LONG-TERM RELATIONSHIPS: STATIC GAINS AND DYNAMIC INEFFICIENCIES

David Hémous
University of Zürich

Morten Olsen
University of Copenhagen

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 noncontractible. 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 noncooperative equilibrium. We contrast this with a setting with relationship-specific innovations that we show are encouraged by the establishment of relational contracts. We illustrate the predictions of the model 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)

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*"

The editor in charge of this paper was Nicola Gennaioli.

Acknowledgments: 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). Hémous is a Research Affiliate at CEPR.

E-mail: david.hemous@econ.uzh.ch (Hémous); mo@dce.au.dk (Olsen)

Journal of the European Economic Association 2018 16(2):383–435 DOI: 10.1093/jeea/jvx019

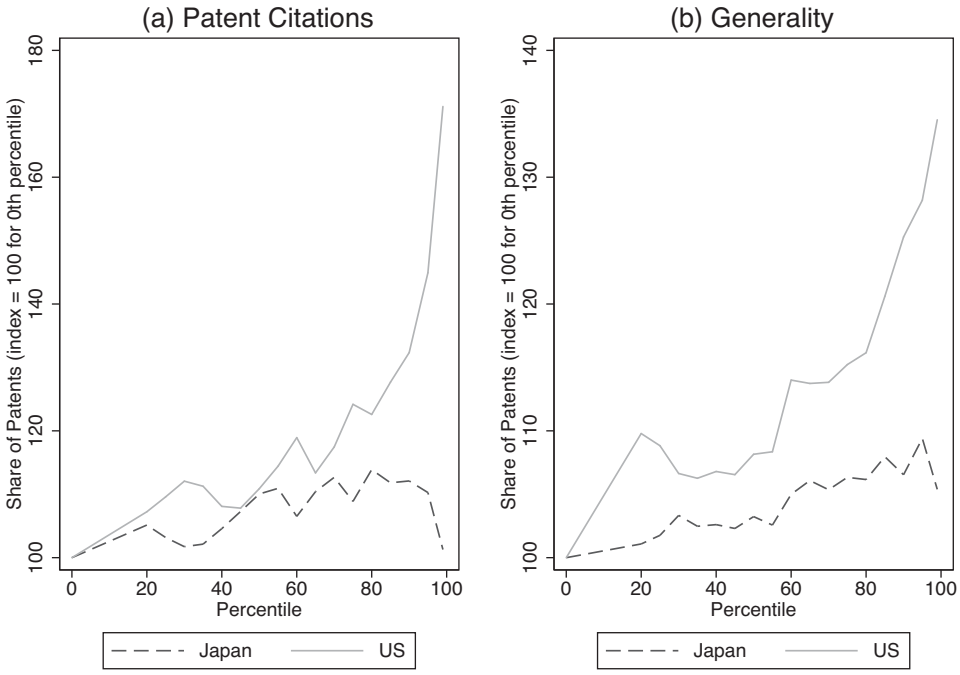

FIGURE 1. Relative importance of Japanese and United States patents.

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: although both countries are innovative, the U.S. innovations tend to be broader.

Figure 1 provides an illustration of this using patents with at least one citation filed with the European Patent Office (EPO). Out of all patents filed with the EPO 24.3% are from the United States and 21.5% are Japanese. The figure shows that the share of US patents increases by more than 70% when focusing on the top 1% most cited patents, whereas it remains largely constant for Japanese patents. 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 the same pattern for generality: the United States has disproportionately more patents with high generality measure. In Section 5, we perform a more systematic analysis and show that the generality of U.S. 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 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 nonrepeated framework, noncontractibility 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

1. The 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 patent i has received from patent class j and n_i is the total number of patent classes citations received by patent i.

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 that 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 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; 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]ierarchical 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 that 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

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.

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, 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 whereas 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 toward the suppliers with whom he has already worked. This implies that an outsider with a

4. Bessen (2008) estimates that more general patents have more private value (controlling for the number of citations). The broadest of innovations is 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; 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 United States, external innovation are responsible for 80% of economic growth.

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

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, Antrás, and Helpman 2007) and empirical (Boehm 2013, Cowan and Neut 2007; 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, 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) that 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 (Online Appendix B).

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.

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. Cooperation and Rigidity of Relationships

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 infinitelylived 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

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

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 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 $(\theta_{jk}A_k)^{1/(\sigma-1)}$). Producing one high quality intermediate input requires one unit of 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 that 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.

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 $R(x) \equiv x^{(\sigma-1)/\sigma}$, and joint profits as $\theta_{jk}A_k\Pi(x)$ where $\Pi(x) \equiv x^{(\sigma-1)/\sigma} - x$. 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.

2.1. 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. We further make the classic assumption that revenues and expenditures of the parties are nonverifiable 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

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.

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.

2.2. *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 whereas the other suppliers use $\gamma^{-1}A$.

2.3. *Timeline*

The overall timeline within each period is as follows (Figure 2):

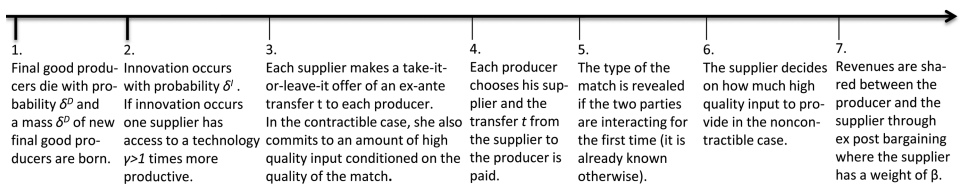

FIGURE 2. Timeline.

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.

2.4. 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 2.6). 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 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.

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).

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.

We define a "cooperating good match supplier" as a good match supplier with whom no deviation triggering a punishment has occurred. Similarly a "noncooperating 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 that 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 noncooperating 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.*

(S4) *A cooperating good match supplier becomes noncooperating 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 noncooperating 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.*

2.5. 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 that 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).

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

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

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 toward the innovator or stay with the now outdated good match supplier (per S3). This decision depends on parameters and is the subject of Section 2.6. 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

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

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.

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

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

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

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

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 noncooperating, 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, whereas the cooperating good match supplier can capture the surplus of a relationship with her over the producer's second best option. Therefore, we obtain

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

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 noncooperating good match to only invest n from now on. Per S5, he optimally decides between trying a new supplier or sticking with this noncooperating good match. Here, we focus on parameters for which the former occurs, so that the continuation value of a noncooperating good match supplier (with that producer) is 0 (see Online 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

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

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).

16. Technically, the presence of a noncooperating 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 noncooperating good match. Nevertheless, if the producer always prefers trying a new supplier to staying with a noncooperating good match, then $V_0^T = V_0^{T,n}$ and (8) holds (see Appendix A.1 that deals with this issue rigorously).

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

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

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 that 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. that is, $y^* \geq x^*$ (with equality if $x^* = m$). Appendix A.1 derives explicitly the right-hand side of both IC constraints that 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 nonexisting. The switching cost here is the risk of finding a bad match.[18] A good match

17. Online 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 noncooperating 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 "pairwise 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

supplier benefits from having been revealed as such. This informational advantage acts as a fixed cost that pushes the producer to stick to the same supplier, who can then capture the associated rents. The prospect of capturing these rents induces 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.

2.6. 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

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

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 that 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 following period (that innovator is no longer viewed as a threat).[20] As a result,

such players (similar to our bad matches here) serves as an expected cost of establishing a new relationship and enables cooperation.

19. If θ is low enough, it is in fact impossible to build an equilibrium that 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.

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

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, whereas 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

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

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 the following proposition.

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 encourages 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.

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.

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 2.6 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 results hold more generally. Every period, one supplier gets a new idea that turns into a useful innovation with probability δ^I if the potential innovator invests $A\psi(\delta^I)$ (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

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.

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.

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

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

Finally, in the cooperative case, we get

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

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

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 that 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 $\omega = \delta^D/(1 - (1 - \delta^D)b)$.[24] Hence, assuming that the steady state

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.

has been reached, the innovator solves the problem

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

for $K \in \{con, Nash, coop\}$. We denote by $Z_K = \omega V_{I,K}^{s,b}\left(\delta^I\right) + (1-\omega)\,V_{I,K}^{s,g}\left(\delta^I\right)$, the expected total value of an innovator. The first order condition $\psi'(\delta^I) = \gamma Z_K$ 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

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

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 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 that 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 that 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 toward $\delta^{Nash} > \delta^{coop}$, 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

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

As shown in Online 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

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

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 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 (15–17), we obtain

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

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 $\delta^{Nash} > \delta^{coop}$ ($\gamma < (1 - b(1-\delta^D))^{-1}$ 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 Online 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*

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.

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

$\delta^D < \theta(\Pi(n)/\Pi(m))$, then for an intermediate range of innovation sizes,

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

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 $\delta^D < \theta(\Pi(n)/\Pi(m))$, then the innovation rate is higher in the Nash than the contractible case, $\delta^{Nash} > \delta^{coop}$, if and only if

$$\gamma \in \left(\frac{1 - \delta^D \frac{\Pi(m)}{\Pi(n)}}{1 - \delta^D \frac{\Pi(m)}{\Pi(n)} - b\,(1-\theta)\left(1-\delta^D\right)}, \frac{1}{1 - b\left(1-\delta^D\right)} \right).$$

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 $\delta^{Nash} > \delta^{coop}$.

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. Although 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 $\delta^D < \theta(\Pi(n)/\Pi(m))$, and for an intermediate range of innovation sizes,*

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

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 $\psi(\delta^I)AN$), so that the innovation problem is independent of the number of products.

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.

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 Online 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.

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 that 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

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 that, as ours, features heterogeneity on the supplier side.

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 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 (Figure 3).

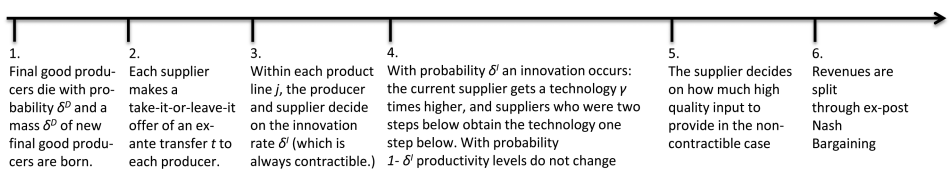

FIGURE 3. Timeline.

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 that 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

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

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 Online 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 Online 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 whereas 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

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

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

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

Since $x^* > n$, $\Pi(n) < \Pi(x^*)$ 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 ($\delta_0^{I,coop} > \delta^{I,Nash}$), whereas 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 Online 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 U.S. and Japanese labor markets and suggest that lower labor turnover 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 noncontractible 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 semiconductor 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 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, etc.) 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

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 United States 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 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 × 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

a touchscreen). In 2008, Apple's retail share was 2%, in 2015 it reached 44% (Euromonitor International 2015).[32]

As a result, although Japan was the role model for business relationships in the management literature in the 1980s and the early 1990s, 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 that again focuses on the 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 Online 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) that 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 339,681 Japanese or U.S. patents (44.6% are US patents).[33] Table 1 shows summary statistics on the two variables

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" that 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 toward 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 Online Appendix B.10, we consider a model in which relationships also allow unproductive firms to remain in operation, though without a financial sector.

33. The generality data are computed using citations by EPO patents to the Japanese or U.S. patent referred by their they EPO number but not their USPTO or JPO numbers. As a consequence, although there are more U.S. than Japanese patents with positive citation, there are more Japanese than U.S. patents with a generality number.

TABLE 1. Summary statistics.

	Count	Mean	Median	Std. Dev	Min	Max
Generality	339,681	0.40	0.48	0.28	0.00	0.93
Differentiated	339,567	0.98	1.00	0.04	0.68	1.00

TABLE 2. Regression results for 2 digit NACE.

	(I) Generality	(II) Generality	(III) Generality
United States	0.055***	0.027***	0.024
	(13.60)	(6.83)	(1.04)
United States × Differentiated			0.044***
			(3.11)
Fixed effects	None	NACE, Year	NACE, Year
Observations	339,681	339,681	339,567

Standardized beta coefficients; t statistics in parentheses. Std. errors clustered at NACE × country level for (III). $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

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 analog of Figure 1, which shows that U.S. 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 U.S. patent is 5.5% 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 a NACE code and reduces the estimate by around half.[34] 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 that slightly reduces the number of observations (details in Online Appendix C).[35] We include fixed effects for the NACE codes as well as 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

34. 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.

35. PATSTAT attributes 2–4 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 Online Appendix shows that our results are similar when we use the lowest level of disaggregation available.

differentiated products, which is both substantially positive and significant (standard errors are clustered at NACE × country level). The difference in generality between a U.S and a Japanese patent increases by 4.4% of a standard deviation when moving from one sector to a sector that 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 Online Appendix C, we use a larger set of countries to perform an analogous analysis and find results consistent with our theory.

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 that 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 Online Appendix B.8). With exogenous probability, in some periods, suppliers can engage in relationship-specific innovations, whereas in others a potential supplier gets the opportunity to undertake a general innovation that 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.[36] Second, the higher is the rate of relationship 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).[37] At the same time, for a given line, the general innovator's rents are proportional to the technology used with a coefficient that 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 that pushes toward less general innovation in the equilibrium where more relationship specific innovation occurs.[38] 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* $\delta^D < \theta(\Pi(n)/\Pi(m))$, *then for an intermediate range of innovation sizes,*

$$\gamma^A \in \left(\frac{1 - \delta^D \frac{\Pi(m)}{\Pi(n)}}{1 - \delta^D \frac{\Pi(m)}{\Pi(n)} - b\,(1-\theta)\left(1-\delta^D\right)}, \frac{1}{1 - b\left(1-\delta^D\right)} \right)$$

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.[39]

36. Note that as in Section 4, we assumed that relationship-specific innovations are contractible. Were 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.

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

38. 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.

39. 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

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; whereas 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 $\psi^A\left(\delta^A\right) = \widetilde{\psi}^A\left(\delta^A\right)^{\psi}/\psi$ and $\psi^B\left(\delta^B\right) = \widetilde{\psi}^B\left(\delta^B\right)^{\psi}/\psi$. 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

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.

good match he knew. This assumption simplifies the exposition in the contractible and Nash cases.[40]

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.[41] 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 $\left(1-\delta^D\right)/(1+\rho) \times ((1-\delta^I)V_1^s + \delta^I \gamma W_1^s)A_t$.[42] 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

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

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 toward 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^*

40. 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.

41. 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.

42. 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.

is ambiguous and the arrival of an innovation may weaken cooperation in established relationship. Nevertheless, in Online Appendix B.9, we show that $\Delta \geq (1+\rho-b(1-\delta^D))/(\gamma(1+\rho)-b(1-\delta^D))$ is a sufficient condition to ensure that $y^* \geq x^*$.

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

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

with $K \equiv (1-b)(1-\delta^D)(1-\delta^I)/(1+\rho-(1-\delta^D)(1-\delta^I)(1-\Delta)) > 0$. 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 $\gamma^{-1}\Pi(y^*)$) relative to switching to a frontier good match supplier (with profits $\Pi(x^*)$) 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

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

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.

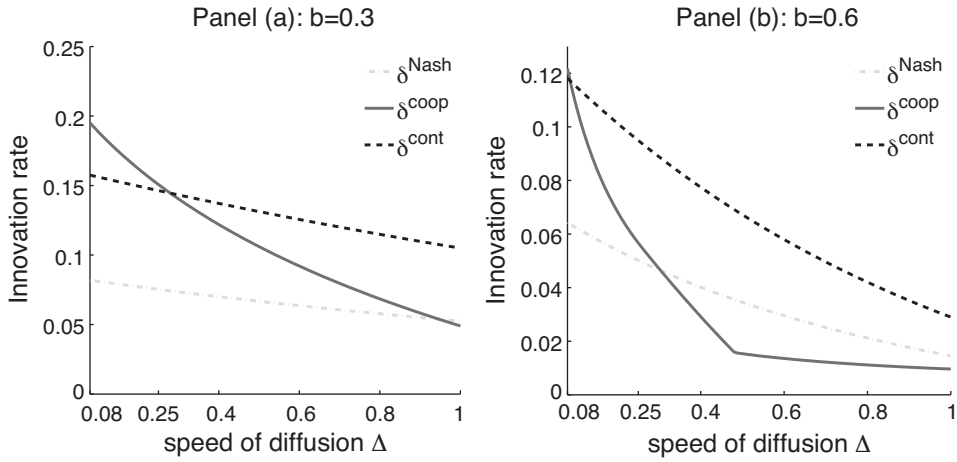

FIGURE 4. Innovation rate and speed of diffusion.

PROPOSITION 6. *Consider parameters such that*

$$\Delta > \frac{1+\rho-b\left(1-\delta^{D}\right)}{\gamma\left(1+\rho\right)-b\left(1-\delta^{D}\right)}$$

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 4 that 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 $b = 0.6$.[43] In both cases, for fast diffusion, the innovation rate in the cooperative case is lower than both in the

43. 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 $\psi(\delta) = \delta^{2}/5$.

Nash and contractible 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 that 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. although 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.

Appendix: 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 noncooperating good match in a period without innovation and with a new outdated supplier over a noncooperating 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 noncooperating good match in a period without innovation and with a new outdated supplier over a noncooperating good match in a period with innovation, the other cases are included in Online 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 Online 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.[44] 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 ($z = s$), 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

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

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

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 the total value of the producer and the innovator ($V_I^{T,g}$) is higher than the surplus value of the old relationship ($W_1^T - V_A^s$):[45]

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

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

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

With probability $1-b$ the relationship turns out to be good delivering profits $\Pi(x^*)$ 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 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

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

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

45. 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.

general the incentive constraints obey

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

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 noncooperating 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[46]

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

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

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

$$\left(1-\delta^I\right) V_1^s + \delta^I \gamma W_1^s = \frac{1+\rho}{1+\rho-b\left(1-\delta^D\right)\delta^I\gamma} \tag{A.10}$$
$$\times\left(\left(1-\delta^I\right) V_1^s + \delta^I\gamma\left(\frac{1}{\gamma}\Pi\left(y^*\right) - \left((1-b)\,\Pi\left(x^*\right) + b\theta\Pi(n)\right)\right)^+\right).$$

Finally note that V_1^T must satisfy (4) that 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 noncooperating 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 noncooperating good match (this is the general expression, see footnote 16). Indeed,

46. That still requires that switching to the innovator is a better option than switching to a potential noncooperating good match when the producer knows one (see footnote 44).

16. Technically, the presence of a noncooperating 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

the outside option for the producer is to start a new relationship, but should he do so, he would now know a noncooperating 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 noncooperating good match or to start a new relationship, now knowing two noncooperating 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 noncooperating good match.

As mentioned in the text, depending on parameters, there is a number of different 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 Online 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 noncooperating good match, and, in periods with innovation, the producer would prefer both the innovator or an outdated new supplier to an (outdated) noncooperating 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 $\left(1-\delta^I\right) V_1^s + \delta^I \gamma W_1^s$ that 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.

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

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

relationship when the producer knows a noncooperating good match. Nevertheless, if the producer always prefers trying a new supplier to staying with a noncooperating good match, then $V_0^T = V_0^{T,n}$ and (8) holds (see Appendix A.1 that deals with this issue rigorously).

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

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

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 ($\Pi(x^*) - \theta\Pi(n)$) and in periods with innovation ($\gamma^{-1}(\Pi(y^*) - \theta\Pi(n))$). 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

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

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:

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

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

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

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

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

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 Online 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 A.1 [Symmetry and Information (SI)]. (i) *For any history belonging to* $\cup_k H_t^n(j,k)$ *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* $\cup_k H_t^n(j,k)$ *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's 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 A.1(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 A.1(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 A.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 A.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 A.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.[47] Finally, we impose the following condition.

CONDITION A.4. [No Investment in Bad Matches]. *Normalized investment levels in bad matches are given by the Nash investment level, n.*

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 A.1. *In any symmetric SPNE satisfying Conditions A.1–A.4, agents' strategies are given as in Proposition 1.*

47. 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.

Proof. It is direct to check that the strategies of Proposition 1 obey Conditions A.1–A.4. Online Appendix B.2 shows that Conditions A.1–A.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 Online Appendix B.3.

PROPOSITION A.2. (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 A.1. When a producer would always rather try a new supplier than work with a noncooperative 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 ($\gamma b(1-\delta^D)(2-\delta^I) < 1+\rho$ 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 that 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 toward 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 noncooperating 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.[48] We prove the following proposition (where the innovation rate in the alternative cooperative case refers to the highest equilibrium level) in Online Appendix B.6.

PROPOSITION A.3. (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* ($\rho < \left(\gamma/\delta^{coop2} - 1\right)\left(1 - b\left(1 - \delta^{D}\right)\right) + b\left(1 - \delta^{D}\right)\delta^{coop2}\left(\gamma - 1\right)$ *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* $\gamma \in \left(\gamma^{cont}, \gamma^{coop2}\right)$ *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 Online Appendix B.6, we show that producers would switch to the innovator if and only if

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

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 that occurs if the innovator turns out to be a

48. 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.

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 toward 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.[49] As an innovator captures more value from producers who are not in an ongoing good match relationship, 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 A.3.[50] 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 A.3).[51]

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

49. In the cooperative case, the share of producers previously not in a good match is given by $(\delta^D + b\delta^I(1-\delta^D))/(1-b(1-\delta^D)(1-\delta^I))$ when $\gamma > \gamma^{coop2}$ but by $\delta^D/(1-b(1-\delta^D))$ when $\gamma < \gamma^{coop2}$ or in the Nash or contractible cases.

50. The condition in the Proposition A.3 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.

51. 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.

Nash case for an intermediate range of innovation sizes provided that the death rate of producers is low enough.[52]

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-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.[53] 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, and so forth—firms are more likely to choose business partners already in operation.

In the cooperative case, we consider an equilibrium that is similar to that described in Section 2. In particular, there is no cooperation in bad matches and there are 3

52. We obtain that $\gamma^{Nash} = \gamma^{cont} = [1 - b + b\theta - (1-\delta^D)b^2(1-\theta)/(1+\rho-b(1-\delta^D)\ (1-\delta^I + \delta^I\gamma))]^{-1}$. 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.

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

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, 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 $\gamma \in \left(\gamma^{Nash}, \gamma^{coop}\right)$, 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 $\gamma \in \left(\gamma^{Nash}, \gamma^{coop}\right)$ 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 Online 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 Online Appendix B.10 the following proposition.

PROPOSITION A.4. (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 standpoint, 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.[54] 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.

References

Acemoglu, Daron and Jorn-Steffen Pischke (1998). "Why Do Firms Train? Theory and Evidence." *The Quarterly Journal of Economics*, 113, 79.

Acemoglu, Daron, Philippe Aghion, and Fabrizio Zilibotti (2003). "Distance to Frontier, Selection, and Economic Growth." *Journal of European Economic Association*, 4, 37–74.

Acemoglu, Daron, Pol Antràs, and Elhanan Helpman (2007). "Contracts and Technology Adoption." *American Economic Review*, 97(3), 916–943.

Aghion, Philippe and Patrick Bolton (1987). "Contracts as a Barrier to Entry." *The American Economic Review*, 77(3), 388–401.

Aghion, Philippe, Ufuk Akcigit, and Peter Howitt (2015). "Lessons from Schumpeterian Growth Theory." *The American Economic Review: Papers & Proceedings*.

Akcigit, Ufuk and Bill Kerr (2016). "Growth through Heterogeneous Innovations." *Working Paper*.

Allen, Franklin, Jun Qian, and Meijun Qian (2005). "Law, Finance, and Economic Growth in China." *Journal of Financial Economics*, 77, 57–116.

Allen, Franklin, Rajesh Chakrabarti, Sankar De, Jun Qian, and Meijun Qian (2006). "Financing Firms in India." *World Bank Policy Research Working Paper no. 3975*.

Allen, Franklin, Rajesh Chakrabarti, Sankar De, Jun Qian, and Qian Meijun (2008). *Emerging Giants: China and India in the World Economy* Law, Institutions and Finance in China and India.

Banerjee, Abhijit and Esther Duflo (2000). "Reputation Effects and the Limits of Contracting: A Study of the Indian Software Industry." *The Quarterly Journal of Economics*, 115, 989–1017.

54. 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.

Bessen, James (2008). "The Value of U.S. Patents by owner and patent characteristics." *Research Policy*, 37, 932–945.
Blinder, Alan and Alan B. Krueger (1996). "Labor Turnover in the USA and Japan: A Tale of Two Countries." *Pacific Economic Review*, 1, 27–57.
Board, Simon (2011). "Relational Contracts and the Value of Loyalty." *The American Economic Review*, 101, 3349–3367.
Boehm, Johannes (2013). "The Impact of Contract Enforcement Costs on Outsourcing and Aggregate Productivity."
Bolton, Michele, Roger Malmrose, and William Ouchi (1994). "The Organization of Innovation in the United States and Japan: Neoclassical and Relational Contracting." *Journal of Management Studies*, 31, 653–679.
Bonfiglioli, Aslessandra and Gino A. Gancia (2014). "Growth, Selection and Appropriate Contracts." *Review of Economic Dynamics*, 17, 21–38.
Bresnahan, T. and M. Trajtenberg (1995). "General Purpose technologies 'Engines of Growth'?" 65, 83.
Brown, Martin, Armin Falk, and Ernst Fehr (2004). "Relational Contracts and the Nature of Market Interactions." *Econometrica*, 72, 747–780.
Calzolari, Giacomo and Giancarlo Spagnolo (2009). "Relational Contracts and Competitive Screening."
Colinson, Simon and David Wilson (2006). "Inertia in Japanese Organizations: Knowledge Management Routines and Failure to Innovate." *Organization Studies*, 27, 1359–1387.
Cowan, Kevin and Alejandro Neut (2007). "Intermediate Goods, Institutions and Output per Worker." *Central Bank of Chile Working Papers*, 420.
Dore, Ronald (1983). "Goodwill and the Spirit of Market Capitalism." *The British Journal of Sociology*, 34, 459–482.
Dujarric, Robert and Andrei Hagiu (2009). "Capitalizing on Innovation: The Case of Japan." *Harvard Business School Working Paper 09-114*.
International, Euromonitor (2015). "Mobile Phones in Japan." http://www.euromonitor.com/mobile-phones-in-japan/report.
Fauli-Oller, Ramon, Joel Sandonis, and Juana Santamaria (2011). "Downstream mergers and upstream investment." *The Manchester School*, 79, 884–898.
Francois, Patrick and Joanne Roberts (2003). "Contracting Productivity Growth." *The Review of Economic Studies*, 70, 59–85.
Ghosh, Parikshit and Debraj Ray (1996). "Cooperation in Community Interaction Without Information Flows." *The Review of Economic Studies*, 63, 491–519.
Grossman, Sanford J. and Oliver Hart (1986). "The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration." *The Journal of Political Economy*, 94, 691.
Hall, Bronwyn and Manuel Trajtenberg (2004). "Uncovering GPTs with Patent Data." *NBER Working paper 10901*.
Hall, Bronwyn, Adam Jaffe, and Manuel Trajtenberg (2001). "The NBER Patent Citation Data File: Lessons, Insights and Methodological Tools." *NBER Working Paper 8498*.
Helper, Susan and Rebecca Henderson (2014). "Management Practices, Relational Contracts, and the Decline of General Motors." *NBER Working paper 19867*.
Helper, Susan (1991). "Strategy and Irreversibility in Supplier Relations: The Case of the U.S. Automobile Industry." *The Business History Review*, 65, 781–824.
Helpman, Elhanan (1998). *General Purpose Technologies and Economic Growth*. MIT Press.
Hoshi, Takeo and Anil Kashyap (2004). "Japan's Financial Crisis and Economic Stagnation." *The Journal of Economic Perspectives*, 18(1), 3–26.
Hsieh, Chang-Tai and Peter Klenow (2009). "Misallocation and Manufacturing TFP in China and India." *Quarterly Journal of Economics*, 124, 771–807.
Inderst, Roman and Christian Wey (2011). "Countervailing Power and Dynamic Efficiency." *Journal of the European Economic Association*, 9, 702–720.
Inderst, Roman and Greg Shaffer (2007). "Retail Mergers, Buyer Power and Product Variety." *The Economic Journal*, 117, 45–67.

Jaffe, Adam B., Manuel Trajtenberg, and Rebecca Henderson (1993). "Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations." 108(3), 577–598.
Johnson, Simon, John McMillan, and Christopher Woodruff (2002). "Courts and Relational Contracting." *Journal of Law, Economics and Organization*, 18, 221–277.
Jovanovic, Boyan and Peter L. Rousseau (2005). "General Purpose Technologies." 18.
Kranton, Rachel E. (1996). "Reciprocal Exchange: A Self-Sustaining System." *The American Economic Review*, 86(4), 830–851.
Kushida, Kenji (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, Stewart (1963). "Non-Contractual Relations in Business: A Preliminary Study." *American Sociological Review*, 28, 55.
Macchiavello, Rocco and Ameet Morjaria (2015). "The Value of Relationships: Evidence from a Supply Shock to Kenyan Rose Exports." *American Economic Review*, 105(9), 2911–2945.
Macchiavello, Rocco and Ameet Morjaria (2014). "Competition and Relational Contracts: Evidence from Rwanda's Coffee Mills."
MacLeod, W. Bentley and James M. Malcomson (1989). "Implicit Contracts, Incentive Compatibility, and Involuntary Unemployment." *Econometrica*, 57, 447–480.
Nunn, Nathan (2007). "Relationship-Specificity, Incomplete Contracts, and the Pattern of Trade." *Quarterly Journal of Economics*, 122, 569–600.
Pudelko, Markus and Mark Mendenhall (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, James E. (1999). "Networks Versus Markets in International Trade." *Journal of International Economics*, 48, 7–35.
Thompson, Peter and Melanie Fox-Kean (2005). "Patent Citations and the Geography of Knowledge Spillovers: A Reassessment." 95, 450–460.
Trajtenberg, Manuel, Rebecca Henderson, and Adam B. Jaffe (1997). "University Versus Corporate Patent: A Window on the Basicness of Invention." 5, 19–50.
Uzzi, Brian (1996). "The Sources and Consequences of Embeddedness for the Economic Performance of Organizations: The Network Effect." *American Sociological Review*, 61, 674–698.
Uzzi, Brian (1997). "Social Structure and Competition in Interfirm Networks: The Paradox of Embeddedness." *Administrative Science Quarterly*, 42, 35–67.

SUPPLEMENTARY DATA

Supplementary data are available at *JEEA* online.