\right) - \theta\Pi\left(n\right)\right)\right)}{1 + \rho - b\left(1-\delta^D\right)\left(1 - \delta^I + \delta^I\gamma\right)},
\end{aligned}$$

straightforward, but somewhat tedious algebra demonstrates that the condition $\delta^{coop2}\left(1 + \rho - b\left(1-\delta^D\right)\left(1 + \delta^{coop2}\left(\gamma - 1\right)\right)\right) < \gamma\left(1 - b\left(1 - \delta^D\right)\right)$ is sufficient to ensure that the incentive to innovate is lower: $\delta^{coop2} < \delta^{con}$.

B.7 Appendix: Proofs for section 4

B.7.1 Contractible and Nash cases

Here we show that the solution must satisfy $\delta_0^I = \delta_1^I$ and $V_0^T = V_1^T$. Taking the difference between (23) and (25) and using (24) one obtains that:

$$V_1^T - V_0^T = \frac{(1+\rho)\left(\psi\left(\delta_0^I\right) - \psi\left(\delta_1^I\right) + \left(\delta_1^I - \delta_0^I\right)\psi'\left(\delta_1^I\right)\right)}{1+\rho-(1-\delta_0^I)(1-\delta^D)}.$$

Next taking the difference between (26) and (24), one gets:

$$\psi'\left(\delta_0^I\right) - \psi'\left(\delta_1^I\right) = \frac{\left(1-\delta^D\right)\left(\psi\left(\delta_0^I\right) - \psi\left(\delta_1^I\right) + \left(\delta_1^I - \delta_0^I\right)\psi'\left(\delta_1^I\right)\right)}{1+\rho-(1-\delta_0^I)(1-\delta^D)}$$

The LHS is increasing in δ_0^I (since ψ is convex). On the RHS, the denominator is positive and increasing in δ_0^I, while the numerator is positive and decreasing in δ_0^I (once again thanks to the convexity of ψ), therefore the RHS is decreasing in δ_0^I. As a result this equation has a unique solution: $\delta_0^I = \delta_1^I$. In return, this ensures that $V_0^T = V_1^T$.

B.7.2 Cooperative case

Here, we describe the cooperative equilibrium in more details. To do so, we first spell out the strategies followed by the different agents, then we derive the results written in the main text, characterize the equilibrium level of cooperation and prove the existence of the equilibrium.

Strategies The strategies are characterized as follows:

- An augmented supplier refers to a supplier who has access to a technology which is higher than the fringe (if deviations have occurred there could be more than one augmented supplier).

- An augmented supplier with whom no deviation ever occurred offers an ex-ante transfer which allows her to capture the full surplus of the relationship over the second best option for the producer (namely going with a new supplier or choosing one with whom a deviation has occurred). If she is chosen by the producer, she invests x^*.

- Non-augmented suppliers with whom no deviation ever occurred, offers an ex-ante transfer which make them break even. If they are chosen, they invest x^* if there is an innovation and n otherwise.

- An augmented supplier with whom a deviation occurred, offers an ex-ante transfer that allows her to break even if either there are several suppliers in her situation, or she cannot offer the highest value for their joint relationship. She offers an ex-ante transfer that allows her to capture the surplus of a relationship with her over starting a new relationship if the producer does not know any other augmented supplier and if that surplus is positive. She invests n.

- A non-augmented supplier with whom a deviation occurred, offers an ex-ante transfer that allows her to break even and invests n.

- A producer chooses the supplier that offers him the highest value. If several suppliers offer the same value, he chooses one with whom the joint value is the highest.

- x^* is the highest level of cooperation within $(n, m]$ which does not violate the incentive constraint of the supplier.

- The innovation rate is chosen so as to maximize the joint value of the relationship.

Proof of Proposition 4 We first need to prove that $V_0^T < V_1^T$. To show that, we first use (27) to derive

$$V_0^T = \frac{(1+\rho)\left[\left(1-\delta_0^I\right)\Pi(n) + \delta_0^I\gamma\Pi(x^*) - \psi\left(\delta_0^I\right)\right] + \delta_0^I\gamma\left(1-\delta^D\right)V_1^T}{1+\rho-\left(1-\delta_0^I\right)\left(1-\delta^D\right)}. \tag{B.64}$$

Since $x^* > n$, we obtain:

$$V_0^T < \frac{(1+\rho)\left[\left(1-\delta_0^I+\delta_0^I\gamma\right)\Pi(x^*) - \psi\left(\delta_0^I\right)\right] + \delta_0^I\gamma\left(1-\delta^D\right)V_1^T}{1+\rho-\left(1-\delta_0^I\right)\left(1-\delta^D\right)}.$$

We denote the right hand side as $f\left(\delta_0^I\right)$. We then get that

$$V_0^T < f\left(\delta_0^I\right) \leq \max_{\delta} f(\delta).$$

To find $\max_{\delta} f(\delta)$, we take a first order condition and obtain that the solution $\widetilde{\delta}$ must satisfy:

$$\begin{aligned}
&\psi'\left(\widetilde{\delta}\right)\left(1+\rho-\left(1-\widetilde{\delta}\right)\left(1-\delta^D\right)\right)-\left(1-\delta^D\right)\psi\left(\widetilde{\delta}\right)\\
&=\left((\gamma-1)\left(\rho+\delta^D\right)-\left(1-\delta^D\right)\right)\Pi\left(x^*\right)+\gamma\frac{1-\delta^D}{1+\rho}V_1^T\left(\rho+\delta^D\right).
\end{aligned}$$

Since the LHS is increasing in $\widetilde{\delta}$ and the RHS is independent of it, this uniquely defines $\widetilde{\delta}$. Using (23) and (24) with $z = x^*$, we can check that $\widetilde{\delta} = \delta_1^I$ satisfies the previous equation. Then, using (23), one gets

$$V_0^T < f\left(\delta_1^I\right) = V_1^T.$$

Comparing (28) and (24) with $z = m, n$ or x^*, it is then direct that $\delta_0^{I,coop} > \delta_1^{I,coop}$ and that $\delta_0^{I,coop} > \delta^{I,Nash}$. Further, we get that if x^* is close to n, $\delta_0^{I,coop}$ is close to $\delta^{I,Nash}$ (and lower than $\delta^{I,cont}$), and if x^* is close to m, $\delta_0^{I,coop} > \delta^{I,cont}$.

The growth rate of the economy in the cooperative case depends on the share of producers who know an augmented supplier and their average productivity. Nevertheless, this growth rate must be larger in the cooperative case than in the Nash case (since the innovation rate is larger whether the producer knows an innovator or not). Similarly, as long as $\delta_0^{I,coop} < \delta^{I,cont}$ (which is true if x^* is close to n), growth is lower in the cooperative than contractible case. If on the other hand $\delta_0^{I,coop} > \delta^{I,cont}$ and $\delta_1^{I,coop} = \delta^{I,cont}$ (which is obtained if $x^* = m$), then the innovation rate is weakly higher (and in some line strictly higher) in the cooperative case than in the contractible one, leading to a higher growth rate in the former.

Characterizing the equilibrium Here we first characterize the equilibrium and then prove its existence. There are two possible cases, either after a deviation the producer stays with an augmented supplier which has deviated, or he ignores that firm and looks for a new supplier. Depending on whether the producer knows one or more deviating firms, his behavior may be different, because he can capture a different value from a relationship with such a non-cooperating good match. However, we demonstrate below that the number of known non-cooperating good match suppliers in fact does not matter for the producer's decision.

Case where the producer does not stick with a non-cooperating good match. In this section, we consider the case where a producer does not stick with a non-cooperating good match (regardless of the number of known non-cooperating good match suppliers). We characterize the level of cooperation x^*, the condition under which this scenario applies and demonstrate that if a producer does not stick with a non-cooperating good match if he only knows one of them, then he will not do so if he knows more than one of them either.

Characterizing x^*. Consider a producer and supplier on equilibrium path, then the incentive compatibility constraint for the supplier is given by

$$\varphi(x) \leq \frac{1-\delta^D}{1+\rho}\left(V_1^s - V_N^s\right), \tag{B.65}$$

where V_N^s is the value the innovator would capture at the beginning of the following period should a deviation occurs (so that the supplier would play the Nash level). Note that no γ term appear here because the technology level at the time of investment in x is the same as the technology level at the beginning of the following period. Since we have assumed that the producer would not want to work with the supplier after the deviation, we obtain $V_N^s = 0$, such that the incentive constraint is

$$\varphi(x) \leq \frac{1-\delta^D}{1+\rho} V_1^s.$$

Following the innovator's strategy, the value she captures from a relationship with the producer corresponds to the surplus over the producer's second best option. Namely we have

$$V_1^p = \gamma^{-1} V_0^{T,n} \text{ and } V_1^s = V_1^T - \gamma^{-1} V_0^{T,n}. \tag{B.66}$$

Indeed, at the beginning of a period, the second best option of the producer is to look for another supplier, whose technology is γ lower (recall that the V's are normalized by the supplier's technology here, which is why γ^{-1}appears). Importantly, if the producer where to switch to a new supplier, he would now be off-equilibrium path and knowing one innovator-non-cooperating good match. The value from a relationship in that situation $V_{0,n}^T$ may differ from the value V_0^T on equilibrium path.

In fact, we can write the law of motion:

$$V_0^{T,n} = -\psi\left(\delta_0^{I,n}\right) + \left(1-\delta_0^{I,n}\right)\left(\Pi\left(n\right) + \frac{1-\delta^D}{1+\rho}V_0^{T,n}\right) + \delta_0^{I,n}\gamma\left(\Pi\left(x_n^*\right) + \frac{1-\delta^D}{1+\rho}V_1^{T,n}\right). \tag{B.67}$$

$V_1^{T,n}$ denotes the joint value of a relationship with an innovator (non non-cooperating good match) when the producer knows an innovator-non-cooperating good match and x_n^* is the cooperation level in that case. If the new supplier fails to innovate, then the producer's situation does not change and by assumption he would then prefer to stay away from the innovator-non-cooperating good match (who plays Nash); so that the continuation value in that case is $V_0^{T,n}$. If on the other hand, an innovation occurs, the producer knows both an innovator-non non-cooperating good match and an innovator-non-cooperating good match and the joint value of the relationship is $V_1^{T,n}$. $\delta_0^{I,n}$ maximizes the joint value $V_0^{T,n}$ (when deriving the law of motion for any V_X^T below, the notation δ_X^I denotes the equilibrium innovation rate, which must maximize V_X^T).

To go further, we need to think about the IC constraint of an innovator when the producer knows an innovator who deviated. This can be written as

$$\varphi\left(x_n\right) \leq \frac{1-\delta^D}{1+\rho}\left(V_1^{s,n} - V_{2N}^s\right), \tag{B.68}$$

where the index n in $V_1^{s,n}$ indicates that the producer knows a non-cooperating good match and in V_{2N}^s that he knows at least 2. When the producer knows at least 2 innovator non-cooperating good match suppliers, Bertrand competition ensures that he captures the whole value, hence, in all cases we will have that $V_{2N}^s = 0$. Further with Bertrand competition, $V_1^{s,n} = V_1^{T,n} - \gamma^{-1}V_0^{T,2n}$, where $V_0^{T,2n}$ indicates the joint value of a new relationship when the producer knows at least 2 non-cooperating good match suppliers (indeed whether a producer knows two or more non-cooperating good match suppliers does not matter since with Bertrand competition he would receive the same offers of ex-ante transfers by the non-cooperating good match—namely, one which allows him to capture the whole value). This ensures that $V_1^{T,n}$ and x_n^* also apply when the producer knows more than 1 non-cooperating good match.

Then, since we have assumed that regardless of the number of non-cooperating good match suppliers the producer would rather keep looking for new suppliers, $V_0^{T,2n}$ must obey the same law of motion as $V_0^{T,n}$ given by (B.67). This ensures that $V_1^{s,n} = V_1^s$ so that $x_n^* = x^*$. In return we then obtain $V_0^{T,n} = V_0^T$ and $\delta_0^{I,n} = \delta_0^I$.

Using (23) with $z = x^*$, we get that

$$V_1^T = \frac{(1+\rho)\left(\left(1-\delta_1^I+\delta_1^I\gamma\right)\Pi\left(x^*\right)-\psi\left(\delta_1^I\right)\right)}{1+\rho-\left(1-\delta^D\right)\left(1-\delta_1^I+\delta_1^I\gamma\right)}, \tag{B.69}$$

while using (27) we get (B.64). Combining this equation with (B.69), we obtain:

$$\begin{aligned} V_1^s &= V_1^T - \gamma^{-1}V_0^T \\ &= \frac{1+\rho}{1+\rho-\left(1-\delta^D\right)\left(1-\delta_0^I\right)}\left(\begin{array}{c} \frac{\left(\rho+\delta^D\right)\left(\left(1-\delta_1^I+\delta_1^I\gamma\right)\Pi\left(x^*\right)-\psi\left(\delta_1^I\right)\right)}{1+\rho-\left(1-\delta^D\right)\left(1-\delta_1^I+\delta_1^I\gamma\right)} \\ -\gamma^{-1}\left(\left(1-\delta_0^I\right)\Pi\left(n\right)+\delta_0^I\gamma\Pi\left(x^*\right)-\psi\left(\delta_0^I\right)\right) \end{array}\right) \end{aligned}$$

Therefore we have that the IC constraint can be written as

$$\varphi\left(x^*\right) \leq \frac{1-\delta^D}{1+\rho-\left(1-\delta^D\right)\left(1-\delta_0^I\right)}\left(\begin{array}{c} \frac{\left(\rho+\delta^D\right)\left(\left(1-\delta_1^I+\delta_1^I\gamma\right)\Pi\left(x^*\right)-\psi\left(\delta_1^I\right)\right)}{1+\rho-\left(1-\delta^D\right)\left(1-\delta_1^I+\delta_1^I\gamma\right)} \\ -\gamma^{-1}\left(\left(1-\delta_0^I\right)\Pi\left(n\right)+\delta_0^I\gamma\Pi\left(x^*\right)-\psi\left(\delta_0^I\right)\right) \end{array}\right). \tag{B.70}$$

$x^* = m$ if (B.70) holds in that case, or x^* is such that (B.70) holds with equality

Condition under which the producer does not stay with the non-cooperating good match. We have assumed that the innovator would rather try a new supplier than stay with an innovator who has deviated. We need to check under which conditions, this is an equilibrium. To do that, we derive the joint value of a producer who knows a non-cooperating good match and decides to stay with her. This joint value obeys:

$$V_N^T = -\psi\left(\delta_N^I\right) + \left(1-\delta_N^I\right)\left(\Pi\left(n\right) + \frac{1-\delta^D}{1+\rho}\gamma^{-1}V_0^T\right) + \delta_N^I\left(\gamma\Pi\left(n\right) + \frac{1-\delta^D}{1+\rho}V_0^T\right). \tag{B.71}$$

If there is no innovation then in the following period, by assumption, the producer would rather try a new supplier (with a lower technology). If innovation occurs, the producer would rather try a new supplier as well (and the technology of that new supplier is the same as today). Moreover the strategy of a non-cooperating good match is to invest the Nash level n. The innovation rate must satisfy:

$$\psi'\left(\delta_N^I\right) = (\gamma-1)\left(\Pi\left(n\right) + \frac{1-\delta^D}{1+\rho}\gamma^{-1}V_0^T\right).$$

Since $x^* > n$ and $\gamma V_1^T - V_0^T > (\gamma-1)\gamma^{-1}V_0^T$, then it must be that $\delta_0^I > \delta_N^I$.

Using (B.71) and (B.64), we find that

$$
\begin{aligned}
&\gamma^{-1}V_0^T - V_N^T \\
&= \frac{\left(1+\rho-\left(1-\delta^D\right)\left(1-\delta_N^I+\delta_N^I\gamma\right)\right)}{1+\rho-\left(1-\delta^D\right)\left(1-\delta_0^I\right)}\left(\left(1-\delta_0^I\right)\gamma^{-1}\Pi\left(n\right)+\delta_0^I\Pi\left(x^*\right)-\gamma^{-1}\psi\left(\delta_0^I\right)+\delta_0^I\frac{1-\delta^D}{1+\rho}V_1^T\right) \\
&- \left(\left(1-\delta_N^I+\delta_N^I\gamma\right)\Pi\left(n\right)-\psi\left(\delta_N^I\right)\right).
\end{aligned}
$$

Therefore the strategies described form an equilibrium if x^* satisfies the IC constraint (with equality unless $x^* = m$) and the following condition is satisfied

$$
\frac{\left(1-\delta_0^I\right)\gamma^{-1}\Pi\left(n\right)+\delta_0^I\Pi\left(x^*\right)-\gamma^{-1}\psi\left(\delta_0^I\right)+\delta_0^I\frac{1-\delta^D}{1+\rho}V_1^T}{1+\rho-\left(1-\delta^D\right)\left(1-\delta_0^I\right)} \geq \frac{\left(1-\delta_N^I+\delta_N^I\gamma\right)\Pi\left(n\right)-\psi\left(\delta_N^I\right)}{1+\rho-\left(1-\delta^D\right)\left(1-\delta_N^I+\delta_N^I\gamma\right)}. \tag{B.72}
$$

Case where the producer would stay with a non-cooperating good match if he knows at least 2 of them but not if he only knows one of them. Here, we show that this case is impossible. Assume otherwise, then we still have $\gamma^{-1}V_0^{T,n} > V_N^T$ (a producer would rather try a new supplier than stick with a single non-cooperating good match) but $\gamma^{-1}V_0^{T,2n} < V_{2N}^T$: a producer would rather work with an innovator-non-cooperating good match than a new supplier when he knows at least two non-cooperating good match suppliers. Then, the value of starting a relationship with a new supplier for a producer who knows at least 2 innovator-non-cooperating good match suppliers, $V_0^{T,2n}$, obeys the following law of motion:

$$
V_0^{T,2n} = -\psi\left(\delta_0^{I,2n}\right)+\left(1-\delta_0^{I,2n}\right)\left(\Pi\left(n\right)+\frac{1-\delta^D}{1+\rho}\gamma V_{2N}^T\right)+\delta_0^{I,2n}\gamma\left(\Pi\left(x_{2n}^*\right)+\frac{1-\delta^D}{1+\rho}V_1^{T,2n}\right). \tag{B.73}
$$

If there is no innovation, then in the following period, the producer should revert back to choosing one of the two non-cooperating good match suppliers (by assumption). In that case, he will capture the full joint value of the relationship (because the two non-cooperating good match suppliers Bertrand compete). If there is an innovation, then the producer will now be working with an augmented supplier, while simultaneously knowing two innovator-non-cooperating good match suppliers. The level of cooperation x_{2n}^* could in principle be different from x_n^*.

The IC constraint that determines x_n^* is given by (B.68) with $V_N^{s,2n} = 0$ and $V_1^{s,n} = V_1^{T,n} - V_{2N}^T$: indeed, should the producer not stick with an innovator who has not deviated he could either try a new supplier or go to the non-cooperating good match he knows, but he would now know two non-cooperating good match suppliers. The latter is by assumption the second best option here.

The IC constraint that determines x_{2n}^{*} is:

$$\varphi\left(x_{2n}^{*}\right) \leq \frac{1-\delta^{D}}{1+\rho}\left(V_{1}^{s,2n}-V_{2N}^{s}\right).$$

We still have $V_{2N}^{s}=0$ and $V_{1}^{s,2n}=V_{1}^{T,2n}-V_{2N}^{T}$ as the second best option of the producer is to go with a non-cooperating good match now that he knows at least two of them. Hence we must have that $x_{2n}^{*}=x_{n}^{*}$, $\delta_{0}^{I,2n}=\delta_{0}^{I,n}$ and $V_{1}^{T,2n}=V_{1}^{T,n}$. This allows us to rewrite (B.73) as:

$$V_{0}^{T,2n}=-\psi\left(\delta_{0}^{I,2n}\right)+\left(1-\delta_{0}^{I,2n}\right)\left(\Pi\left(n\right)+\frac{1-\delta^{D}}{1+\rho}\gamma V_{2N}^{T}\right)+\delta_{0}^{I,2n}\gamma\left(\Pi\left(x_{n}^{*}\right)+\frac{1-\delta^{D}}{1+\rho}V_{1}^{T,n}\right). \tag{B.74}$$

The joint value of a relationship between a producer and an innovator-non-cooperating good match when the producer only knows one such non-cooperating good match, still obeys (B.71) but with $V_{0}^{T,n}$ instead of V_{0}^{T} (as we cannot establish that they are the same here), hence:

$$V_{N}^{T}=-\psi\left(\delta_{N}^{I}\right)+\left(1-\delta_{N}^{I}\right)\left(\Pi\left(n\right)+\frac{1-\delta^{D}}{1+\rho}\gamma^{-1}V_{0}^{T,n}\right)+\delta_{N}^{I}\left(\gamma\Pi\left(n\right)+\frac{1-\delta^{D}}{1+\rho}V_{0}^{T,n}\right). \tag{B.75}$$

Finally, the joint value of a relationship between a producer and an innovator-non-cooperating good match, when the producer knows at least 2 such non-cooperating good match suppliers is given by:

$$V_{2N}^{T}=-\psi\left(\delta_{2N}^{I}\right)+\left(1-\delta_{2N}^{I}\right)\left(\Pi\left(n\right)+\frac{1-\delta^{D}}{1+\rho}V_{2N}^{T}\right)+\delta_{2N}^{I}\left(\gamma\Pi\left(n\right)+\frac{1-\delta^{D}}{1+\rho}V_{0}^{T,n}\right). \tag{B.76}$$

If no innovation occurs then the best option is by assumption to stick with a non-cooperating good match. If innovation occurs though, the producer would now know only 1 innovator-non-cooperating good match (the one with whom he has just worked), indeed the other non-cooperating good match suppliers will not have access to the frontier technology. By assumption, in the following period, the producer should then try a new supplier (whose technology predates the last innovation) instead of staying with the single innovator-non-cooperating good match.

From (B.67) and (B.74) and using that $\delta_0^{I,2n}$ maximizes $V_0^{T,2n}$, we have

$$V_0^{T,2n} - V_0^{T,n} \geq \left(1-\delta_0^{I,n}\right)\frac{1-\delta^D}{1+\rho}\left(\gamma V_{2N}^T - V_0^{T,n}\right) > 0, \tag{B.77}$$

by assumption. Using that $\gamma^{-1}V_0^{T,n} > V_N^T$ and that $V_{2N}^T > \gamma^{-1}V_0^{T,2n}$, we get $V_{2N}^T > V_N^T$.

Moreover, using that δ_N^I maximizes the RHS in (B.75), we get:

$$V_N^T \geq -\psi\left(\delta_{2N}^I\right) + \left(1-\delta_{2N}^I\right)\left(\Pi(n) + \frac{1-\delta^D}{1+\rho}\gamma^{-1}V_0^{T,n}\right) + \delta_{2N}^I\left(\gamma\Pi(n) + \frac{1-\delta^D}{1+\rho}V_0^{T,n}\right). \tag{B.78}$$

Further using that $\gamma^{-1}V_0^{T,n} > V_N^T$, we obtain:

$$V_N^T > -\psi\left(\delta_{2N}^I\right) + \left(1-\delta_{2N}^I\right)\left(\Pi(n) + \frac{1-\delta^D}{1+\rho}V_N^T\right) + \delta_{2N}^I\left(\gamma\Pi(n) + \frac{1-\delta^D}{1+\rho}V_0^{T,n}\right). \tag{B.79}$$

Take the difference between (B.76) and (B.79) to get

$$V_N^T - V_{2N}^T > 0. \tag{B.80}$$

This contradicts the result we obtained above, which shows that this case is impossible: if a producer would rather look for a new supplier when he knows one innovator-non-cooperating good match, he should also do so if he knows more than one innovator-non-cooperating good match.

Case where the producer stays with the non-cooperating good match In this section, we consider the case where a producer sticks with an innovator-non-cooperating good match if he does not know a non-non-cooperating good match innovator. We assume that the producer does so regardless of the number of known non-cooperating good match suppliers. As above, we first characterize the level of cooperation x^* and the condition under which this scenario applies. Then we show that if a producer prefers a non-cooperating good match to trying a new supplier who does not have access to the latest technology when he knows one non-cooperating good match, then he must also prefer doing so when he knows several non-cooperating good match suppliers.

Characterizing the level of cooperation x^*.

Consider a producer who is matched with an innovator with whom no deviation ever occurred and further assume that the producer does not know any innovator-non-cooperating good match (this corresponds to what happens after the first successful innovation on equilibrium path). The IC constraint faced by the innovator still obeys (B.65). V_1^p and V_1^s are still determined by (B.66) since the second best option of the producer is to start a new relationship with a firm with an inferior technology but now knowing one non-cooperating good match. The difference is that $V_N^s \neq 0$: by assumption in case of a deviation the producer would rather stick with an innovator who has deviated than try a new supplier with an inferior technology (the only outside option here). Hence we obtain that

$$V_N^s = V_N^T - \gamma^{-1} V_{0,n}^T \text{ and } V_N^p = \gamma^{-1} V_0^{T,n}.$$

Therefore the IC constraint can be written as

$$\varphi\left(x^*\right) \leq \frac{1-\delta^D}{1+\rho}\left(V_1^T - V_N^T\right).$$

Note that V_N^T obeys (23) with $z = n$, and that the joint value of a relationship with a non-cooperating good match when the producer knows at least 2 non-cooperating good match suppliers (V_{2N}^T) also obeys the same law of motion (since the producer would always prefer to stick with an innovator-non-cooperating good match rather than trying a new supplier with an inferior technology). Hence we get $V_N^T = V_{2N}^T = V_1^{T,Nash}$ with $\delta_N^I = \delta_{2N}^I = \delta^{I,Nash}$, so that

$$V_N^T = \frac{(1+\rho)\left(\left(1-\delta_N^I+\delta_N^I\gamma\right)\Pi(n) - \psi\left(\delta_N^I\right)\right)}{1+\rho-\left(1-\delta^D\right)\left(1-\delta_N^I+\delta_N^I\gamma\right)} \tag{B.81}$$

Further V_1^T is still given by (B.69), combined with (B.81), we obtain that the level of cooperation x^* is characterized by the IC constraint:

$$\varphi\left(x^*\right) \leq \left(1-\delta^D\right)\left(\frac{\left(1-\delta_1^I+\delta_1^I\gamma\right)\Pi\left(x^*\right)-\psi\left(\delta_1^I\right)}{1+\rho-\left(1-\delta^D\right)\left(1-\delta_1^I+\delta_1^I\gamma\right)} - \frac{\left(1-\delta_N^I+\delta_N^I\gamma\right)\Pi(n)-\psi\left(\delta_N^I\right)}{1+\rho-\left(1-\delta^D\right)\left(1-\delta_N^I+\delta_N^I\gamma\right)}\right). \tag{B.82}$$

Therefore $x^* = m$ if (B.82) holds in that case and otherwise x^* is such that (B.82) holds with equality.

Condition under which the producer prefers an innovator-non-cooperating good match to a new supplier.

We want to derive the conditions under which it is indeed the case that $V_N^T > \gamma^{-1} V_0^{T,n}$.

To do that, we first need to characterize $V_0^{T,n}$. The incentive constraint faced by an augmented supplier who is in a relationship with a producer that already knows a non-cooperating good match is given by (B.68). Since the producer knows already one innovator-non-cooperating good match, then $V_{2N}^s = 0$. Furthermore $V_1^{s,n} = V_1^{T,n} - V_N^T$, since the producer's outside option next period is to start with one of the non-cooperating good match (with whom he would capture the entire surplus), and as explained above $V_{2N}^T = V_N^T$. Therefore, the IC constraint is still given by (B.82) with x_n^* instead of x^*, which implies that $V_1^{T,n} = V_1^T$, $\delta_1^{I,n} = \delta_1^I$ and $x_n^* = x^*$.

The value of a new relationship when a producer already knows exactly one innovator-non-cooperating good match obeys the following law of motion:

$$V_0^{T,n} = -\psi\left(\delta_0^{I,n}\right) + \left(1-\delta_0^{I,n}\right)\left(\Pi(n) + \frac{1-\delta^D}{1+\rho}V_0^{T,n}\right) + \delta_0^{I,n}\gamma\left(\Pi(x^*) + \frac{1-\delta^D}{1+\rho}V_1^T\right). \tag{B.83}$$

Indeed, the continuation value of the producer if there is no innovation is $V_0^{T,n}$ only since the non-cooperating good match would capture the surplus of the relationship; while, following an innovation, the level of cooperation is given by x^*. We then directly obtain that $\delta_0^{I,n} = \delta_0^I$ and that $V_0^{T,n} = V_0^T$.

Using (B.64) and (B.81), one gets

$$\gamma^{-1}V_0^T - V_N^T = \tag{B.84}$$

$$(1+\rho)\left(\frac{\left(\left(1-\delta_0^I\right)\gamma^{-1}\Pi(n) + \delta_0^I\Pi(x^*) - \gamma^{-1}\psi\left(\delta_0^I\right)\right) + \delta_0^I\frac{1-\delta^D}{1+\rho}V_1^T}{1+\rho-\left(1-\delta_0^I\right)\left(1-\delta^D\right)} - \frac{\left(1-\delta_N^I+\delta_N^I\gamma\right)\Pi(n) - \psi\left(\delta_N^I\right)}{1+\rho-\left(1-\delta_N^I+\delta_N^I\gamma\right)\left(1-\delta^D\right)}\right)$$

To ensure that producers want to stick with the innovator-non-cooperating good match, the RHS of this equation must be weakly negative. In other words, we obtain an equilibrium provided that the weak opposite of (B.72) holds.

Ruling out the possibility for the producer not to stay with a non-cooperating good match if he knows at least 2 of them. Furthermore, the value of starting a relationship with a new supplier when the producer knows at least two innovator-non-cooperating good match suppliers obeys the following law of motion:

$$V_0^{T,2n} = -\psi\left(\delta_0^{I,2n}\right) + \left(1-\delta_0^{I,2n}\right)\left(\Pi(n) + \frac{1-\delta^D}{1+\rho}\max\left(\gamma V_N^T, V_0^{T,2n}\right)\right) + \delta_0^{I,2n}\gamma\left(\Pi(x^*) + \frac{1-\delta^D}{1+\rho}V_1^T\right).$$

Indeed, if no innovation occurs the producer will then decide whether to try a supplier without the frontier technology or a non-cooperating good match. Since there are several of each, the producer captures the whole value of the relationship. We either have $V_N^T > \gamma^{-1} V_0^{T,2n}$, that is a non-cooperating good match is preferred to a new supplier with an outdated technology regardless of the number of non-cooperating good match suppliers—which is what we have assumed; or $V_N^T < \gamma^{-1} V_0^{T,2n}$. In that case though, $V_0^{T,2n}$ obeys the same law of motion as $V_0^{T,n}$, namely (B.83), we then get $V_0^{T,2n} = V_0^{T,n} < \gamma V_N^T$: which is a contradiction. If the producer prefers a non-cooperating good match to a relationship with a supplier with a non-frontier technology, he will do so regardless of the number of non-cooperating good match suppliers (still we will have $V_0^{T,2n} \neq V_0^{T,n}$ and $\delta_0^{I,2n} \neq \delta_0^{I,n}$).

Existence We have derived necessary conditions for the existence of an equilibrium obtained with the strategies we described. It is direct to check that these are also sufficient conditions. Therefore, the last thing to do is to ensure that there exists a x^* such that all conditions are satisfied. That is we must show that either i) m satisfies the IC constraint (B.70) together with (B.72) or m satisfies (B.82) together with the opposite of (B.72); or ii) the IC constraint (B.70) binds and (B.72) holds or the IC constraint (B.82) binds and the opposite of (B.72) holds.

To do that we first show that the IC constraint does not bind when x^* is close to n. For x^* close to n, $\delta_1^I \approx \delta_0^I \approx \delta^{I,Nash}$ and we obtain that $V_1^T \approx V_0^T \approx V_1^{T,Nash} = V_N^T$. Therefore the opposite of (B.72) holds. Further for x^* close to but above n, (B.82) holds as

$$\begin{aligned}\frac{\left(1-\delta_1^I+\delta_1^I\gamma\right)\Pi\left(x^*\right)-\psi\left(\delta_1^I\right)}{1+\rho-\left(1-\delta^D\right)\left(1-\delta_1^I+\delta_1^I\gamma\right)} &\geq \frac{\left(1-\delta_N^I+\delta_N^I\gamma\right)\Pi\left(x^*\right)-\psi\left(\delta_N^I\right)}{1+\rho-\left(1-\delta^D\right)\left(1-\delta_N^I+\delta_N^I\gamma\right)} \\ &> \frac{\left(1-\delta_N^I+\delta_N^I\gamma\right)\Pi\left(n\right)-\psi\left(\delta_N^I\right)}{1+\rho-\left(1-\delta^D\right)\left(1-\delta_N^I+\delta_N^I\gamma\right)}.\end{aligned}$$

The first inequality uses that δ_1^I maximizes $\frac{(1-\delta+\delta\gamma)\Pi(x^*)-\psi(\delta)}{1+\rho-(1-\delta^D)(1-\delta+\delta\gamma)}$ and the second that $\Pi\left(x\right)$ is increasing over (n,m).

If there is a x^* such that (B.82) binds while the opposite of (B.72) holds, then an equilibrium exists. Otherwise, there must exist a $\overline{x}$ such that (B.72) holds with equality at $\overline{x}$ (and holds strictly above $\overline{x}$) with (B.82) not binding over $[n,\overline{x}]$. Note that at $\overline{x}$, $V_N^T = \gamma^{-1} V_0^T$, and as result (B.70) and (B.82) are identical. Therefore, by continuity (B.70) still does not bind for x just above $\overline{x}$. By continuity, an equilibrium exists: either the IC constraint never binds and the appropriate equilibrium condition at m ((B.72) or its opposite) is satisfied, or the IC constraint binds and the appropriate equilibrium condition holds.

B.8 Appendix: Combined model

B.8.1 Model description

The model combined the baseline model of general innovation of section 2 with the relationship-specific innovation model of section 4. As in the latter model, the frontier technology can now vary in each line. A general innovation pushes the frontier by a factor γ^A in each line, but it is imitated after one period (so that all suppliers get access to the frontier technology in that line at the beginning of the next period). A relationship specific innovation pushes the frontier by a factor γ^B in the line in which it occurs. The innovator is the only one with this technology until a further general or relationship specific innovation, in which case the innovator gets access to the new frontier technology and all other firms get access to the previous frontier technology. The two types of innovations do not occur in the same period, instead a period is either one where general innovation may happen (with probability ν) or one where relationship specific innovation may happen (with probability $1-\nu$). For relationship specific innovation the innovation cost is $\psi^B\left(\delta^B\right)A_j$ where ψ^B is a convex function of the innovation rate δ^B and A_j is the pre-innovation frontier technology in line j. For general innovation, the innovation cost is $\psi^A\left(\delta^A\right)\widetilde{A}$ where ψ^A is a convex function of the innovation rate δ^B and $\widetilde{A}$ is the average pre-innovation frontier technology in the economy. To ensure a steady-state, we assume that the potential innovator cannot observe when the last general innovation occurs.

As in the baseline model, there are good and bad matches. Cooperation is only possible in good matches, moreover, we also assume that relationship-specific innovation is impossible in bad matches.[62] The nature of a match is revealed before relationship-specific innovation or investment are undertaken but after a producer has decided to start working with a supplier (so after a potential general innovation). If a relationship-specific innovation occurs in a good match supplier, we will refer to the supplier as an "augmented" good match until she is not the frontier supplier (by opposition we will talk of a "regular" good match otherwise). As in the relationship-specific model, if a

[62]This is a simplifying assumption, without it the innovation rate would simply be lower in bad matches and we would still obtain very similar results.

producer dies, he is replaced by a new producer and for that line the technology level (pre-innovation) is equal to the average technology level in the economy (pre-innovation). Note that the baseline model is obtained in the specific case where $\nu = 1$ and the relationship-specific innovation only model is obtained for $\nu = 0$ and $b = 0$.

We look at a cooperative equilibrium which has the same characteristics as that of sections 2 and 4. In bad matches (or after a deviation occurred in the personal history of the producer and the supplier), normalized investment level is n. Producers with a good match supplier can be in four different scenarios: 1) the good match supplier has access to the frontier technology and she is a "regular" good match; 2) the good match supplier has access to a frontier technology and is an "augmented" good match; 3) in a period where a general innovation occurred, the good match supplier does not have access to the frontier technology and she is an "outdated" good match. In case 1), the investment level in equilibrium is the same and denoted x_1^*, in case 2) the investment level is denoted x_2^* and in case 3) it is denoted y^*. The baseline model made it clear why we needed two different levels of investment depending on whether the supplier had access to the frontier technology or not. In addition, the level of cooperation may differ depending on whether a relationship-specific innovation was the most recent innovation in the line or not, since the producer outside option is different. This was the case in section 4 already: before the relationship-specific investment there was no cooperation and afterwards some cooperation, the difference is that here because of the presence of bad matches, some cooperation is possible right away; in line with that section we assume that $x_1^* \leq x_2^*$. Finally, we denote δ_1^B the relationship-specific innovation rate in periods where such innovations are possible in a "regular" good match, while δ_2^B denotes the same rate for an "augmented" good match (in section 4 the corresponding notations were δ_0^I and δ_1^I, here we move the subscripts to 1 and 2 to have notations that mirror those of the value functions).

We still assume that a supplier forgives a producer who tries out a general innovator if that innovator turns out to be bad match. Finally, we consider parameters for which off-equilibrium path the producer would rather try out a new supplier rather than staying with a non-cooperating good match supplier playing the Nash level of investment (or equivalently, we assume that a producer forgets the identity of previous good matches if he starts working with one). Figure 3 summarizes the model by providing a timeline.

To use the same notations as in both sections 2 and 4, I normalize value functions by the frontier technology just after a general innovation has occurred in general innovation

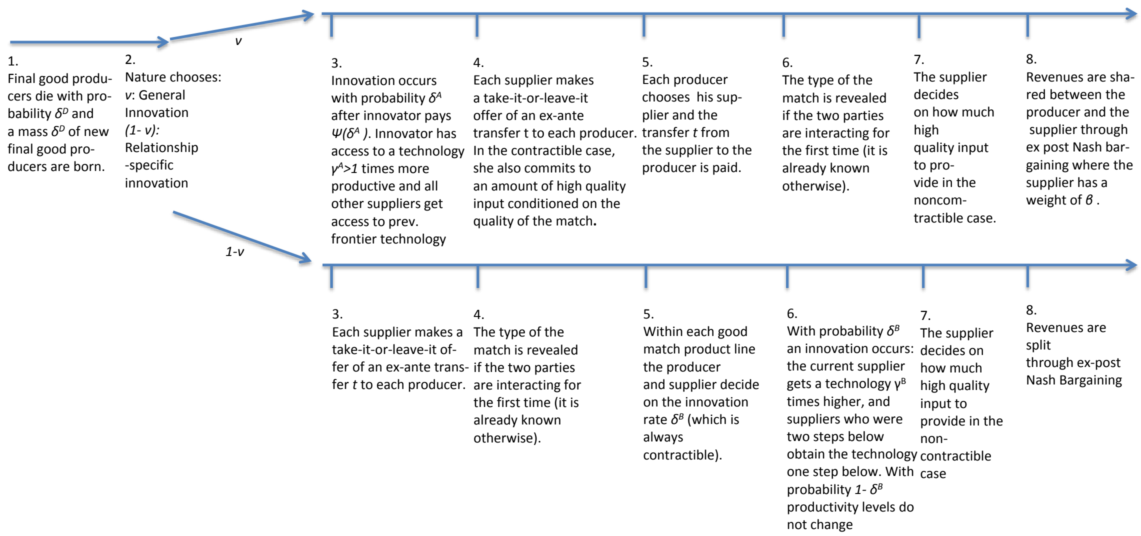


Figure 3: Timeline

periods but before a relationship-specific innovation has occurred in relationship-specific innovation periods. I use the notations V, W with indexes 0, 1 or A and superscripts s, p, T or g and b exactly as in the baseline model for periods where only general innovation is possible. The superscript 2 is used to denote an augmented good match (while 1 is for a regular good match). I denote the value functions by U in periods where relationship-specific innovations are possible. The superscripts s, p, T are used as before and the indexes $0, 1$ and 2 are used to denote a new match, a regular good match and an augmented good match.

B.8.2 Value functions and equilibrium description

We derive the value functions to show that the cooperative equilibrium exists and eventually describe the innovation incentives.

Relationship-specific innovation period First consider a relationship-specific innovation period. Then the (normalized) joint value of a producer together with a new supplier, U_0^T, obeys:

$$U_0^T = (1-b)\,U_1^T + b\theta\Pi\,(n) + b\frac{1-\delta^D}{1+\rho}\left[\nu\left(\left(1-\delta^A\right)V_0^T + \delta^A\gamma^A W_0^T\right) + (1-\nu)\,U_0^T\right]. \tag{B.85}$$

With probability $1-b$, the new supplier is a good match (which cannot have been augmented yet), leading to the joint value U_1^T. With probability b, it is a bad match. The flow of profits is then given by $\theta\Pi(n)$. In the next period, if the producer survives, there are two cases. With probability $1-\nu$, the next period is also one of relationship-specific innovations and the situation is the same as today (the producer value is then the full joint value U_0^T). With probability ν, the next period is one of potential general innovation leading to a producer value of V_0^T, with probability $1-\delta^A$, or $\gamma^A W_0^T$, with probability δ^A: if a general innovation does indeed occur the frontier moves by a factor γ^A but the producer only captures the value of a relationship with his second-best option available, namely starting a new relationship with an outdated good match.

Similarly the joint value of a relationship with a regular good match supplier obeys:

$$U_1^T = -\psi^B\left(\delta_1^B\right) + \left(1-\delta_1^B\right)\left(\Pi\left(x_1^*\right) + \frac{1-\delta^D}{1+\rho}\left[\nu\left(\left(1-\delta^A\right)V_1^T + \delta^A\gamma^A W_1^T\right) + (1-\nu)\,U_1^T\right]\right)$$
$$+\,\delta_1^B\gamma^B\left(\Pi\left(x_2^*\right) + \frac{1-\delta^D}{1+\rho}\left(\nu\left(\left(1-\delta^A\right)V_2^T + \delta^A\gamma^A W_1^T\right) + (1-\nu)\,U_2^T\right)\right). \tag{B.86}$$

In the current period, the supplier invests $\psi^B\left(\delta_1^B\right)$ (after having learned the nature of the match). The innovation fails with probability $1-\delta_1^B$, in which case the investment level will be x_1^*. With probability δ_1^B, the innovation succeeds, the frontier moves by a factor γ^B, the investment level is x_2^* and the match becomes an augmented good match. Note that if in the next period a general innovation does occur (which happens with probability $\nu\delta^A$), then the augmented good match supplier loses her advantage since she ceases to be the frontier supplier for that line and the joint value is $\gamma^B\gamma^A W_1^T$ or $\gamma^A W_1^T$ depending on whether the relationship specific innovation occurs today or not.

The joint value of a relationship with an augmented good match obeys:

$$U_2^T = -\psi\left(\delta_2^B\right) + \left(1-\delta_2^B+\delta_2^B\gamma^B\right)\left(\Pi\left(x_2^*\right) + \frac{1-\delta^D}{1+\rho}\left(\nu\left(\left(1-\delta^A\right)V_2^T + \delta^A\gamma^A W_1^T\right) + (1-\nu)\,U_2^T\right)\right), \tag{B.87}$$

if the innovation succeeds the frontier moves by a factor γ^B but nothing else changes.

Within a regular good match, the incentive compatibility constraint can be written as:

$$\varphi\left(x_1^*\right) \leq \frac{1-\delta^D}{1+\rho}\left(\nu\left(\left(1-\delta^A\right)V_1^s + \delta^A\gamma^A W_1^s\right) + (1-\nu)\,U_1^s\right). \tag{B.88}$$

In the next period, a supplier who cooperate captures U_1^s if it is a period where relationship-specific innovation may occurs, while she captures V_1^s in a general innovation period without an actual general innovation and W_1^s (with a higher technology level) in a period where a general innovation did occur. Similarly in an augmented good match, the incentive compatibility constraint can be written as:

$$\varphi\left(x_2^*\right) \leq \frac{1-\delta^D}{1+\rho}\left(\nu\left(\left(1-\delta^A\right) V_2^s+\delta^A \gamma^A W_1^s\right)+(1-\nu) U_2^s\right). \tag{B.89}$$

In equilibrium, the innovation rates δ_1^B and δ_2^B maximize U_1^T and U_2^T so that the following first order conditions must hold:

$$\begin{aligned}\psi'\left(\delta_1^B\right) = {} & \gamma^B\left(\Pi\left(x_2^*\right)+\frac{1-\delta^D}{1+\rho}\left(\nu\left(\left(1-\delta^A\right) V_2^T+\delta^A \gamma^A W_1^T\right)+(1-\nu) U_2^T\right)\right) \\ & -\left(\Pi\left(x_1^*\right)+\frac{1-\delta^D}{1+\rho}\left(\nu\left(\left(1-\delta^A\right) V_1^T+\delta^A \gamma^A W_1^T\right)+(1-\nu) U_1^T\right)\right).\end{aligned} \tag{B.90}$$

$$\psi'\left(\delta_2^B\right)=\left(\gamma^B-1\right)\left(\Pi\left(x_2^*\right)+\frac{1-\delta^D}{1+\rho}\left(\nu\left(\left(1-\delta^A\right) V_2^T+\delta^A \gamma^A W_1^T\right)+(1-\nu) U_2^T\right)\right). \tag{B.91}$$

General innovation period where innovation failed We turn to the case of a general innovation period where innovation failed. Such a period is identical to the previous case except that relationship specific innovations are ruled out. Yet since such an innovation is undertaken before the relationship specific investment is undertaken, its absence does not affect investment levels (conditional on the match being regular or augmented at the time of the investment level). Therefore we get that the joint values with a new supplier, a regular good match and an augmented good match obey (these equations are the pendant to (B.85), (B.86) and (B.87) in the previous case):

$$V_0^T=(1-b) V_1^T+b \theta \Pi(n)+b \frac{1-\delta^D}{1+\rho}\left[\nu\left(\left(1-\delta^A\right) V_0^T+\delta^A \gamma^A W_0^T\right)+(1-\nu) U_0^T\right]; \tag{B.92}$$

$$V_1^T=\Pi\left(x_1^*\right)+\frac{1-\delta^D}{1+\rho}\left(\nu\left(\left(1-\delta^A\right) V_1^T+\delta^A \gamma^A W_1^T\right)+(1-\nu) U_1^T\right); \tag{B.93}$$

$$V_2^T=\Pi\left(x_2^*\right)+\frac{1-\delta^D}{1+\rho}\left(\nu\left(\left(1-\delta^A\right) V_2^T+\delta^A \gamma^A W_1^T\right)+(1-\nu) U_2^T\right). \tag{B.94}$$

Furthermore, the incentive compatibility constraints are still given by (B.88) and (B.89).

Equilibrium properties: $x_2^* \geq x_1^*$ **and** $\delta_2^B \leq \delta_1^B$ We are here going to show that the equilibrium can indeed feature $x_2^* \geq x_1^*$ and that we must have $\delta_2^B \leq \delta_1^B$. To do that first note that Bertrand competition implies:

$$U_1^s = U_1^T - U_0^T \text{ and } U_2^s = U_2^T - \frac{1}{\gamma^B} U_0^T. \tag{B.95}$$

$$V_1^s = V_1^T - V_0^T \text{ and } V_2^s = V_2^T - \frac{1}{\gamma^B} V_0^T. \tag{B.96}$$

Moreover, combining (B.86) with (B.93) and (B.87) with (B.94), we get:

$$U_1^T = -\psi\left(\delta_1^B\right) + \left(1 - \delta_1^B\right) V_1^T + \delta_1^B \gamma^B V_2^T, \tag{B.97}$$

$$U_2^T = -\psi\left(\delta_2^B\right) + \left(1 - \delta_2^B + \delta_2^B \gamma^B\right) V_2^T. \tag{B.98}$$

Using (B.93) and (B.94), we can rewrite (B.90) and (B.91) as:

$$\psi'\left(\delta_1^B\right) = \gamma^B V_2^T - V_1^T \text{ and } \psi'\left(\delta_2^B\right) = \left(\gamma^B - 1\right) V_2^T. \tag{B.99}$$

Using (B.93), (B.94) and (B.97), (B.98) and then (B.99), we can derive:

$$\begin{aligned} &\left(V_2^T - V_1^T\right)\left(1 - \frac{1-\delta^D}{1+\rho}\left(\nu\left(1-\delta^A\right) + (1-\nu)\left(1-\delta_1^B\right)\right)\right) \\ &= \Pi\left(x_2^*\right) - \Pi\left(x_1^*\right) + (1-\nu)\frac{1-\delta^D}{1+\rho}\left(\psi\left(\delta_1^B\right) - \psi\left(\delta_2^B\right) - \left(\delta_1^B - \delta_2^B\right)\psi'\left(\delta_2^B\right)\right). \end{aligned} \tag{B.100}$$

Since ψ is convex, $\psi\left(\delta_1^B\right) - \psi\left(\delta_2^B\right) - \left(\delta_1^B - \delta_2^B\right)\psi'\left(\delta_2^B\right) \geq 0$, and since $x_2^* \geq x_1^*$, then we get that $V_2^T \geq V_1^T$. From (B.99), we then obtain that $\delta_1^B \geq \delta_2^B$ and the inequality is strict unless $x_1^* = x_2^*$.

Further, using (B.95) and (B.96), we get

$$\begin{aligned} &\nu\left(\left(1-\delta^A\right)V_2^s + \delta^A\gamma^A W_1^s\right) + (1-\nu)U_2^s - \left(\nu\left(\left(1-\delta^A\right)V_1^s + \delta^A\gamma^A W_1^s\right) + (1-\nu)U_1^s\right) \\ &= \nu\left(1-\delta^A\right)\left(V_2^T - \frac{1}{\gamma^B}V_0^T\right) + (1-\nu)\left(U_2^T - \frac{1}{\gamma^B}U_0^T\right) - \left[\nu\left(1-\delta^A\right)\left(V_1^T - V_0^T\right) + (1-\nu)\left(U_1^T - U_0^T\right)\right] \end{aligned}$$

Then plug in (B.97), (B.98) and (B.99) to get:

$$\begin{aligned} &\nu\left(\left(1-\delta^A\right)V_2^s + \delta^A\gamma^A W_1^s\right) + (1-\nu)U_2^s - \left(\nu\left(\left(1-\delta^A\right)V_1^s + \delta^A\gamma^A W_1^s\right) + (1-\nu)U_1^s\right) \\ &= \left(\nu\left(1-\delta^A\right) + (1-\nu)\left(1-\delta_1^B\right)\right)\left(V_2^T - V_1^T\right) + (1-\nu)\left(\psi\left(\delta_1^B\right) - \psi\left(\delta_2^B\right) - \left(\delta_1^B - \delta_2^B\right)\psi'\left(\delta_2^B\right)\right). \end{aligned}$$

As established before the last line is weakly positive as $V_2^T \geq V_1^T$, this in return implies that the IC constraint faced by suppliers in an augmented match is laxer than that faced by suppliers in a regular good match, which justifies that the equilibrium features $x_2^B \geq x_1^B$ (with equality if and only if $x_1^B = x_2^B = m$).

General innovation period when innovation succeeded Finally, we look at the value functions in a general innovation period when innovation succeeded. The joint value with a new outdated producer is given by:

$$W_0^T = (1-b)\,W_1^T + b\theta\frac{1}{\gamma^A}\Pi(n) + b\frac{1-\delta^D}{1+\rho}\left[\nu\left(\left(1-\delta^A\right)V_0^T + \delta^A\gamma^A W_0^T\right) + (1-\nu)\,U_0^T\right]. \tag{B.101}$$

The logic is the same as in the baseline model except that the continuation value must take into account that with probability $1-\nu$ the following period is one where relationship specific innovations are possible (so that n case of bad match this period, the continuation value is U_0^T). The joint value with an outdated good match is then given by:

$$W_1^T = \frac{1}{\gamma^A}\Pi(y^*) + \frac{1-\delta^D}{1+\rho}\left[\nu\left(\left(1-\delta^A\right)V_1^T + \delta^A\gamma^A W_1^T\right) + (1-v)\,U_1^T\right]. \tag{B.102}$$

Note that there are no outdated augmented good matches, because the last relationship specific innovation becomes freely available when the next general innovation becomes available. Then, using (B.93), we get:

$$V_1^T - W_1^T = \Pi(x_1^*) - \frac{1}{\gamma^A}\Pi(y^*). \tag{B.103}$$

The joint value of a relationship with the innovator when the producer does not know a good match obeys:

$$V_I^{T,b} = V_0^T$$

as in the baseline model. Therefore we still have through Bertrand Competition, using (B.92), (B.101) and (B.103) that:

$$V_I^{s,b} = V_0^T - W_0^T = (1-b)\left(\Pi(x_1^*) - \frac{1}{\gamma^A}\Pi(y^*)\right) + b\theta\left(1 - \frac{1}{\gamma^A}\right)\Pi(n). \tag{B.104}$$

The joint value of a relationship with the innovator when the producer knows a good match is given by:

$$V_I^{T,g} = (1-b)\,V_1^T + b\theta\Pi\,(n) + b\frac{1-\delta^D}{1+\rho}\left(\nu\left(\left(1-\delta^A\right)V_1^p + \delta^A\gamma^A W_1^p\right) + (1-\nu)\,U_1^p\right). \tag{B.105}$$

The logic is the same as in the baseline model, if the innovator turns out to be a bad match, then the producer can return to the previous good match supplier and earns V_1^p, W_1^p or U_1^p depending on the situation (note that the next period the outdated producer starts with the new frontier technology but cannot be an augmented good match).

As in the baseline model, the value of an outdated good match producer who is not picked by the producer is still positive and given by:

$$V_A^s = b\frac{1-\delta^D}{1+\rho}\left(\nu\left(\left(1-\delta^A\right)V_1^s + \delta^A\gamma^A W_1^s\right) + (1-\nu)\,U_1^s\right), \tag{B.106}$$

so that we must have $W_1^s \geq V_A^s$ and $W_1^p \leq W_1^T - V_A^s$. The producer will then switch suppliers whenever $V_I^{T,g} > W_1^T - V_A^s$. Combining (B.102), (B.105) and (B.106), we get:

$$V_I^{s,g} = \left((1-b)\,\Pi\,(x_1^*) + b\theta\Pi\,(n) - \frac{1}{\gamma^A}\Pi\,(y^*)\right)^+, \tag{B.107}$$

exactly as in the baseline model.

Further note that the incentive constraint in an outdated good match obeys:

$$\frac{1}{\gamma^A}\varphi\,(y^*) \leq \frac{1-\delta^D}{1+\rho}\left(\nu\left(\left(1-\delta^A\right)V_1^s + \delta^A\gamma^A W_1^s\right) + (1-\nu)\,U_1^s\right),$$

so that $y^* \geq x_1^*$ with equality if and only if the first best is achievable.

Recall that an innovator does not observe when the last general innovation took place. The expected value of a general innovator, normalized by the pre-innovation average technology ($\widetilde{A}_t$) in the economy is given by

$$Z = \omega V_I^{s,b}\frac{E\left(\widetilde{A}_{jt}|j\in b\right)}{\widetilde{A}_t} + (1-\omega)\,V_I^{s,g}\frac{E\left(\widetilde{A}_{jt}|j\in b\right)}{\widetilde{A}_t},$$

where $\widetilde{A}_{jt}$ denotes the pre-innovation frontier technology in line j and $j \in b$ means that the producer j does not know a (cooperating) good match. When a producer does not know a good match, then he cannot enjoy relationship specific innovations in his line.

Therefore there exists a function $\lambda\left(\delta_1^B, \delta_2^B, \delta^I\right) \geq 1$ such that in expectation

$$E\left(\widetilde{A}_{jt}|j \in g\right) = \lambda\left(\delta_1^B, \delta_2^B, \delta^A\right) E\left(\widetilde{A}_{jt}|j \in b\right),$$

and this function is increasing in δ_1^B,δ_2^B, since the more relationship specific innovation occur, the more the average technology level of good match producers will pull ahead. We can then rewrite

$$Z = \frac{\omega}{\omega + (1-\omega)\lambda\left(\delta_1^B, \delta_2^B, \delta^I\right)} V_I^{s,b} + \frac{(1-\omega)\lambda\left(\delta_1^B, \delta_2^B, \delta^I\right)}{\omega + (1-\omega)\lambda\left(\delta_1^B, \delta_2^B, \delta^I\right)} V_I^{s,g}, \tag{B.108}$$

and the innovation rate solves:

$$\psi'\left(\delta^A\right) = \gamma^A Z. \tag{B.109}$$

Relationship-specific innovation rate Here we rewrite the first order condition (B.99) in function of profits only, which will be useful to prove Proposition 5. Use (B.98) and (B.103) in (B.94) to get:

$$V_2^T = \frac{(1+\rho)\Pi(x_2^*) + \left(1-\delta^D\right)\left(\nu\delta^A\gamma^A\left(V_1^T - \left(\Pi(x_1^*) - \frac{1}{\gamma^A}\Pi(y^*)\right)\right) - (1-\nu)\psi\left(\delta_2^B\right)\right)}{1+\rho - \left(1-\delta^D\right)\left(\nu\left(1-\delta^A\right) + (1-\nu)\left(1-\delta_2^B + \delta_2^B\gamma^B\right)\right)}, \tag{B.110}$$

where we defined

$$\widetilde{\rho} \equiv \rho - \left(1-\delta^D\right)\nu\left(1-\delta^A\right),$$

to simplify our expressions ($1+\widetilde{\rho} > 0$ otherwise the value functions are not well defined).

Then use this equation together with (B.103) and (B.97) in (B.93) to get (after rearranging terms):

$$V_1^T \tag{B.111}$$

$$= \frac{\left[\begin{array}{c} \left(1+\widetilde{\rho} - \left(1-\delta^D\right)(1-\nu)\left(1-\delta_2^B + \delta_2^B\gamma^B\right)\right)\left((1+\rho)\Pi(x_1^*) - \left(1-\delta^D\right)(1-\nu)\psi\left(\delta_1^B\right)\right) \\ -\left(1-\delta^D\right)\left(1+\widetilde{\rho} - \left(1-\delta^D\right)(1-\nu)\left(1-\delta_2^B + \left(\delta_2^B - \delta_1^B\right)\gamma^B\right)\right)\nu\delta^A\gamma^A\left(\Pi(x_1^*) - \frac{1}{\gamma^A}\Pi(y^*)\right) \\ +\left(1-\delta^D\right)(1-\nu)\delta_1^B\gamma^B\left((1+\rho)\Pi(x_2^*) - \left(1-\delta^D\right)(1-\nu)\psi\left(\delta_2^B\right)\right) \end{array}\right]}{\left[\begin{array}{c} \left(1+\widetilde{\rho} - \left(1-\delta^D\right)\left[(1-\nu)\left(1-\delta_1^B\right) + \nu\delta^A\gamma^A\right]\right)\left(1+\widetilde{\rho} - \left(1-\delta^D\right)(1-\nu)\left(1-\delta_2^B + \delta_2^B\gamma^B\right)\right) \\ -(1-\nu)\delta_1^B\gamma^B\left(1-\delta^D\right)^2\nu\delta^A\gamma^A \end{array}\right]}.$$

Combining this expression with (B.110), we get:

$$V_2^T \tag{B.112}$$
$$= \frac{\left[\begin{array}{c} \left(1+\widetilde{\rho}-\left(1-\delta^D\right)\left[(1-\nu)\left(1-\delta_1^B\right)+\nu\delta^A\gamma^A\right]\right)\left((1+\rho)\Pi\left(x_2^*\right)-\left(1-\delta^D\right)(1-\nu)\psi\left(\delta_2^B\right)\right) \\ +\left(1-\delta^D\right)\nu\delta^A\gamma^A\left[\begin{array}{c}\left((1+\rho)\Pi\left(x_1^*\right)-\left(1-\delta^D\right)(1-\nu)\psi\left(\delta_1^B\right)\right) \\ -\left(1+\widetilde{\rho}-\left(1-\delta^D\right)(1-\nu)\left(1-\delta_1^B\right)\right)\left(\Pi\left(x_1^*\right)-\frac{1}{\gamma^A}\Pi\left(y^*\right)\right)\end{array}\right]\end{array}\right]}{\left[\begin{array}{c}\left(1+\widetilde{\rho}-\left(1-\delta^D\right)\left[(1-\nu)\left(1-\delta_1^B\right)+\nu\delta^A\gamma^A\right]\right)\left(1+\widetilde{\rho}-\left(1-\delta^D\right)(1-\nu)\left(1-\delta_2^B+\delta_2^B\gamma^B\right)\right) \\ -(1-\nu)\delta_1^B\gamma^B\left(1-\delta^D\right)^2\nu\delta^A\gamma^A\end{array}\right]}.$$

Combining this with (B.99), we get

$$\psi'\left(\delta_2^B\right)/\left(\gamma^B-1\right) \tag{B.113}$$
$$= \frac{\left[\begin{array}{c} \left(1+\widetilde{\rho}-\left(1-\delta^D\right)\left[(1-\nu)\left(1-\delta_1^B\right)+\nu\delta^A\gamma^A\right]\right)\left((1+\rho)\Pi\left(x_2^*\right)-\left(1-\delta^D\right)(1-\nu)\psi\left(\delta_2^B\right)\right) \\ +\left(1-\delta^D\right)\nu\delta^A\gamma^A\left[\begin{array}{c}\left((1+\rho)\Pi\left(x_1^*\right)-\left(1-\delta^D\right)(1-\nu)\psi\left(\delta_1^B\right)\right) \\ -\left(1+\widetilde{\rho}-\left(1-\delta^D\right)(1-\nu)\left(1-\delta_1^B\right)\right)\left(\Pi\left(x_1^*\right)-\frac{1}{\gamma^A}\Pi\left(y^*\right)\right)\end{array}\right]\end{array}\right]}{\left[\begin{array}{c}\left(1+\widetilde{\rho}-\left(1-\delta^D\right)\left[(1-\nu)\left(1-\delta_1^B\right)+\nu\delta^A\gamma^A\right]\right)\left(1+\widetilde{\rho}-\left(1-\delta^D\right)(1-\nu)\left(1-\delta_2^B+\delta_2^B\gamma^B\right)\right) \\ -(1-\nu)\delta_1^B\gamma^B\left(1-\delta^D\right)^2\nu\delta^A\gamma^A\end{array}\right]}$$

B.8.3 Proof of Proposition 5

Nash case The Nash equilibrium obeys the same equation except that the investment level is always n. Combining (B.99) and (B.100) implies that in the Nash equilibrium $\delta_1^{B,Nash} = \delta_2^{B,Nash} = \delta^{B,Nash}$ and $V_1^{T,Nash} = V_2^{T,Nash}$. Moreover, we get (using (B.113) but replacing x_1^*, x_2^* and y^* by n and δ_1^B, δ_2^B by a single $\delta^{B,Nash}$):

$$\frac{\psi'\left(\delta^{B,Nash}\right)}{\left(\gamma^B-1\right)}$$
$$= \frac{\left(1+\rho-\left(1-\delta^D\right)\nu\delta^{A,Nash}\left(\gamma^A-1\right)\right)\Pi(n)-\left(1-\delta^D\right)(1-\nu)\psi\left(\delta^{B,Nash}\right)}{1+\rho-\left(1-\delta^D\right)\left(\nu\left(1-\delta^{A,Nash}+\delta^{A,Nash}\gamma^A\right)+(1-\nu)\left(1-\delta^{B,Nash}+\delta^{B,Nash}\gamma^B\right)\right)}.$$

Establishing part b) We first establish part b) of the proposition. (B.99) implies that:

$$\psi'\left(\delta_1^B\right) = \psi'\left(\delta_2^B\right) + V_2^T - V_1^T.$$

Using (B.100), we then get:

$$\psi'\left(\delta_1^B\right)=\frac{\left[\begin{array}{c}\psi'\left(\delta_2^B\right)\left(1+\widetilde{\rho}-\left(1-\delta^D\right)(1-\nu)\left(1-\delta_2^B\right)\right)\\ +\left(\Pi\left(x_2^*\right)-\Pi\left(x_1^*\right)\right)(1+\rho)+(1-\nu)\left(1-\delta^D\right)\left(\psi\left(\delta_1^B\right)-\psi\left(\delta_2^B\right)\right)\end{array}\right]}{1+\widetilde{\rho}-\left(1-\delta^D\right)(1-\nu)\left(1-\delta_1^B\right)}. \tag{B.114}$$

Because of the convexity of ψ, we have $\psi\left(\delta_1^B\right)-\psi\left(\delta_2^B\right)\leq\left(\delta_1^B-\delta_2^B\right)\psi'\left(\delta_1^B\right)$, using the expression above we then get:

$$\psi\left(\delta_1^B\right)-\psi\left(\delta_2^B\right)\leq\left(\delta_1^B-\delta_2^B\right)\psi'\left(\delta_2^B\right)+\frac{(1+\rho)\left(\delta_1^B-\delta_2^B\right)\left(\Pi\left(x_2^*\right)-\Pi\left(x_1^*\right)\right)}{1+\widetilde{\rho}-\left(1-\delta^D\right)(1-\nu)\left(1-\delta_2^B\right)}. \tag{B.115}$$

Rearrange (B.113) to obtain:

$$\begin{array}{l}\frac{\psi'\left(\delta_2^B\right)}{\gamma^B-1}\left(\begin{array}{c}\left(1+\widetilde{\rho}-\left(1-\delta^D\right)\left[(1-\nu)\left(1-\delta_1^B\right)+\nu\delta^A\gamma^A\right]\right)\left(1+\widetilde{\rho}-\left(1-\delta^D\right)(1-\nu)\left(1-\delta_2^B+\delta_2^B\gamma^B\right)\right)\\ -(1-\nu)\,\delta_1^B\gamma^B\left(1-\delta^D\right)^2\nu\delta^A\gamma^A\end{array}\right)\\ =\left[\begin{array}{c}\left(1+\widetilde{\rho}-\left(1-\delta^D\right)\left[(1-\nu)\left(1-\delta_1^B\right)+\nu\delta^A\gamma^A\right]\right)(1+\rho)\,\Pi\left(x_2^*\right)\\ +\left(1-\delta^D\right)\nu\delta^A\gamma^A\left[(1+\rho)\,\Pi\left(x_1^*\right)-\left(1+\widetilde{\rho}-\left(1-\delta^D\right)(1-\nu)\left(1-\delta_1^B\right)\right)\left(\Pi\left(x_1^*\right)-\frac{1}{\gamma^A}\Pi\left(y^*\right)\right)\right]\\ -\left(1+\widetilde{\rho}-\left(1-\delta^D\right)(1-\nu)\left(1-\delta_1^B\right)\right)\left(1-\delta^D\right)(1-\nu)\,\psi\left(\delta_2^B\right)\\ -\left(1-\delta^D\right)\nu\delta^A\gamma^A\left(1-\delta^D\right)(1-\nu)\left(\psi\left(\delta_1^B\right)-\psi\left(\delta_2^B\right)\right)\end{array}\right]\end{array}$$

Then use (B.115) and $y^*\geq x_1^*$ to get:

$$\begin{array}{l}\frac{\psi'\left(\delta_2^B\right)}{\gamma^B-1}\left(1+\widetilde{\rho}-\left(1-\delta^D\right)(1-\nu)\left(1-\delta_1^B\right)\right)\left(\left(1+\widetilde{\rho}-\left(1-\delta^D\right)\left(\nu\delta^A\gamma+(1-\nu)\left(1-\delta_2^B+\delta_2^B\gamma^B\right)\right)\right)\right)\\ \geq\left[\begin{array}{c}\left(1+\widetilde{\rho}-\left(1-\delta^D\right)\left[(1-\nu)\left(1-\delta_1^B\right)+\nu\delta^A\gamma^A\right]\right)(1+\rho)\,\Pi\left(x_2^*\right)\\ +\left(1-\delta^D\right)\nu\delta^A\gamma^A\left[1+\rho-\left(1+\widetilde{\rho}-\left(1-\delta^D\right)(1-\nu)\left(1-\delta_1^B\right)\right)\left(1-\frac{1}{\gamma^A}\right)\right]\Pi\left(x_1^*\right)\\ -\left(1+\widetilde{\rho}-\left(1-\delta^D\right)(1-\nu)\left(1-\delta_1^B\right)\right)\left(1-\delta^D\right)(1-\nu)\,\psi\left(\delta_2^B\right)\\ -\nu\delta^A\gamma^A\left(1-\delta^D\right)^2(1-\nu)\,\frac{(1+\rho)\left(\delta_1^B-\delta_2^B\right)\left(\Pi\left(x_2^*\right)-\Pi\left(x_1^*\right)\right)}{1+\widetilde{\rho}-\left(1-\delta^D\right)(1-\nu)\left(1-\delta_2^B\right)}\end{array}\right].\end{array}$$

Further reorder terms (and use that $1+\widetilde{\rho}-\left(1-\delta^D\right)(1-\nu)\left(1-\delta_1^B\right)>0$) to obtain:

$$\begin{array}{l}\psi'\left(\delta_2^B\right)\left(\left(1+\widetilde{\rho}-\left(1-\delta^D\right)\left(\nu\delta^A\gamma+(1-\nu)\left(1-\delta_2^B+\delta_2^B\gamma^B\right)\right)\right)\right)\\ \geq\left(\gamma^B-1\right)\left[\begin{array}{c}\left(1+\rho-\left(1-\delta^D\right)\nu\delta^A\left(\gamma^A-1\right)\right)\Pi\left(x_1^*\right)-\left(1-\delta^D\right)(1-\nu)\,\psi\left(\delta_2^B\right)\\ +\frac{\left(1+\widetilde{\rho}-\left(1-\delta^D\right)\left(\nu\delta^A\gamma^A+(1-\nu)\left(1-\delta_2^B\right)\right)\right)(1+\rho)\left(\Pi\left(x_2^*\right)-\Pi\left(x_1^*\right)\right)}{1+\widetilde{\rho}-\left(1-\delta^D\right)(1-\nu)\left(1-\delta_2^B\right)}\end{array}\right].\end{array}$$

Using that $x_2^* \geq x_1^*$ and that $1+\widetilde{\rho}-\left(1-\delta^D\right)\left(\nu\delta^A\gamma^A+(1-\nu)\left(1-\delta_2^B\right)\right)>0$ for value functions to be defined, we get:

$$\psi'\left(\delta_2^B\right) \geq \frac{\left(\gamma^B-1\right)\left[\left(1+\rho-\left(1-\delta^D\right)\nu\delta^A\left(\gamma^A-1\right)\right)\Pi\left(x_1^*\right)-\left(1-\delta^D\right)(1-\nu)\psi\left(\delta_2^B\right)\right]}{1+\rho-\left(1-\delta^D\right)\left(\nu\left(1-\delta^A+\delta^A\gamma^A\right)+(1-\nu)\left(1-\delta_2^B+\delta_2^B\gamma^B\right)\right)}.$$

Further, note that

$$\begin{aligned}
&\frac{\partial}{\partial\delta^A}\frac{\left[\left(1+\rho-\left(1-\delta^D\right)\nu\delta^A\left(\gamma^A-1\right)\right)\Pi\left(x_1^*\right)-\left(1-\delta^D\right)(1-\nu)\psi\left(\delta_2^B\right)\right]}{1+\rho-\left(1-\delta^D\right)\left(\nu\left(1-\delta^A+\delta^A\gamma\right)+(1-\nu)\left(1-\delta_2^B+\delta_2^B\gamma^B\right)\right)}\\
&=\frac{\left(\Pi\left(x_1^*\right)\left(\nu+(1-\nu)\left(1-\delta_2^B+\delta_2^B\gamma^B\right)\right)-(1-\nu)\psi\left(\delta_2^B\right)\right)\left(1-\delta^D\right)^2\nu\left(\gamma^A-1\right)}{\left[1+\rho-\left(1-\delta^D\right)\left(\nu\left(1-\delta^A+\delta^A\gamma^A\right)+(1-\nu)\left(1-\delta_2^B+\delta_2^B\gamma^B\right)\right)\right]^2}>0
\end{aligned}$$

under the assumption that $\psi\left(\delta_2^B\right)<\left(\nu/(1-\nu)+1-\delta_2^B+\delta_2^B\gamma^B\right)$. Therefore, if $\delta^{A,coop} \geq \delta^{A,Nash}$, we get

$$\psi'\left(\delta_2^B\right) \geq \frac{\left(\gamma^B-1\right)\left[\left(1+\rho-\left(1-\delta^D\right)\nu\delta^{A,Nash}\left(\gamma^A-1\right)\right)\Pi\left(x_1^*\right)-\left(1-\delta^D\right)(1-\nu)\psi\left(\delta_2^B\right)\right]}{1+\rho-\left(1-\delta^D\right)\left(\nu\left(1-\delta^{A,Nash}+\delta^{A,Nash}\gamma^A\right)+(1-\nu)\left(1-\delta_2^B+\delta_2^B\gamma^B\right)\right)}.$$

Finally, $x_1^* > n$, therefore one gets:

$$\begin{aligned}
&\psi'\left(\delta_2^B\right)\left(1+\rho-\left(1-\delta^D\right)\left(\nu\left(1-\delta^{A,Nash}+\delta^{A,Nash}\gamma^A\right)+(1-\nu)\left(1-\delta_2^B+\delta_2^B\gamma^B\right)\right)\right)\\
&-\left(\gamma^B-1\right)\left[\left(1+\rho-\left(1-\delta^D\right)\nu\delta^{A,Nash}\left(\gamma^A-1\right)\right)\Pi(n)-\left(1-\delta^D\right)(1-\nu)\psi\left(\delta_2^B\right)\right]\\
&>0.
\end{aligned}$$

Noting that the left-hand side is an increasing function of δ_2^B and that $\delta^{B,Nash}$ is the solution to the left-hand side being equal to 0, we obtain

$$\delta_1^{B,coop} \geq \delta_2^{B,coop} > \delta^{B,Nash}.$$

Proof of Part a) Assume that $\delta_1^{B,coop} \geq \delta_2^{B,coop} > \delta^{B,Nash}$ (otherwise we know for sure that we must have $\delta^{A,coop} < \delta^{A,Nash}$ per part b)). Then one gets that $\lambda\left(\delta_1^{B,coop},\delta_2^{B,coop},\delta^A\right) > \lambda\left(\delta^{B,Nash},\delta^{B,Nash},\delta^A\right)$, since $V_I^{s,b} > V_I^{s,g}$, this factor pushes towards relatively less general innovation in the cooperative than in the Nash case. Other than that the expressions for the incentive to innovate is the same as in the baseline model, therefore sufficient conditions under which $\delta^{A,coop} < \delta^{A,Nash}$ in the baseline model are still sufficient now (but necessary conditions need not be so any more).

B.8.4 Proof of Remark 3

Assume the same exogenous innovation rates in the Nash and the cooperative cases with $\delta_2^B \leq \delta_1^B$. Then one can use (B.113) and obtain:

$$\frac{\psi'\left(\delta_{2,0}^{B,coop}\right)}{\gamma^B - 1} = \frac{\left[\begin{array}{c} \left(1+\widetilde{\rho} - \left(1-\delta^D\right)\left[\left(1-\nu\right)\left(1-\delta_1^B\right)+\nu\delta^A\gamma^A\right]\right)\left(\left(1+\rho\right)\Pi\left(x_2^*\right) - \left(1-\delta^D\right)\left(1-\nu\right)\psi\left(\delta_2^B\right)\right) \\ +\left(1-\delta^D\right)\nu\delta^A\gamma^A\left[\begin{array}{c} \left(\left(1+\rho\right)\Pi\left(x_1^*\right) - \left(1-\delta^D\right)\left(1-\nu\right)\psi\left(\delta_1^B\right)\right) \\ -\left(1+\widetilde{\rho}-\left(1-\delta^D\right)\left(1-\nu\right)\left(1-\delta_1^B\right)\right)\left(\Pi\left(x_1^*\right) - \frac{1}{\gamma^A}\Pi\left(y^*\right)\right) \end{array}\right] \end{array}\right]}{\left[\begin{array}{c} \left(1+\widetilde{\rho} - \left(1-\delta^D\right)\left[\left(1-\nu\right)\left(1-\delta_1^B\right)+\nu\delta^A\gamma^A\right]\right)\left(1+\widetilde{\rho}-\left(1-\delta^D\right)\left(1-\nu\right)\left(1-\delta_2^B+\delta_2^B\gamma^B\right)\right) \\ -\left(1-\nu\right)\delta_1^B\gamma^B\left(1-\delta^D\right)^2\nu\delta^A\gamma^A \end{array}\right]}$$

where $\widetilde{\rho}$ is defined as above with the exogenous innovation rate δ^A, and similarly,

$$\frac{\psi'\left(\delta_{2,0}^{B,Nash}\right)}{\gamma^B - 1} = \frac{\left[\begin{array}{c} \left(1+\widetilde{\rho} - \left(1-\delta^D\right)\left[\left(1-\nu\right)\left(1-\delta_1^B\right)+\nu\delta^A\gamma^A\right]\right)\left(\left(1+\rho\right)\Pi\left(n\right) - \left(1-\delta^D\right)\left(1-\nu\right)\psi\left(\delta_2^B\right)\right) \\ +\left(1-\delta^D\right)\nu\delta^A\gamma^A\left[\begin{array}{c} \left(\left(1+\rho\right)\Pi\left(n\right) - \left(1-\delta^D\right)\left(1-\nu\right)\psi\left(\delta_1^B\right)\right) \\ -\left(1+\rho-\left(1-\delta^D\right)\left[\left(1-\nu\right)\left(1-\delta_1^B\right)+\nu\left(1-\delta^A\right)\right]\right)\left(1-\frac{1}{\gamma^A}\right)\Pi\left(n\right) \end{array}\right] \end{array}\right]}{\left[\begin{array}{c} \left(1+\widetilde{\rho} - \left(1-\delta^D\right)\left[\left(1-\nu\right)\left(1-\delta_1^B\right)+\nu\delta^A\gamma^A\right]\right)\left(1+\widetilde{\rho}-\left(1-\delta^D\right)\left(1-\nu\right)\left(1-\delta_2^B+\delta_2^B\gamma^B\right)\right) \\ -\left(1-\nu\right)\delta_1^B\gamma^B\left(1-\delta^D\right)^2\nu\delta^A\gamma^A \end{array}\right]}$$

At the same time (B.104), (B.107), (B.108) and (B.109) give:

$$\psi'\left(\delta_0^{A,coop}\right) = \gamma^A\left(\begin{array}{c} \frac{(1-\omega)\lambda\left(\delta_1^B,\delta_2^B,\delta^A\right)}{\omega+(1-\omega)\lambda\left(\delta_1^B,\delta_2^B,\delta^A\right)}\left(\left(1-b\right)\Pi\left(x_1^*\right)+b\theta\Pi\left(n\right)-\frac{1}{\gamma^A}\Pi\left(y^*\right)\right)^+ \\ +\frac{\omega}{\omega+(1-\omega)\lambda\left(\delta_1^B,\delta_2^B,\delta^I\right)}\left[\left(1-b\right)\left(\Pi\left(x_1^*\right)-\frac{\Pi(y^*)}{\gamma^A}\right)+b\theta\left(1-\frac{1}{\gamma^A}\right)\Pi\left(n\right)\right] \end{array}\right),$$

$$\psi'\left(\delta_0^{A,Nash}\right) = \gamma^A\left(\begin{array}{c} \frac{(1-\omega)\lambda\left(\delta_1^B,\delta_2^B,\delta^A\right)}{\omega+(1-\omega)\lambda\left(\delta_1^B,\delta_2^B,\delta^A\right)}\left(1-b+b\theta-\frac{1}{\gamma^A}\right)^+ \\ +\frac{\omega}{\omega+(1-\omega)\lambda\left(\delta_1^B,\delta_2^B,\delta^I\right)}\left(1-\frac{1}{\gamma^A}\right)\left[1-b+b\theta\right] \end{array}\right)\Pi\left(n\right).$$

We then obtain

$$\left(\frac{\delta_0^{A,Nash}/\delta_{2,0}^{B,Nash}}{\delta_0^{A,coop}/\delta_{2,0}^{B,coop}}\right)^{\psi-1}=\frac{\psi'\left(\delta_0^{A,Nash}\right)/\psi'\left(\delta_{2,0}^{B,Nash}\right)}{\psi'\left(\delta_0^{A,coop}\right)/\psi'\left(\delta_{2,0}^{B,coop}\right)}=\frac{A}{B}$$

with

$$A\equiv\frac{\left(\frac{(1-\omega)\lambda\left(\delta_1^B,\delta_2^B,\delta^A\right)}{\omega+(1-\omega)\lambda\left(\delta_1^B,\delta_2^B,\delta^A\right)}\left(1-b+b\theta-\frac{1}{\gamma^A}\right)^{+}+\frac{\omega}{\omega+(1-\omega)\lambda\left(\delta_1^B,\delta_2^B,\delta^I\right)}\left(1-\frac{1}{\gamma^A}\right)[1-b+b\theta]\right)}{\left(\begin{array}{c}\frac{(1-\omega)\lambda\left(\delta_1^B,\delta_2^B,\delta^A\right)}{\omega+(1-\omega)\lambda\left(\delta_1^B,\delta_2^B,\delta^A\right)}\left((1-b)+b\theta\frac{\Pi(n)}{\Pi\left(x_1^*\right)}-\frac{1}{\gamma^A}\frac{\Pi(y^*)}{\Pi\left(x_1^*\right)}\right)^{+}\\+\frac{\omega}{\omega+(1-\omega)\lambda\left(\delta_1^B,\delta_2^B,\delta^I\right)}\left[(1-b)\left(1-\frac{\Pi(y^*)}{\gamma^A\Pi\left(x_1^*\right)}\right)+b\theta\left(1-\frac{1}{\gamma^A}\right)\frac{\Pi(n)}{\Pi\left(x_1^*\right)}\right]\end{array}\right)};$$

and

$$B\equiv\frac{\left[\begin{array}{c}\left(1+\widetilde{\rho}-\left(1-\delta^D\right)\left[(1-\nu)\left(1-\delta_1^B\right)+\nu\delta^A\gamma^A\right]\right)\left((1+\rho)-\left(1-\delta^D\right)(1-\nu)\frac{\psi\left(\delta_2^B\right)}{\Pi(n)}\right)\\+\left(1-\delta^D\right)\nu\delta^A\gamma^A\left[\begin{array}{c}\left((1+\rho)\Pi(n)-\left(1-\delta^D\right)(1-\nu)\frac{\psi\left(\delta_1^B\right)}{\Pi(n)}\right)\\-\left(1+\widetilde{\rho}-\left(1-\delta^D\right)(1-\nu)\left(1-\delta_1^B\right)\right)\left(1-\frac{1}{\gamma^A}\right)\end{array}\right]\end{array}\right]}{\left[\begin{array}{c}\left(1+\widetilde{\rho}-\left(1-\delta^D\right)\left[(1-\nu)\left(1-\delta_1^B\right)+\nu\delta^A\gamma^A\right]\right)\left((1+\rho)\frac{\Pi\left(x_2^*\right)}{\Pi\left(x_1^*\right)}-\left(1-\delta^D\right)(1-\nu)\frac{\psi\left(\delta_2^B\right)}{\Pi\left(x_1^*\right)}\right)\\+\left(1-\delta^D\right)\nu\delta^A\gamma^A\left[\begin{array}{c}\left(1+\rho-\left(1-\delta^D\right)(1-\nu)\frac{\psi\left(\delta_1^B\right)}{\Pi\left(x_1^*\right)}\right)\\-\left(1+\widetilde{\rho}-\left(1-\delta^D\right)(1-\nu)\left(1-\delta_1^B\right)\right)\left(1-\frac{1}{\gamma^A}\frac{\Pi(y^*)}{\Pi\left(x_1^*\right)}\right)\end{array}\right]\end{array}\right]}.$$

Since $\frac{\Pi(n)}{\Pi\left(x_1^*\right)}<1$ and $\frac{\Pi(y^*)}{\Pi\left(x_1^*\right)}\geq 1$ then $A>1$. In addition, $\frac{\Pi\left(x_2^*\right)}{\Pi\left(x_1^*\right)}\geq 1$ and $\frac{\Pi(y^*)}{\Pi\left(x_1^*\right)}\geq 1$ plus $\frac{\psi\left(\delta_2^B\right)}{\Pi\left(x_1^*\right)}<\frac{\psi\left(\delta_2^B\right)}{\Pi(n)}$ and $\frac{\psi\left(\delta_1^B\right)}{\Pi\left(x_1^*\right)}<\frac{\psi\left(\delta_2^B\right)}{\Pi(n)}$ imply that $B<1$, so that $A/B>1$ and

$$\frac{\delta_0^{A,Nash}}{\delta_{2,0}^{B,Nash}}>\frac{\delta_0^{A,coop}}{\delta_{2,0}^{B,coop}}.$$

If instead exogenous innovation is free, then the ψ terms disappear from B and we would get $B\leq 1$ so that we still have $A/B>1$.

Using (B.99) together with (B.111) and (B.110)—these expressions are valid as they do not use (B.99) for δ_1^B—we obtain:

$$\psi'\left(\delta_{1,0}^{B,coop}\right) = \frac{\left[\begin{array}{c} \gamma^B\left(1+\widetilde{\rho}-\left(1-\delta^D\right)\left[1-\nu+\nu\delta^A\gamma^A\right]\right)(1+\rho)\,\Pi\left(x_2^*\right) \\ -\left(1+\widetilde{\rho}-\left(1-\delta^D\right)\left[1-\nu+\nu\delta^A\gamma^A\right]\right)\left(\gamma^B\psi\left(\delta_2^B\right)-\psi\left(\delta_1^B\right)\right)\left(1-\delta^D\right)(1-\nu) \\ -\left(1+\widetilde{\rho}-\left(1-\delta^D\right)\left(\nu\gamma^B\delta^A\gamma^A+(1-\nu)\left(1-\delta_2^B+\delta_2^B\gamma^B\right)\right)\right)(1+\rho)\,\Pi\left(x_1^*\right) \\ -\left(1-\delta^D\right)\left(\gamma^B-1\right)\left((1-\nu)\,\delta_2^B+\nu\delta^A\gamma^A\right)\psi\left(\delta_1^B\right)\left(1-\delta^D\right)(1-\nu) \\ -\left(\gamma^B-1\right)\left(\Pi\left(x_1^*\right)-\frac{1}{\gamma^A}\Pi\left(y^*\right)\right)\left(1-\delta^D\right)\nu\delta^A\gamma^A\left(1+\widetilde{\rho}-\left(1-\delta^D\right)(1-\nu)\left(1-\delta_2^B\right)\right) \end{array}\right]}{\left[\begin{array}{c} \left(1+\widetilde{\rho}-\left(1-\delta^D\right)\left[(1-\nu)\left(1-\delta_1^B\right)+\nu\delta^A\gamma^A\right]\right)\left(1+\widetilde{\rho}-\left(1-\delta^D\right)(1-\nu)\left(1-\delta_2^B+\delta_2^B\gamma^B\right)\right) \\ -(1-\nu)\,\delta_1^B\gamma^B\left(1-\delta^D\right)^2\nu\delta^A\gamma^A \end{array}\right]},$$

and similarly

$$\psi'\left(\delta_{1,0}^{B,Nash}\right) = \frac{\begin{array}{c} \gamma^B\left(1+\widetilde{\rho}-\left(1-\delta^D\right)\left[1-\nu+\nu\delta^A\gamma^A\right]\right)(1+\rho)\,\Pi(n) \\ -\left(1+\widetilde{\rho}-\left(1-\delta^D\right)\left[1-\nu+\nu\delta^A\gamma^A\right]\right)\left(\gamma^B\psi\left(\delta_2^B\right)-\psi\left(\delta_1^B\right)\right)\left(1-\delta^D\right)(1-\nu) \\ -\left(1+\widetilde{\rho}-\left(1-\delta^D\right)\left(\nu\gamma^B\delta^A\gamma^A+(1-\nu)\left(1-\delta_2^B+\delta_2^B\gamma^B\right)\right)\right)(1+\rho)\,\Pi(n) \\ -\left(1-\delta^D\right)\left(\gamma^B-1\right)\left((1-\nu)\,\delta_2^B+\nu\delta^A\gamma^A\right)\psi\left(\delta_1^B\right)\left(1-\delta^D\right)(1-\nu) \\ -\left(\gamma^B-1\right)\left(1-\frac{1}{\gamma^A}\right)\Pi(n)\left(1-\delta^D\right)\nu\delta^A\gamma^A\left(1+\widetilde{\rho}-\left(1-\delta^D\right)(1-\nu)\left(1-\delta_2^B\right)\right) \end{array}}{\left[\begin{array}{c} \left(1+\widetilde{\rho}-\left(1-\delta^D\right)\left[(1-\nu)\left(1-\delta_1^B\right)+\nu\delta^A\gamma^A\right]\right)\left(1+\widetilde{\rho}-\left(1-\delta^D\right)(1-\nu)\left(1-\delta_2^B+\delta_2^B\gamma^B\right)\right) \\ -(1-\nu)\,\delta_1^B\gamma^B\left(1-\delta^D\right)^2\nu\delta^A\gamma^A \end{array}\right]}.$$

Therefore, we get:

$$\left(\frac{\delta_0^{A,Nash}/\delta_{1,0}^{B,Nash}}{\delta_0^{A,coop}/\delta_{1,0}^{B,coop}}\right)^{\psi-1} = \frac{\psi'\left(\delta_0^{A,Nash}\right)/\psi'\left(\delta_{1,0}^{B,Nash}\right)}{\psi'\left(\delta_0^{A,coop}\right)/\psi'\left(\delta_{1,0}^{B,coop}\right)} = \frac{A}{C},$$

with

$$C \equiv \frac{\begin{array}{c}\gamma^B\left(1+\widetilde{\rho}-\left(1-\delta^D\right)\left[1-\nu+\nu\delta^A\gamma^A\right]\right)(1+\rho)\\ -\left(1+\widetilde{\rho}-\left(1-\delta^D\right)\left(\nu\gamma^B\delta^A\gamma^A+(1-\nu)\left(1-\delta_2^B+\delta_2^B\gamma^B\right)\right)\right)(1+\rho)\\ -\left(\gamma^B-1\right)\left(1-\frac{1}{\gamma^A}\right)\left(1-\delta^D\right)\nu\delta^A\gamma^A\left(1+\widetilde{\rho}-\left(1-\delta^D\right)(1-\nu)\left(1-\delta_2^B\right)\right)\\ -\left(1+\widetilde{\rho}-\left(1-\delta^D\right)\left[(1-\nu)+\nu\delta^A\gamma^A\right]\right)\frac{\left(\gamma^B\psi\left(\delta_2^B\right)-\psi\left(\delta_1^B\right)\right)}{\Pi(n)}\left(1-\delta^D\right)(1-\nu)\\ -\left(1-\delta^D\right)\left(\gamma^B-1\right)\left((1-\nu)\,\delta_2^B+\nu\delta^A\gamma^A\right)\frac{\psi\left(\delta_1^B\right)}{\Pi(n)}\left(1-\delta^D\right)(1-\nu)\end{array}}{\begin{array}{c}\gamma^B\left(1+\widetilde{\rho}-\left(1-\delta^D\right)\left[(1-\nu)+\nu\delta^A\gamma^A\right]\right)(1+\rho)\frac{\Pi\left(x_2^*\right)}{\Pi\left(x_1^*\right)}\\ -\left(1+\widetilde{\rho}-\left(1-\delta^D\right)\left(\nu\gamma^B\delta^A\gamma^A+(1-\nu)\left(1-\delta_2^B+\delta_2^B\gamma^B\right)\right)\right)(1+\rho)\\ -\left(\gamma^B-1\right)\left(1-\frac{1}{\gamma^A}\frac{\Pi(y^*)}{\Pi\left(x_1^*\right)}\right)\left(1-\delta^D\right)\nu\delta^A\gamma^A\left(1+\widetilde{\rho}-\left(1-\delta^D\right)(1-\nu)\left(1-\delta_2^B\right)\right)\\ -\left(1+\widetilde{\rho}-\left(1-\delta^D\right)\left[1-\nu+\nu\delta^A\gamma^A\right]\right)\frac{\gamma^B\psi\left(\delta_2^B\right)-\psi\left(\delta_1^B\right)}{\Pi\left(x_1^*\right)}\left(1-\delta^D\right)(1-\nu)\\ -\left(1-\delta^D\right)\left(\gamma^B-1\right)\left((1-\nu)\,\delta_2^B+\nu\delta^A\gamma^A\right)\frac{\psi\left(\delta_1^B\right)}{\Pi\left(x_1^*\right)}\left(1-\delta^D\right)(1-\nu)\end{array}}.$$

$\frac{\Pi\left(x_2^*\right)}{\Pi\left(x_1^*\right)} \geq 1$ and $\frac{\Pi(y^*)}{\Pi\left(x_1^*\right)} \geq 1$ both push towards $C \leq 1$, we then get that $C < 1$ provided that

$$\begin{aligned}&\left(1+\rho-\left(1-\delta^D\right)\left[(1-\nu)+\nu\left(1-\delta^A+\delta^A\gamma^A\right)\right]\right)\left(\gamma^B\psi\left(\delta_2^B\right)-\psi\left(\delta_1^B\right)\right)\\ &+\left(1-\delta^D\right)\left(\gamma^B-1\right)\left((1-\nu)\,\delta_2^B+\nu\delta^A\gamma^A\right)\psi\left(\delta_1^B\right)\\ &>0,\end{aligned}$$

a sufficient condition is then that $\gamma^B\psi\left(\delta_2^B\right) \geq \psi\left(\delta_1^B\right)$. We then have

$$\frac{\delta_0^{A,Nash}}{\delta_{1,0}^{B,Nash}} > \frac{\delta_0^{A,coop}}{\delta_{1,0}^{B,coop}}.$$

If innovation is free when exogenous then the ψ terms in C disappear and we directly have $C \leq 1$.

B.9 Slow diffusion of innovation

We consider a cooperative equilibrium where at the beginning of any relationship a good match cooperates as much as possible whether at the frontier or not, while there is no cooperation in bad matches. As explained in the text, the equilibrium is characterized by the levels of cooperation in frontier good matches (x^*) and in outdated good matches (y^*). It is direct to derive the IC constraints (10) and (29).

On equilibrium path, a producer switches between suppliers (favoring those with the frontier technology) until he finds a good match. Once he has found one, he optimally decides between switching to an innovator (when innovation occurs) or staying with the outdated good match supplier. If it is optimal to switch to the innovator and the innovator turns out to be a good match, he stays in a relationship with that innovator. If she turns out to be a bad match, then in the following period, the producer resumes his relationship with his old supplier if that supplier obtained the frontier technology and tries another frontier firm otherwise.

In this appendix, we first derive the condition under which a producer who knows a good match switches to the innovator in the cooperative case. Then, we look at the corresponding condition in the contractible or Nash cases. Afterward, we describe in more details the IC constraints and check for the existence of the equilibrium. Finally we derive the equations determining the innovation rates in the three cases.

B.9.1 Switching in the cooperative equilibrium

We focus on the cooperative equilibrium. As in the baseline model, Bertrand competition ensures that

$$V_1^s = V_1^T - V_0^{T,b}. \tag{B.116}$$

V_1^T is the joint value with a (cooperating) good match supplier. $V_0^{T,b}$ is the joint value of starting a relationship with a frontier firm without the option to fall back on a cooperating outdated good match, which we indicated through the superscript b. Making this distinction is now helpful since a producer working with an outdated good match supplier may try to start a new relationship with a frontier supplier (even in periods without innovation), without necessarily being punished for doing so (as the outdated supplier would forgive if the frontier supplier turns out to be a bad match). In this case, the outside option of a producer working with a frontier supplier is to try another frontier supplier, but his previous good match would not forgive him for doing so.

Consider now a cooperating outdated good match. If the producer were to try a frontier firm, her expected value does not become 0 because she could get the producer back in the following period if the frontier firm turns out to be a bad match. As before, we denote by V_A^s, the expected value of an outdated good match when the producer tries out a frontier firm. Bertrand competition ensures that the value of an outdated good match W_1^s must satisfy $W_1^s \geq V_A^s$. Following the same reasoning as in Appendix A.1.1, we get that (A.2) holds. That is $W_1^p = V_I^{p,g}$, where $V_I^{p,g}$ is the value that a producer obtains in a relationship with the innovator when the producer already knows a cooperating outdated good match, and W_1^p is the value the producer captures with an outdated good match. This ensures that a producer (who knows a good match) switches to the innovator if (A.3) holds that is $V_I^{T,g} > W_1^T - V_A^s$, where $V_I^{T,g}$ is the joint value of a relationship with the innovator (when the producer knows a cooperating outdated good match). As a result, we get that (as in the baseline model):

$$W_1^s = V_A^s + \left(W_1^T - V_A^s - V_I^{T,g}\right)^+. \tag{B.117}$$

Note that $V_I^{T,g}$ is the same as $V_0^{T,g}$, the joint value of starting a relationship with any frontier firm in a period without innovation when the producer also knows a cooperating outdated good match. Therefore in periods without innovation, a producer who knows a good match switches to a frontier firm under the same circumstances (that is whenever $V_I^{T,g} = V_0^{T,g} > W_1^T - V_A^s$). However, Bertrand competition ensures that $V_0^{p,g} = V_0^{T,g}$ since in periods without innovation there is more than one firm with the frontier technology, so that, generally $V_0^{p,g} \neq V_I^{T,g}$.

We obtain the law of motion

$$V_I^{T,g} = V_0^{T,g} = (1-b)\, V_1^T + b\left(\theta\Pi(n) + \frac{1-\delta^D}{1+\rho}\left(\left(1-\delta^I\right)\left(\Delta V_1^p + (1-\Delta)\max\left(V_0^{T,g}, W_1^p\right)\right) + \delta^I \gamma W_1^p\right)\right) \tag{B.118}$$

With probability $1-b$ the match is good, and the value is given by V_1^T. With probability b the match is bad, generating profits $\theta\Pi(n)$ this period; in the following period, there are three possibilities. i) No innovation occurred but the previous good match got access to the frontier technology, then the producer resumes his previous relationship and obtains V_1^p. ii) No innovation occurred and the previous good match did not inherit the technology, then the producer optimally decides between staying with his previous good match or trying another frontier supplier; since there are now several suppliers the producer would get the full value of a relationship with a frontier firm if that is higher than $W_1^T - V_A^s$; otherwise he stays with the outdated supplier and gets $W_1^p = W_1^T - \left(V_A^s + \left(W_1^T - V_A^s - V_I^{T,g}\right)\right) = V_0^{T,g}$ in that case. iii) An innovation occurs, in which case the producer can secure γW_1^p (as the frontier has moved one step). We can then rewrite more simply:

$$V_I^{T,g} = V_0^{T,g} = (1-b)\, V_1^T + b\left(\theta\Pi\,(n) + \frac{1-\delta^D}{1+\rho}\left(\left(1-\delta^I\right)\left(\Delta V_1^p + (1-\Delta)\, V_0^{T,g}\right) + \delta^I\gamma W_1^p\right)\right). \tag{B.119}$$

The expected value of an outdated supplier when the producer tries a good match, V_A^s, obeys the following law of motion:

$$V_A^s = b\frac{1-\delta^D}{1+\rho}\left(\left(1-\delta^I\right)\Delta V_1^s + \left(\left(1-\delta^I\right)(1-\Delta) + \delta^I\gamma\right) W_1^s\right). \tag{B.120}$$

With probability $1-b$, the producer met a good match and therefore the outdated supplier turns into a non-cooperating good match and her value becomes 0. Otherwise, she resumes her relationship with the producer in the following period and obtains V_1^s if she imitates the frontier technology and no innovation occurred. If she does not get access to the frontier technology (either because of a new innovation or because the previous one did not diffuse), the supplier's normalized value is W_1^s.

Combining the two, we get that

$$\begin{aligned} & V_0^{T,g} + V_A^s \\ = \; & (1-b)\, V_1^T + b\left(\theta\Pi\,(n) + \frac{1-\delta^D}{1+\rho}\left(\left(1-\delta^I\right)\left(\Delta V_1^T + (1-\Delta)\left(V_0^{T,g} + V_A^s + \left(W_1^T - V_A^s - V_0^{T,g}\right)^+\right)\right.\right.\right. \end{aligned}$$

The joint value of a relationship with the frontier supplier is still given by (4), which we reproduce here:

$$V_1^T = \Pi\,(x^*) + \frac{1-\delta^D}{1+\rho}\left(\left(1-\delta^I\right) V_1^T + \delta^I\gamma W_1^T\right). \tag{B.122}$$

The joint value of a relationship with an outdated producer obeys the following law of motion:

$$W_1^T = \gamma^{-1}\Pi\,(y^*) + \frac{1-\delta^D}{1+\rho}\left(\left(1-\delta^I\right)\left(\Delta V_1^T + (1-\Delta)\left(W_1^T + \left(V_0^{T,g} + V_A^s - W_1^T\right)^+\right)\right) + \delta^I\gamma W_1^T\right). \tag{B.123}$$

The current flow of profits is given by $\gamma^{-1}\Pi\,(y^*)$ since the producer does not have access to the frontier technology. In the following period, the relationship becomes a frontier one if the supplier gets access to the frontier technology (which occurs with probability $\left(1-\delta^I\right)\Delta$). If the technology does not diffuse then the producer should try a frontier supplier if $V_0^{T,g} > W_1^T - V_A^s$, in that case he obtains $V_0^{T,g}$ (since there are several frontier firms) and the supplier gets the expected value $W_1^s = V_A^s$; on the other hand if $V_0^{T,g} < W_1^T - V_A^s$, the producer stays with the outdated good match and they obtain W_1^T together. If another innovation occurs, then the innovator would be the only one with the frontier technology and would obtain any surplus of a relationship with her, hence the joint value of the producer and the (previous) supplier is W_1^T.

Combine (B.121), (B.122) and (B.123) to obtain:

$$\begin{aligned} & V_0^{T,g} + V_A^s - W_1^T \\ = \; & (1-b)\,\Pi(x^*) + b\theta\Pi(n) - \gamma^{-1}\Pi(y^*) + \frac{1-\delta^D}{1+\rho}\left(1-\delta^I\right)(1-\Delta)\times \\ & \left[(1-b)\,V_1^T + b\left(V_0^{T,g} + V_A^s + \left(W_1^T - V_A^s - V_0^{T,g}\right)^+\right) - \left(W_1^T + \left(V_0^{T,g} + V_A^s - W_1^T\right)^+\right)\right]. \end{aligned} \tag{B.124}$$

Next, using (B.122) and (B.123), we get:

$$V_1^T - W_1^T = \Pi(x^*) - \gamma^{-1}\Pi(y^*) + \frac{1-\delta^D}{1+\rho}\left(1-\delta^I\right)(1-\Delta)\left(V_1^T - W_1^T - \left(V_0^{T,g} + V_A^s - W_1^T\right)^+\right).$$

Hence

$$V_1^T - W_1^T = \frac{(1+\rho)\left(\Pi(x^*) - \gamma^{-1}\Pi(y^*)\right)}{1+\rho-(1-\delta^D)(1-\delta^I)(1-\Delta)} - \frac{\left(1-\delta^D\right)\left(1-\delta^I\right)(1-\Delta)}{1+\rho-(1-\delta^D)(1-\delta^I)(1-\Delta)}\left(V_0^{T,g} + V_A^s - W_1^T\right)^+. \tag{B.125}$$

Plugging this expression in (B.124), we get

$$\begin{aligned} & V_0^{T,g} + V_A^s - W_1^T \\ = \; & (1-b)\,\Pi(x^*) + b\theta\Pi(n) - \gamma^{-1}\Pi(y^*) \\ & + \frac{\left(1-\delta^D\right)\left(1-\delta^I\right)(1-\Delta)(1-b)}{1+\rho-(1-\delta^D)(1-\delta^I)(1-\Delta)}\left[\left(\Pi(x^*) - \gamma^{-1}\Pi(y^*)\right) - \left(V_0^{T,g} + V_A^s - W_1^T\right)^+\right] \end{aligned}$$

This implies that $V_0^{T,g} + V_A^s - W_1^T > 0$ if and only if (30) holds. Further

$$\left(V_0^{T,g} + V_A^s - W_1^T\right)^+ = \tag{B.126}$$

$$\frac{\left(b\left(1+\rho-\left(1-\delta^D\right)\left(1-\delta^I\right)(1-\Delta)\right)\left(\theta\Pi(n) - \gamma^{-1}\Pi(y^*)\right) + (1+\rho)(1-b)\left(\Pi(x^*) - \gamma^{-1}\Pi(y^*)\right)\right)^+}{1+\rho-b\,(1-\delta^D)(1-\delta^I)(1-\Delta)}.$$

B.9.2 Switching in the contractible and Nash cases

In the contractible and Nash cases the same logic applies, since once a producer has found two good matches he can only remember the last one. Therefore, as in the equilibrium path described above, a producer would only return to the his last good match after having tried out the frontier firm (which implies that whether the technology diffuses to suppliers with whom he worked before does not matter). We can then directly copy the previous equations but replacing all investment levels by m in the contractible case and n in the Nash case. Therefore a producer switches if and only if (31) holds.

B.9.3 IC constraints in the cooperative case

To compare x^* and y^*, we need to compare the right-hand side of (10) and (29). To do that we first combine (B.117) and (B.120) to get

$$V_A^s = \frac{b\left(1-\delta^D\right)\left(\left(1-\delta^I\right)\Delta V_1^s + \left(\left(1-\delta^I\right)(1-\Delta)+\delta^I\gamma\right)\left(W_1^T - V_A^s - V_I^{T,g}\right)^+\right)}{1+\rho-b\left(1-\delta^D\right)\left(\left(1-\delta^I\right)(1-\Delta)+\delta^I\gamma\right)},$$

and

$$W_1^s = \frac{b\left(1-\delta^D\right)\left(1-\delta^I\right)\Delta V_1^s + (1+\rho)\left(W_1^T - V_A^s - V_I^{T,g}\right)^+}{1+\rho-b\left(1-\delta^D\right)\left(\left(1-\delta^I\right)(1-\Delta)+\delta^I\gamma\right)}. \tag{B.127}$$

Define $IC_y \equiv \gamma\left(1-\delta^I\right)\Delta V_1^s + \gamma\left(\left(1-\delta^I\right)(1-\Delta)+\delta^I\gamma\right)W_1^s$ and $IC_x \equiv \left(1-\delta^I\right)V_1^s + \delta^I\gamma W_1^s$, then we can rewrite (10) and (29) as

$$\varphi(x^*) \leq \frac{1-\delta^D}{1+\rho}IC_x \text{ and } \varphi(y^*) \leq \frac{1-\delta^D}{1+\rho}IC_y.$$

Then, using (B.127), we get:

$$\begin{aligned} & IC_y - IC_x \\ = \; & \frac{1}{1+\rho-b\left(1-\delta^D\right)\left(\left(1-\delta^I\right)(1-\Delta)+\delta^I\gamma\right)} \\ & \left(\begin{array}{c} \left((1+\rho)(\gamma\Delta-1)+b\left(1-\delta^D\right)(1-\Delta)\left(\left(1-\delta^I\right)+\delta^I\gamma\right)\right)\left(1-\delta^I\right)V_1^s \\ +\left(\left(1-\delta^I\right)(1-\Delta)+\delta^I(\gamma-1)\right)\gamma(1+\rho)\left(W_1^T-V_A^s-V_I^{T,g}\right)^+ \end{array}\right). \end{aligned}$$

In equilibrium $V_1^s > 0$, therefore a sufficient condition to ensure that $IC_y - IC_x \geq 0$ is that $(1+\rho)(\gamma\Delta-1)+b\left(1-\delta^D\right)(1-\Delta)\left(\left(1-\delta^I\right)+\delta^I\gamma\right) \geq 0$, which is satisfied for any δ^I as long as $\Delta > \left(1+\rho - b\left(1-\delta^D\right)\right) / \left(\gamma\left(1+\rho\right) - b\left(1-\delta^D\right)\right)$.

To ensure that an equilibrium exists, we must check that the IC constraints are not binding at n. This requires finding an expression for V_1^s. To do that, first note that the value of a relationship with a frontier firm for a supplier who does not know a cooperating outdated good match, $V_0^{T,b}$ follows:

$$V_0^{T,b} = (1-b)\, V_1^T + b\theta\Pi\left(n\right) + \frac{1-\delta^D}{1+\rho} b\left(\left(1-\delta^I\right) V_0^{T,b} + \delta^I \gamma W_0^T\right). \tag{B.128}$$

With probability $1-b$, the producer meets a good match. Otherwise, the situation is the same in the next period if there has been no innovation, while if an innovation occurs, the producer would try the innovator but would only capture his outside option, namely starting a relationship with an outdated supplier (as there is only one frontier firm then, $V_I^{p,b} = W_0^T$).

Similarly, the law of motion of W_0^T is:

$$W_0^T = (1-b)\, W_1^T + b\theta\gamma^{-1}\Pi\left(n\right) + b\frac{1-\delta^D}{1+\rho}\left(\left(1-\delta^I\right) V_0^{T,b} + \delta^I \gamma W_0^T\right). \tag{B.129}$$

With probability $1-b$, the producer meets an outdated good match, generating the joint value W_1^T. Otherwise, the producer gets current profit $\theta\gamma^{-1}\Pi\left(n\right)$ (with an outdated bad match), and in the following period, he tries one of the frontier suppliers and capture the full value if no innovation occurs, while he can only capture the value of a relationship with a new outdated supplier if an innovation occurs.

Combining (B.128) and (B.129), we get

$$V_0^{T,b} - W_0^T = (1-b)\left(V_1^T - W_1^T\right) + b\theta\left(1-\gamma^{-1}\right)\Pi\left(n\right). \tag{B.130}$$

Further, combining (B.122), (B.129) and using (B.130), we get

$$\begin{aligned} V_1^s &= V_1^T - V_0^{T,b} \\ &= b\left(\Pi\left(x^*\right) - \theta\Pi\left(n\right) + \tfrac{1-\delta^D}{1+\rho}\left(\left(1-\delta^I+\delta^I\gamma\right) V_1^s - b\delta^I\gamma\left(V_1^T - W_1^T\right) + \delta^I b\theta\left(\gamma-1\right)\Pi\left(n\right)\right)\right). \end{aligned}$$

Therefore, using (B.125), we get:

$$V_1^s = \frac{b\left(1+\rho\right)}{\left(1+\rho-\left(1-\delta^D\right)\left(1-\delta^I\right)\left(1-\Delta\right)\right)\left(1+\rho-b\left(1-\delta^D\right)\left(1-\delta^I+\delta^I\gamma\right)\right)} L\left(x^*, y^*\right),$$

with

$$
\begin{aligned}
L\left(x^{*},y^{*}\right) \equiv\ & \left(1+\rho-\left(1-\delta^{D}\right)\left(1-\delta^{I}\right)(1-\Delta)-b\delta^{I}\gamma\left(1-\delta^{D}\right)\right)\left(\Pi\left(x^{*}\right)-\theta\Pi\left(n\right)\right) \\
& +b\delta^{I}\left(1-\delta^{D}\right)\left(\Pi\left(y^{*}\right)-\theta\Pi\left(n\right)\right) \\
& -b\delta^{I}\gamma\frac{\left(1-\delta^{D}\right)^{2}}{1+\rho}\left(1-\delta^{I}\right)(1-\Delta)\left(\theta\left(1-\gamma^{-1}\right)\Pi\left(n\right)+\left(V_{0}^{T,g}+V_{A}^{s}-W_{1}^{T}\right)^{+}\right)
\end{aligned}
$$

Using (B.126), we get:

$$
\begin{aligned}
& L\left(x^{*},y^{*}\right) \\
\geq\ & \Pi\left(x^{*}\right)-\theta\Pi\left(n\right)-\frac{b\delta^{I}\gamma\left(1-\delta^{D}\right)\left(\Pi(x^{*})-\gamma^{-1}\Pi(y^{*})\right)}{1+\rho-(1-\delta^{D})(1-\delta^{I})(1-\Delta)}+\frac{b\delta^{I}\gamma\left(1-\delta^{D}\right)}{1+\rho}\theta\left(1-\gamma^{-1}\right)\Pi\left(n\right)+\frac{b\delta^{I}\gamma}{1+\rho}\times \\
& \frac{\left(1-\delta^{D}\right)^{2}\left(1-\delta^{I}\right)(1-\Delta)}{1+\rho-b(1-\delta^{D})(1-\delta^{I})(1-\Delta)}\left[b\left(\theta\Pi\left(n\right)-\gamma^{-1}\Pi\left(y^{*}\right)\right)+\frac{(1+\rho)(1-b)\left(\Pi(x^{*})-\gamma^{-1}\Pi(y^{*})\right)}{1+\rho-(1-\delta^{D})(1-\delta^{I})(1-\Delta)}\right]
\end{aligned}
$$

Hence

$$
\begin{aligned}
\frac{L\left(n,n\right)}{\Pi\left(n\right)} \geq\ & (1-\theta)\left(1-\frac{b\delta^{I}\gamma\left(1-\delta^{D}\right)}{1+\rho}\left(\left(1-\gamma^{-1}\right)+\frac{b\left(1-\delta^{D}\right)\left(1-\delta^{I}\right)(1-\Delta)}{1+\rho-b\left(1-\delta^{D}\right)\left(1-\delta^{I}\right)(1-\Delta)}\right)\right) \\
\geq\ & (1-\theta)\left(\frac{1+\rho-b\left(1-\delta^{D}\right)\left(1-\delta^{I}+\delta^{I}\gamma\right)(1-\Delta)}{1+\rho-b\left(1-\delta^{D}\right)\left(1-\delta^{I}\right)(1-\Delta)}+\frac{b\delta^{I}\left(1-\delta^{D}\right)}{1+\rho}\right)>0
\end{aligned}
$$

The last inequality is obtained because we must have $\rho>\delta^{I}\left(\gamma-1\right)$ to ensure that utility is finite. Therefore the IC constraints do not bind at $x^{*}=y^{*}=n$, which ensures that there exist values such that $x^{*},y^{*}\in(n,m]$ and either the IC constraints bind or the first best is achieved without violating the constraints.

Our equilibrium assumes that in periods with an innovation, the best offer that a producer receives comes from either the innovator or an outdated good match if he knows one. Further, the outdated good match is not willing to offer the full value of the relationship to the producer because even if the producer chooses the innovator, she still secures a postie expected value from the relationship. Therefore, in principle, we should check that the offer from a new outdated supplier cannot be better than that of the current outdated good match supplier. First assume that the producer would rather stay with her current supplier than switch to the innovator ($W_{1}^{T}-V_{A}^{s}>V_{0}^{T,g}$), then since the innovator dominates a new outdated supplier, it is clear that the good match supplier's offer is better than that of a new outdated supplier.

Let us then assume that $W_{1}^{T}-V_{A}^{s}<V_{0}^{T,g}$, we must then check that we have $W_{0}^{T}<$

$W_1^T - V_A^s$. Combining (B.116), (B.120) with $W_1^s = V_A^s$, (B.123), and (B.129), we get:

$$\begin{aligned}
& \left(W_1^T - V_A^s - W_0^T\right)\left(1 - b\frac{1-\delta^D}{1+\rho}\delta^I\gamma\right) \\
= \quad & b\gamma^{-1}\left(\Pi\left(y^*\right) - \theta\Pi\left(n\right)\right) + b\frac{1-\delta^D}{1+\rho}\left(\left(1-\delta^I\right)\left(1-\Delta\right)\left(V_0^{T,g} - V_0^{T,b}\right)\right).
\end{aligned}$$

Similarly using (B.118) and (B.128), we get:

$$\left(V_0^{T,g} - V_0^{T,b}\right)\left(1 - b\frac{1-\delta^D}{1+\rho}\left(1-\delta^I\right)\left(1-\Delta\right)\right) = b\frac{1-\delta^D}{1+\rho}\delta^I\gamma\left(W_1^T - V_A^s - W_0^T\right).$$

Therefore both $V_0^{T,g} > V_0^{T,b}$ and $W_1^T - V_A^s - W_0^T > 0$, which ensures that the two best options for a producer in a period with innovation are the innovator and a good match outdated supplier if he knows one.

B.9.4 Endogenous innovation

Finally, we want to determine the innovation rates, which requires to find the reward from innovation Z_K (for $K = cont, coop$ or $Nash$). Z_K is still defined as $Z_K = \omega V_{I,K}^{s,b} + (1-\omega)V_{I,K}^{s,g}$, and we still have that the expected mass of producers who do not know a good match supplier in steady-state is equal to $\omega = \delta^D / \left(1 - b\left(1-\delta^D\right)\right)$.

In the cooperative case, we obtain $V_I^{s,g} = \left(V_I^{T,g} - \left(W_1^T - V_A^s\right)\right)^+$ which is given in (B.126). For producers who do not know an outdated good match, an innovator captures $V_I^{s,b} = V_0^{T,b} - W_0^{T,b}$, which using (B.130) and (B.125) is given by

$$\begin{aligned}
V_I^{s,b} \quad = \quad & (1-b)\frac{(1+\rho)\left(\Pi\left(x^*\right) - \gamma^{-1}\Pi\left(y^*\right)\right) - \left(1-\delta^D\right)\left(1-\delta^I\right)\left(1-\Delta\right)\left(V_0^{T,g} + V_A^s - W_1^T\right)^+}{1+\rho-\left(1-\delta^D\right)\left(1-\delta^I\right)\left(1-\Delta\right)} \\
& + b\theta\left(1-\gamma^{-1}\right)\Pi\left(n\right).
\end{aligned}$$

We then get that

$$\begin{aligned}
& Z_{coop} \\
= \quad & \frac{\Pi\left(x^*\right)}{1-b\left(1-\delta^D\right)}\left(\begin{array}{c}
\frac{\delta^D(1-b)(1+\rho)}{1+\rho-(1-\delta^D)(1-\delta^I)(1-\Delta)}\left(1-\gamma^{-1}\frac{\Pi(y^*)}{\Pi(x^*)}\right) + \delta^D b\theta\left(1-\gamma^{-1}\right)\frac{\Pi(n)}{\Pi(x^*)} \\
+\frac{(1-b)\left(1-\delta^D\right)\left(1+\rho-\left(1-\delta^I\right)(1-\Delta)\right)}{1+\rho-b(1-\delta^D)(1-\delta^I)(1-\Delta)}\left(\begin{array}{c}
b\left(\theta\frac{\Pi(n)}{\Pi(x^*)} - \gamma^{-1}\frac{\Pi(y^*)}{\Pi(x^*)}\right) \\
+\frac{(1+\rho)(1-b)}{1+\rho-(1-\delta^D)(1-\delta^I)(1-\Delta)}\left(1-\gamma^{-1}\frac{\Pi(y^*)}{\Pi(x^*)}\right)
\end{array}\right)^+
\end{array}\right).
\end{aligned}$$

Therefore as in the baseline case Z_{coop} can be written as a function of $\Pi(x^*)$, $\frac{\Pi(y^*)}{\Pi(x^*)}$ and $\frac{\Pi(n)}{\Pi(x^*)}$, which is increasing in $\Pi(x^*)$, decreasing in $\frac{\Pi(y^*)}{\Pi(x^*)}$ and increasing in $\frac{\Pi(n)}{\Pi(x^*)}$. The same expression applies in the contractible and Nash cases if one replaces $\Pi(x^*)$ by $\Pi(m)$ or $\Pi(n)$ respectively and the profit ratios by 1. For ψ sufficiently convex, we can compare the innovation rates across the three cases by comparing the three Z_K.

Therefore, if $\Delta > \frac{1+\rho-b\left(1-\delta^D\right)}{\gamma(1+\rho)-b(1-\delta^D)}$, so that $y^* \geq x^*$, we must have $\delta^{cont} > \delta^{coop}$. Moreover, if relationships break in the Nash but not the cooperative case (so $V_0^{T,g} - V_A^s > W_1^T$ holds in the Nash but not the cooperative cases), then $\delta^{coop} < \delta^{Nash}$ for δ^D small enough (as then Z_{coop} is proportional to δ^D but Z_{Nash} is not).

B.10 Proof of Proposition 10

As in the baseline model, the reward from cooperation for a good match supplier is independent of his current productivity, as her next period productivity is independent of the current one. Denoting IC, this reward from cooperation, we get that the incentive constraints for a good match suppliers are

$$\gamma^2\varphi(x_2) \leq IC,\ \gamma\varphi(x_1) \leq IC \text{ and } \varphi(x_0) \leq IC.$$

The higher is current productivity, the larger is the incentive to deviate, so that we must have $x_0 \geq x_1 \geq x_2$.

Using (1), we can write welfare as

$$U = \frac{1+\rho}{\rho}\left(1 + \int_0^1 \theta_{jk} A_k W(x)\right),$$

with $W(x) \equiv \frac{\sigma}{\sigma-1}x^{\sigma-1} - x$ the normalized social surplus when the supplier's normalized investment is x. Note that on (n,m), W is increasing in x.

First part: We prove that when $\lambda = 1$, welfare is necessarily higher in the cooperative case relative to the Nash case. In this case, producers face the same distribution of productivities for the alternative supplier in the cooperative and Nash cases.

First, consider a producer who draws suppliers whose productivities are such that he switches neither in the Nash nor in the cooperative case. Then, the social welfare from that line is $\gamma^i W(x_i)$ with $i \in \{0,1,2\}$ in the cooperative case and $\gamma^i W(n)$ in the Nash case, and we have $\gamma^i W(x_i) > \gamma^i W(n)$ since $x_i > n$.

Then suppose draws such that the producer switches in both cases. Then the social welfare in the cooperative case is $\gamma^i \left((1-b) W(x_i) + b\theta W(n)\right)$ versus $\gamma^i (1-b+b\theta) W(n)$ in the Nash case. Here as well, we have: $\gamma^i \left((1-b) W(x_i) + b\theta W(n)\right) > \gamma^i (1-b+b\theta) W(n)$.

Finally, let us consider a situation where the producer switches in the cooperative case but not in the Nash case (the reverse being impossible). For instance, assume that the producer does not switch in the cooperative case when the previous good match's productivity is 1 and the alternative supplier's productivity is γ. This requires that $\gamma\left((1-b)\Pi(x_1) + b\theta\Pi(n)\right) < \Pi(x_0)$. At the same time, the switch occurs in the Nash case ($\gamma > \gamma^{Nash}$). Note that:

$$\begin{aligned} & \left((1-b)\Pi(x_1) + b\theta\Pi(n)\right) W(x_0) - \left((1-b) W(x_1) + b\theta W(n)\right) \Pi(x_0) \\ = & \left((1-b)\left[x_1^{\frac{-1}{\sigma}} - x_0^{\frac{-1}{\sigma}}\right] x_1 + b\theta n \left[n^{\frac{-1}{\sigma}} - x_0^{\frac{-1}{\sigma}}\right]\right) \frac{x_0}{\sigma-1} > 0, \end{aligned}$$

since $n < x_0$ and $x_1 \leq x_0$. As a result we must have $\left((1-b)\Pi(x_1) + b\theta\Pi(n)\right)/\Pi(x_0) > \left((1-b) W(x_1) + b\theta W(n)\right)/W(x_0)$. Hence we have that

$$W(x_0) > \gamma\left((1-b) W(x_1) + b\theta W(n)\right) > \gamma(1-b+b\theta) W(n).$$

Therefore in that case too, the social surplus is larger in the cooperative case than in the Nash case. The same logic applies to the other cases with a switch in the cooperative but not the Nash case.

We can then conclude that welfare is strictly higher in the cooperative than in the Nash case.

Second part. For simplicity, we consider parameters such that $x_0 = x_1 = x_2 = m$ (this occurs if b is high and ρ is low) and we focus on γ such that $\gamma^{Nash} = (1-b+b\theta)^{-1} < \gamma < \gamma^{coop} = \left(1-b+b\theta\frac{\Pi(n)}{\Pi(m)}\right)^{-1} < \gamma^2$.

First let us look at the first round of producers in the Nash case. Since $\gamma > \gamma^{Nash}$, we get that the producer will pick the supplier with the highest productivity. Hence the chosen supplier's productivity is 1 only if both the original and the alternative suppliers' productivities are 1, which only happens with probability 1/9. It is given by γ^2 if either supplier got a productivity γ^2, which occurs with probability $\frac{1}{3} + \frac{2}{3}\frac{1}{3} = \frac{5}{9}$ (either the first supplier got γ^2 or he did not but the second one did). Finally the chosen supplier's productivity is γ with probability $1 - 1/9 - 5/9 = 1/3$.

Let us then consider the second round of producers:

- With probability $\frac{1}{3}$, the previous supplier's productivity is γ^2. Therefore that supplier is chosen and the social surplus is given by $\gamma^2 W(n)$.

- With probability $\frac{1}{3}$, the previous supplier's productivity is γ.
  - Further, with probability $\frac{5}{9}$, the alternative supplier's productivity is γ^2. In that case, the producer chooses the alternative supplier and the (expected) social surplus is given by $(1-b+b\theta)\,\gamma^2 W(n)$.
  - Otherwise, the alternative supplier's productivity is weakly lower. In that case, the producer chooses the previous supplier and the social surplus is given by $\gamma W(n)$.
- With probability $\frac{1}{3}$, the previous supplier's productivity is 1.
  - Then, with probability $\frac{5}{9}$, the alternative supplier's productivity is γ^2, so that the social surplus is $(1-b+b\theta)\,\gamma^2 W(n)$.
  - With probability $\frac{1}{3}$, the alternative supplier's productivity is γ, so that the social surplus is $(1-b+b\theta)\,\gamma W(n)$.
  - Finally with probability $\frac{1}{9}$, the alternative supplier's productivity is 1, so that the producer keeps his previous supplier, leading to a social surplus $W(n)$.

Denote by L the set of lines for which the producer belongs to the second round. Then the expected social surplus for these lines in the Nash case is given by:

$$E\left(W_j^{Nash}|j\in L\right)=\frac{W(n)}{3}\left(\gamma^2\left(1+\frac{10}{9}(1-b+b\theta)\right)+\left(\frac{4}{9}+\frac{1}{3}(1-b+b\theta)\right)\gamma+\frac{1}{9}\right). \tag{B.131}$$

Let us similarly consider the cooperative case. Since $\gamma<\gamma^{coop}<\gamma^2$, switching occurs only when the previous producer's technology is 1 while the alternative supplier's one is γ^2. This event occurs with probability 1/9. The chosen supplier's productivity is 1 if the original supplier's productivity is 1 and the alternative supplier did not draw γ^2: this happens with probability $\frac{1}{3}\left(1-\frac{1}{3}\right)=\frac{2}{9}$. The chosen supplier's productivity is γ if the original supplier's productivity is γ (as then she is always picked), which happens with probability $\frac{1}{3}$. Finally the chosen supplier's productivity is γ^2 with probability $1-\frac{1}{3}-\frac{2}{9}=\frac{4}{9}$. Let us then consider the second round of producers:

- With probability $\frac{1}{3}$, the previous supplier's productivity is γ^2. Therefore that supplier is chosen and the social surplus is given by $\gamma^2 W(m)$.

- With probability $\frac{1}{3}$, the previous supplier's productivity is γ. Therefore that supplier is chosen and the social surplus is given by $\gamma W(m)$.
- With probability $\frac{1}{3}$, the previous supplier's productivity is 1.
  - Then, with probability $\frac{4}{9}$, the alternative supplier's productivity is γ^2, so that the social surplus is $((1-b)W(m)+b\theta W(n))\gamma^2$.
  - Otherwise, with probability $\frac{5}{9}$, the previous supplier is chosen and the social surplus is $W(m)$.

The expected social surplus for these lines in the cooperative case is given by:

$$E\left(W_j^{coop}|j\in L\right)=\frac{1}{3}\left(\gamma^2\left(W(m)+\frac{4}{9}(1-b)W(m)+b\theta W(n)\right)+\gamma W(m)+\frac{5}{9}W(m)\right). \tag{B.132}$$

Take the difference between (B.132) and (B.131):

$$\begin{aligned}&E\left(W_j^{coop}|j\in L\right)-E\left(W_j^{Nash}|j\in L\right)\\ =\;&\frac{W(m)}{3}\left[\begin{array}{c}\left(\left(1+\frac{4}{9}(1-b)\right)\gamma^2+\gamma+\frac{5}{9}\right)\left(1-\frac{W(n)}{W(m)}\right)-\frac{1}{9}(\gamma-1)\frac{W(n)}{W(m)}\\ -\left(\frac{2}{3}\gamma+\frac{1}{3}\right)\left[(1-b+b\theta)\gamma-1\right]\frac{W(n)}{W(m)}\end{array}\right].\end{aligned}$$

Since $(1-b+b\theta)\gamma-1>0$, this expression shows that for n sufficiently close to m (which is obtained by increasing β), then $E\left(W_j^{coop}|j\in L\right)<E\left(W_j^{Nash}|j\in L\right)$. When λ is close to 0, then nearly all lines belong to L therefore we also have $U^{Nash}>U^{coop}$.

C Appendix: Data and regression

In this appendix, we describe our data and run some additional regressions involving other countries than the US and Japan.

C.1 Data

We use the patent data set of the OECD (OECD, 2015) which is built on data from the European Patent Office (EPO). The data set records all patent applications to the EPO. The year of a patent corresponds to the earliest year of application. Each patent is associated to a country depending on the address of its inventors (if a patent is associated with inventors from several countries, we weight each patent x country combination according to the share of inventors from each country). We restrict attention to patents granted by 2009 (from 2010, more than 10% of patent applications have been neither granted nor rejected yet, in addition, few patents have had the time to receive citations, which is problematic for the accuracy of our measure of generality). Furthermore we consider only the patents that have never been withdrawn. The generality measure is computed by the OECD and we use it directly (it is computing using citations made by other EPO patents). We drop patents for which the measure is not computed (either because the patents have not received any citations—indicating their low value—or because of missing information).

To compute our "Rauch index," we first use a PATSTAT file which attributes 2, 3 or 4 digit NACE Rev 2 codes (depending on the sector) to each patent. Some patents might have multiple weighted associated NACE codes. Then, we use the liberal classification from Rauch (1999) which labels each 4 digit SITC 2 code as either "goods traded on organized exchange", "reference priced" or "differentiated". We attribute a "Rauch index" 1 to goods which are labeled as differentiated, and give an index of 0 to the other goods. We convert this into SITC 3 codes using the conversion table from `http://econweb.ucsd.edu/~jrauch/rauch_classification.html`. This is close to a one-to-one conversion. We use a conversion table from SITC 3 to NACE Rev 1 from World Integrated Trade Systems (`http://wits.worldbank.org/product_concordance.html`) to convert the SITC 3 to NACE Rev 1. If a NACE Rev 1 is associated with multiple SITC 3 codes we take the average value of the SITC 3 codes. Finally, we convert the NACE Rev 1 into NACE Rev 2 using a concordance table from Eurostat (`http://ec.europa.eu/eurostat/web/nace-rev2/correspondence_tables`). Again, if a NACE Rev 2 code corresponds to multiple NACE Rev 1 we use an unweighted average. We use the two digit NACE Revision 2 in our analysis, which leaves 27 distinct NACE categories with a Rauch index varying between 0.16 and 1. The Rauch index of each sector is given in Table C.1a below.

Data on trust (used for the additional regression below) come from the World Value Survey longitudinal data file 1981-2014. We focus on questions G007_33 which ask to respondents whether they trust people they know personally and G007_34 which ask them whether they trust people they meet for the first time. There are 4 levels of trust and we linearly transform each variable so that they are in $[0, 1]$ with a high value corresponding to a high level of trust. These two questions where only asked from 2005 onward. For each country and wave, we average the answer given by individuals using the weights provided in the dataset (as recommended we censor weights below 0.33 and above 3).[63] For countries which are present in multiple waves, we further average across waves. Then, we take the difference between the trust level towards people met for the first time and trust towards people already known to define our "Dif Trust" variable: a low level indicates a relatively higher trust towards known people which favors the establishment of cooperation in long-term relationships. Since our trust measures are all from 2005, we restrict attention to patents filed after 1995. Furthermore, we focus on countries with more than 100 patents with generality data, though little depends on this choice. Table C.1b reports the full list of countries with their differential trust measure.

C.2 Regression

We first run a regression analogous to that of table 2 but at the lowest level of disaggregation available in the patent data (2 to 4 digit NACE level depending on the sector). The results are in table C.2. There is little difference in the estimates.

We now report results on the larger set of countries which are consistent with our theory that cooperation in an environmental of weak contractability should deter general innovations. Direct evidence for whether countries are in the 'cooperative' or 'Nash' equilibrium is difficult to come by, we use the Dif Trust variable built above using World Values Survey data. A high value corresponds to a high relative trust in strangers, which reduces the scope for the establishment of cooperation in existing relationships. Japan gets the fourth lowest score and France the lowest, whereas the United States is slightly above the middle with South Africa receiving the highest score. Trusting that these values are relatively constant, we include patents filed ten years before and until 2009. We focus on countries with more than 100 patents with generality data, though little depends on this choice. This leaves us with 318,197 patents and a total of 24 countries. We then run the following regression:

$$gen_{i,t,s,c} = \beta_0 + \beta_1 Diff\,Trust_c \times Differentiated_s + \delta_c + \delta_s + \delta_t + \epsilon_{i,t,s,c},$$

where $gen_{i,t,s,c}$ is the generality of patent i, which was filed in year t, corresponds to sector s and whose inventors were from country c, $Differentiated_s$ is the measure of the importance of contractibility issues in sector s and δ_c, δ_s and δ_t are country, sector and

[63] For the United Kingdom we use the value corresponding to Great Britain (which excludes Northern Ireland) and for Serbia we use results conducted for Serbia and Montenegro.

NACE, rev. 2	Label	Degree of Differentiation
11	Beverages	0.16
12	Tobacco products	0.17
24	Basic metals	0.23
19	Coke and refined petroleum products	0.25
10	Food products	0.28
20	Chemicals and chemical products	0.50
17	Paper	0.55
21	Basic pharmaceutical products	0.61
22	Rubber and plastic products	0.78
23	Other non-metallic mineral products	0.78
27	Electrical equipment	0.78
16	Woods and similar	0.79
13	Textiles	0.84
29	Motor vehicles, trailers and semi-trailers	0.85
26	Computer, electronic and optical products	0.88
25	Fabricated metal products	0.91
32	Other manufacturing	0.93
18	Print and repro of recorded media	0.94
43	Specialized construction activities	0.95
15	Leater and related products	0.95
30	Other transport equipment	0.97
31	Furniture	0.98
28	Machinery and Equipment n.e.c.	0.99
14	Wearing apparel	1.00
62	Computer Programming, consultancy	1.00

	Trust person you know	Trust person you just met	Differenc
France	0.87	0.44	-0.
Russia	0.67	0.27	-0.
Turkey	0.67	0.29	-0.
Japan	0.65	0.28	-0.
Great Britain	0.83	0.46	-0.
Spain	0.75	0.39	-0.
South Korea	0.65	0.30	-0.
Singapore	0.73	0.38	-0.
Netherlands	0.71	0.36	-0.
China	0.65	0.30	-0.
Germany	0.71	0.36	-0.
Canada	0.81	0.47	-0.
Australia	0.79	0.45	-0.
USA	0.74	0.41	-0.
Brazil	0.54	0.21	-0.
Hungary	0.70	0.39	-0.
Finland	0.80	0.49	-0.
Norway	0.86	0.56	-0.
Poland	0.65	0.35	-0.
Switzerland	0.77	0.48	-0.
Sweden	0.81	0.54	-0.
Italy	0.57	0.31	-0.
India	0.63	0.38	-0.
South Africa	0.62	0.39	-0.

(a) Extend of differentiation for NACE Rev.2 product categories (details in text)

(b) Trust measures for countries used in regression (details in text)

Table C.1: Country and Product category information used in the regressions

	(I)	(II)	(III)
	Generality	Generality	Generality
US.	0.055***	0.024***	-0.007
	(13.47)	(6.08)	(1.36)
US x Differentiated			0.077 *
			(1.79)
Fixed Effects	No	Year, NACE	Year, NACE
Observations	337913	337913	337799

Standardized beta coefficients; t statistics in parentheses. Std. errors clustered at NACE x country level for (III)

* $p < 0.10$, ** $p < 0.01$, *** $p < 0.001$

Table C.2: Regression Results for 4 digit NACE

	(I)
	Generality
Trust x Differentiated	0.015*
	(1.84)
Fixed Effects	NACE, Country, Year
Observations	318197

Standardized beta coefficients; t statistics in parentheses.

Column (I) uses 2 digit NACE (details in text)

* $p < 0.10$, ** $p < 0.05$

Table C.3: Regression results for 24 countries

year fixed effects. shows this regression to be consistent with our predictions as a high relative trust in strangers (i.e. a lower scope for cooperation in existing relationships) is associated with relatively more general patents in more differentiated industries. The result is no longer statistically significant if we employ the NACE code at 4 digits.