Endogenous Technology Cycles in Dynamic R&D Networks☆

Michael D. König[a], Tim Rogers[b]

[a]*Department of Economics, University of Zurich, Schönberggasse 1, CH 8001 Zurich, Switzerland.*
[b]*Centre for Networks and Collective Behaviour, Department of Mathematical Sciences, University of Bath, BA2 7AY, U.K.*

Abstract

We study the coevolutionary dynamics of knowledge creation and diffusion with the formation of R&D collaboration networks. Differently to previous works, we do not treat knowledge as an abstract scalar variable, but rather represent it as a multidimensional portfolio of technologies. Over time the composition of this portfolio may change due innovations and knowledge spillovers between collaborating firms. The collaborations between firms, in turn, are dynamically adjusted based on the firms' expectations of learning a new technology from their collaboration partners. We show that the interplay between knowledge diffusion, network formation and competition across sectors can give rise to a cyclical pattern in the collaboration intensity, which can be described as a damped oscillation. This theoretical finding recapitulates the novel observation of oscillations in an empirical sample of a large R&D collaboration network over several decades. Finally, we apply our findings to describe how an effective R&D policy can balance subsidies for entrants as well as R&D collaborations between incumbent firms.

Key words: R&D networks, innovation, network formation, technology cycles
JEL: D85, L24, O33

☆We would like to thank Andrea Montanari for numerous comments, guidance and support, and Adrian Etter for the excellent research assistance. Moreover, we would like to thank Paolo Pin, Fernando Vega-Redondo, Jan Eeckhout, Thomas Chaney, Kiminori Matsuyama, Chad Jones, Philippe Aghion, Matt Jackson, Francis Bloch, Armin Schmutzler, Frank Page, Andrei Levchenko, Tim Hellmann, Douglas White, David Dorn, Woody Powell, John Hagedoorn, Michelle Sovinsky, Samuel Kortum, Fabrizio Zilibotti, Pradeep Dubey, Yair Tauman, Hans Gersbach, Maik Schneider and seminar participants at the University of Zurich, Bocconi, Bath University, Stanford University, and the audiences of the 13th SAET Conference on Current Trends in Economics at MINES ParisTech in 2013, the Workshop on Innovations, Patents and Licensing, and the 25th Stony Brook International Conference on Game Theory in 2014 for helpful comments. Michael König acknowledges financial support from Swiss National Science Foundation through research grants PBEZP1–131169 and 100018_140266, and would like to thank SIEPR and the Department of Economics at Stanford University for their kind hospitality during 2010–2012 where parts of this paper have been written. Tim Rogers gratefully acknowledges the support of the Royal Society. The authors further acknowledge the University of Zurich S3IT: Service and Support for Science IT, for providing the support and the computational resources that have contributed to the research results reported in this publication. URL: http://www.s3it.uzh.ch.

Email addresses: `michael.koenig@econ.uzh.ch` (Michael D. König), `T.C.Rogers@bath.ac.uk` (Tim Rogers)

1. Introduction

R&D collaborations play an important role in the creation and diffusion of new technologies, as firms are increasingly relying on external sources of knowledge to develop new products and processes [cf. Powell et al., 2005; Sørensen et al., 2006]. Conversely, new technological opportunities that arise in an evolving technology space impact the formation of R&D collaborations [cf. e.g. Cowan et al., 2007]. In this paper we analyze the two-way influence of innovation, technology diffusion and R&D network formation that is crucial for our understanding of the innovation process. Our analysis reveals (i) a critical threshold in the technology spillover rates between R&D collaborating firms to drive innovation, and (ii) the novel observation of a cyclic dynamic in the formation of R&D collaborations, for which we find empirical evidence using real-world R&D collaboration data.

We develop the first tractable model to study the endogenous coevolution of network formation, knowledge creation and diffusion, in which knowledge is not treated as an abstract scalar variable but considered a diverse portfolio of heterogenous technologies [Cowan and Jonard, 2009]. The technology portfolios change over time through innovation and knowledge spillovers from imitation and learning between collaborating firms. The growth of knowledge is thus an emergent process between innovation and imitation [cf. Jovanovic and Rob, 1989; König et al., 2016]. Conversely, technologies can become obsolete [cf. Adams, 1990; Klette and Kortum, 2004; Sørensen and Stuart, 2000], while R&D collaborations have a finite lifetime, are costly and their profitability is plagued with uncertainty [cf. Fleming, 2001; Harrigan, 1988; Kelly et al., 2002].

Representing the technologies embodied in a firm as a portfolio of different ideas is not only more realistic but has also important implications for the formation of R&D collaborations and the emerging network structure [cf. Baum et al., 2010; Cowan et al., 2007]. First, it allows us to properly account for technological complementarity as an incentive to collaborate. As Jovanovic and Rob [1989] state it "...spillovers of knowledge depend not only on how hard people are trying, but also on the differences in what they know: if all of us know the same thing, we cannot learn from each other". In our model firms only form collaborations when there exists the opportunity for a mutually profitable exchange of technologies. We show that this has important consequences for the characteristics and the stability of the network structure. Second, in the existing models in the literature, which typically treat knowledge as an abstract scalar variable,[1] larger firms typically have lower incentives to form collaborations than smaller firms. However, this contradicts the fact that many collaborations formed, for example, in the biotech sector are between large and small firms [cf. Powell et al., 2005], where the small firm possess knowledge of a technology that is particularly valuable to the larger firm. Here we propose a model in which even large firms have incentives to collaborate with smaller ones when these hold some key technologies of interest.

An important finding of our model is the existence of a *threshold* in the probability of technology spillovers between collaborating firms below which an economy with weak in-house R&D capabilities does not innovate even in the presence of R&D collaborations. This threshold identifies the critical level for spillovers across collaborating firms such that the process of technological diffusion (building up the firms' technology portfolios) in the network dominates the process of knowledge obsolescence (depleting the technology portfolios). The threshold indicates that R&D collaborations can only benefit an economy if firms have developed sufficient "absorptive capacities" to learn and incorporate other firms' technologies [cf. Cohen and Levinthal, 1990; Griffith et al., 2003]. The threshold also indicates that there exists a critical minimal size for innovation clusters (such that spillovers are high enough) to be successful [cf. Duranton et al., 2010].

We further analyze changes of the threshold with respect to various parameters of the model, and, in particular, find that it is lowered by the effect of product competition across sectors. Moreover, we show that the threshold is increasing with the knowledge obsolescence rate [i.e. the "intensity of creative destruction"; see Klette and Kortum, 2004] and the linking cost, while it is decreasing with the productivity of the firms and the alliance duration. Furthermore, we find that this thresh-

[1] Such as, for example, Cabrales et al. [2011], D'Aspremont and Jacquemin [1988], Bloch [1995], Goyal and Joshi [2003], Dawid and Hellmann [2014], Westbrock [2010] and König et al. [2011].

old is a non-monotonic function of the uncertainty (noise) from R&D collaborations [cf. Czarnitzki et al., 2015; Fleming, 2001; Podolny and Page, 1998]. When the noise is large then the threshold is decreasing as lower levels of noise let firms form collaborations that are more profitable. The more profitable collaborations typically involve larger firms with larger technology portfolios, and this leads to more centralized network structures that exhibit better technology diffusion properties [cf. Jackson and Rogers, 2007]. In such networks a technology can diffuse even when the spillover probability is small. In contrast, when the level of noise falls even further, and firms only form the most profitable collaborations, then the selectivity of the firms in terms of their collaboration partners leads to an overall decline in the number of collaborations that are being formed. The resulting network is becoming increasingly sparse, and this weakens its technology diffusion properties. In such networks, a higher technology spillover probability is necessary to guarantee that a technology can diffuse through the economy.

We then characterize the stationary states, and show that the model can endogenously generate power-law distributed productivity levels as well as degrees (number of collaborations), consistent with empirical studies of R&D networks [cf. e.g. Gay and Dousset, 2005; König et al., 2014; Powell et al., 2005]. These highly skewed distributions stem from economies of scale and scope in the propensities of the firms to innovate and the marginal profits firms obtain from forming collaborations.

To test the assumptions and implications of our model, we use a large firm-level panel dataset on R&D collaborations over several decades and various sectors. To motivate the model we provide micro-level evidence illustrating the dynamic interaction between the technology (patent) portfolios of firms and the R&D collaboration network, and we show that neither of them can be studied in isolation. In particular, the existence of technological opportunities through complementary knowledge between firms generates incentives to collaborate, while the existence of collaborations fosters the diffusion of technologies across firms [cf. Jovanovic and MacDonald, 1994; Jovanovic and Rob, 1989].

Moreover, we identify a novel empirical observation, namely, that the R&D collaboration intensity follows a *cyclical* pattern that can be described as a "damped oscillation".[2] A key contribution of this paper is to explain this phenomenon from the existence of endogenous technology cycles, in a tractable framework that is also amenable to policy analysis. In particular, our theoretical analysis indicates that the cyclicality evident in the data is a competition effect. In the early stages after a new technology is discovered, there is a large market for this technology and firms have strong incentives to enter and form collaborations which allow them to get access to the technology. However, once the technology has sufficiently diffused through the network, the market size shrinks, and so do the incentives to collaborate or enter the market, until a new innovation arrives. As a result, the economy experiences periods of high collaborative activity followed by periods of low collaborative activity [cf. Matsuyama, 1999]. Further, the delay stemming from the market adjustment process (when the level of competition and the market size change) gives rise to an over- and undershooting effect in the collaboration intensity compared to the stationary state, and this leads to an oscillation in the collaboration intensity. This has important policy implications. If policy makers want to increase output in the economy by strengthening competition, then our model suggests that a natural side effect is an increased volatility in the network, akin to the Schumpeterian waves of "creative destruction" [cf. Schumpeter, 1934].

After having estimated the parameters of the model, we investigate the effect of industrial policies related to deregulation, subsidies to start-ups and R&D collaborations. A key finding is that when subsidies to R&D collaborations of incumbent firms are too high then they can deter entry, as entrants expect lower profits from stronger competition with more productive incumbents. However, having only few incumbent firms can make the economy more vulnerable to shocks that

[2]This is an oscillation of the average number of collaborations in which the amplitude of the oscillating average is decreasing with time (cf. Figure 2 in Section 2). The oscillations also appear when considering individual sectors separately, or alternative data on R&D collaborations with a different sectoral composition, while the phases of the oscillations differ across sectors. See also Appendix K.4.

force firms to exit, and this can have large negative consequences on the output generated by the economy. We therefore show that an effective R&D policy should balance subsidies for entrants as well as R&D collaborations between incumbent firms.

Relation to the literature. There exists a growing number of empirical studies of R&D networks that document their increasing importance [see e.g. Hagedoorn, 2002]. Only recently, however, has it been recognized that the R&D network structure is highly dynamic. Over time, networks tend to become more dense and increasingly centralized [Hanaki et al., 2010]. Other empirical studies have shown that the propensity to form new alliances by central firms in the network follows a non-monotonic pattern over time [Hagedoorn and van Kranenburg, 2003]. For example, Gulati et al. [2010] find a rise and fall of "small worlds" in the R&D alliance network over time. Various empirical studies have also documented the convergence of firms' knowledge bases in sectors like electronics and the biotechnology industries. In order to explain these phenomenona, in this paper we develop a tractable model in which firms experience decreasing returns from collaborations the more similar their technology portfolios are [cf. Jovanovic and Rob, 1989]. We further show that we can fully replicate the cyclical pattern observed in the data if we also take into account product competition of firms across different sectors [cf. Matsuyama, 1999].

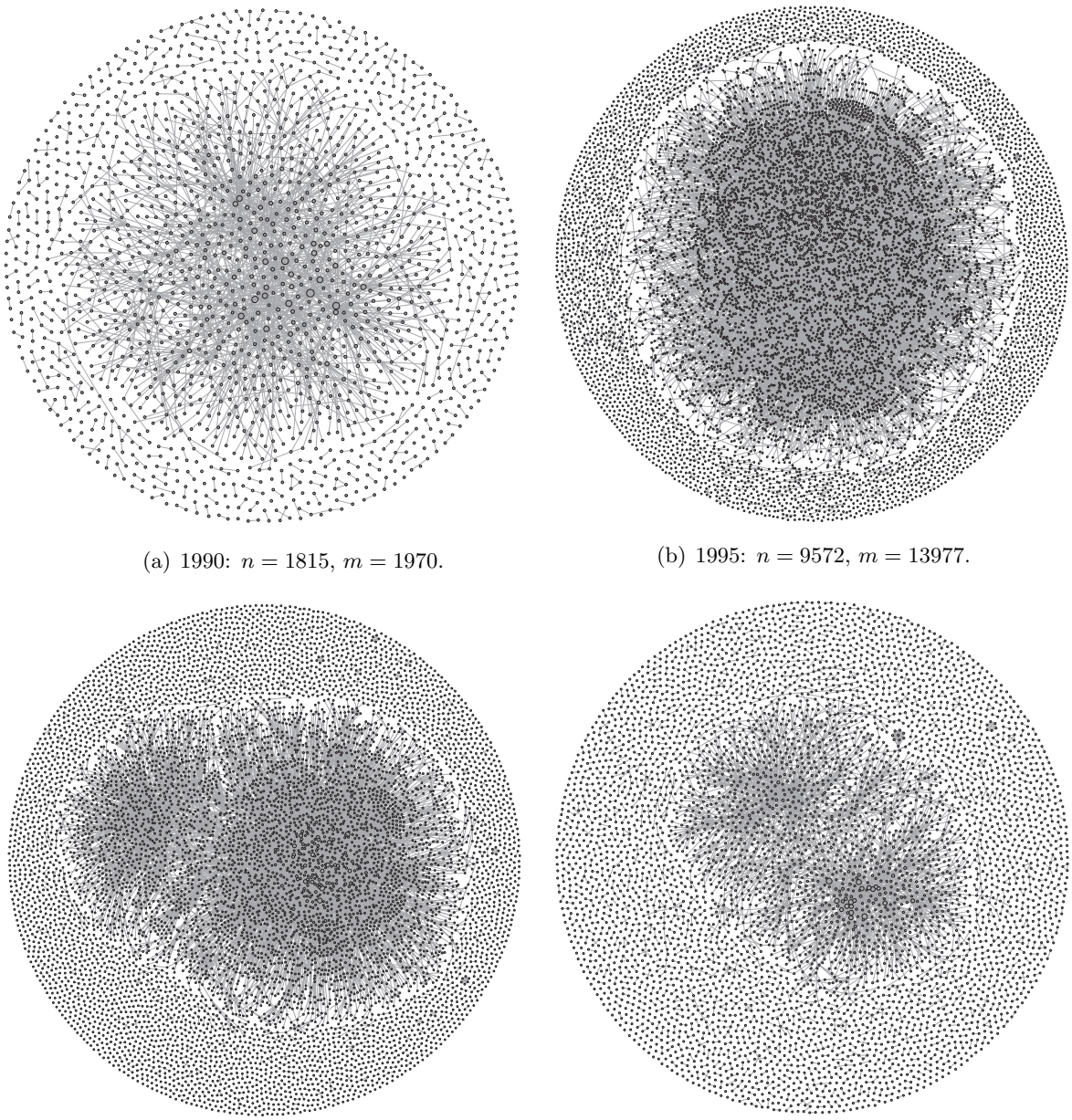
There exists a different strand of literature, seemingly unrelated to R&D networks, in which cyclical patterns of technological change ("innovation waves") have a long history [cf. Schumpeter, 1934]. Cyclical patterns have further been observed empirically in mergers and acquisitions [Golbe and White, 1993], and joint ventures [Gomes-Casseres, 1986]. However, a comprehensive theoretical and empirical study (by showing their existence and providing a theoretical explanation) of cycles in R&D networks is missing so far.

The theoretical analysis of R&D collaborations has attracted some attention in the literature.[1] For example, Goyal and Joshi [2003] and Dawid and Hellmann [2014] have investigated the formation of networks of R&D collaborating firms in which firms can share knowledge about a cost reducing technology. König et al. [2011] study the evolution of R&D networks in which firms form collaborations to maximize their knowledge growth rate through knowledge spillovers from other firms. Relatedly, in Cabrales et al. [2011] agents make costly investments in both, knowledge production and interactions, and there are spillovers between connected agents in the network of interactions. These works, however, abstract from the process of innovation and do not study how such technologies are discovered in the first place. Moreover, in all these works knowledge is treated as an abstract scalar variable instead of a portfolio of different technologies held by a firm. As discussed above, this ignores complementarities in the technology portfolios of collaborating firms, and cannot explain why many collaborations are formed between large and small firms, where the small firm possess knowledge of a key technology that is particularly valuable to the larger firm [cf. Powell et al., 2005].

Another strand of literature has studied the process of knowledge diffusion in an exogenously given (typically communication or social) network. In the mathematics, epidemiology, computer science and physics literature the spread of epidemics on networks has been extensively studied [e.g. Acemoglu et al., 2011].[3] In the economics literature, Morris [2000] provided topological conditions on the network structure under which the adoption of a new technology (in a coordination game played on a fixed network) becomes epidemic. Jackson and Rogers [2007] have analyzed the effect of different, exogenously given network topologies on the spread of innovations and welfare. Similarly, Meagher and Rogers [2004] and Montanari and Saberi [2010] study the process of innovation and knowledge diffusion on an exogenous network structure.[4] These works, however, do not explain the network structure but take it as exogenously given. We depart from this literature, by analyzing the endogenous formation of networks in which innovation and knowledge diffusion takes place.

[3]The model analyzed in this literature is the "susceptible-infective-susceptible" (SIS) model for epidemics spreading on a network [cf. e.g. Jackson and Rogers, 2007]. This model corresponds to the one we study in Section 3.1 for the specific parameter choice of $\alpha = \gamma = 0$ and $N = 1$ with an exogenously given network.

[4]Moreover, empirical evidence for the dependency of new idea creation on the existing stock of ideas through innovation networks is presented in Acemoglu et al. [2016].



(a) 1990: $n = 1815$, $m = 1970$.

(b) 1995: $n = 9572$, $m = 13977$.

(c) 2000: $n = 8258$, $m = 9237$.

(d) 2005: $n = 5487$, $m = 4463$.

Figure 1: Network snapshots for the years (a) 1990, (b) 1995, (c) 2000 and (d) 2005. A node's size indicates its eigenvector centrality (cf. Appendix E).

In particular, for fixed networks, it is the largest eigenvalue of the associated Hashimoto matrix that determines a threshold below which epidemics do not spread [see Karrer et al., 2014; Rogers, 2015]. By contrast here we show that when networks are formed endogenously, this threshold can be reduced to a function of the rates at which neighboring nodes become infected and the infected nodes recover.

There exist only few epidemic spreading models with an endogenously formed network. Notable examples are Gross et al. [2006] and, more recently, Fosco et al. [2010], Blume et al. [2011] and Rogers et al. [2012]. However, these papers do not take into account the incentives of agents to form links. For example, Gross et al. [2006] and Rogers et al. [2012] assume that links are rewired at random. Similarly, in Blume et al. [2011] an agent receives a constant payoff from forming a link, linking decisions are not fully endogenized, and the network is formed according to a specific random process. Fosco et al. [2010] study an endogenously formed network where agents show either good or bad behavior, bad behavior spills over between linked agents, and links involving agents with bad behavior vanish at a higher rate than others. Relatedly, Jackson and Rogers [2007b] analyze

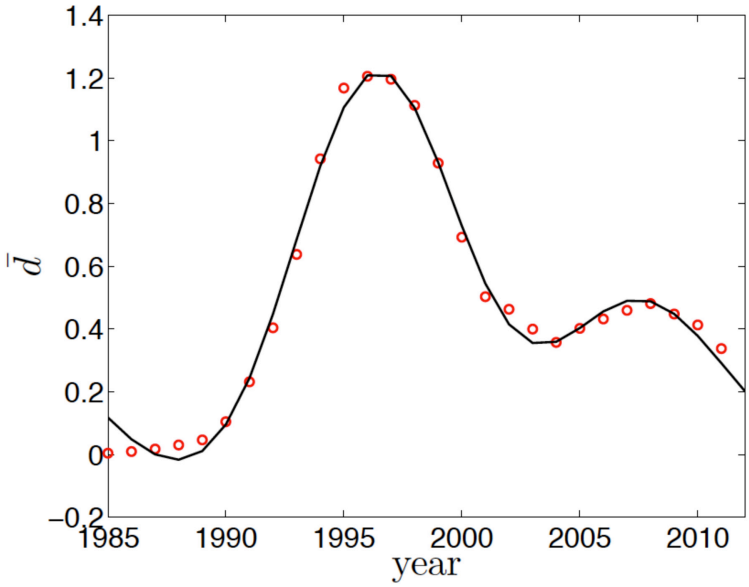


Table: Estimated coefficients with their standard deviations, t-statistics and p-values for the regression model $\bar{d}_t = \frac{1}{2}a_0 + \sum_{j=1}^{3}\left(a_j \cos(\omega_j t) + b_j \sin(\omega_j t)\right)$, with $t \in [0, T]$, the parameters a_0, a_j, b_j, $j = 1, \ldots, 3$ (Fourier coefficients), where the angular frequency is given by $\omega_j \equiv \frac{2\pi}{T} j$, and $\bar{d}_t$ is periodic with period T. See e.g. Hamilton [1994, Chapter 6.2] for further details.

θ	$\hat{\theta}$	$\hat{\sigma}_\theta$	$t_{\hat{\theta}}$	$p_{\hat{\theta}}$
a_0	0.4885***	0.0084	58.4100	0.0000
a_1	−0.4085***	0.0118	−34.5350	0.0000
b_1	−0.0286**	0.0118	−2.4190	0.0247
a_2	0.1648***	0.0118	13.9340	0.0000
b_2	−0.2587***	0.0118	−21.8760	0.0000
a_3	−0.0433***	0.0118	−3.6634	0.0014
b_3	0.0513***	0.0118	4.3366	0.0003

The number of observations is $T = 28$, the error degrees of freedom is 21. The root mean squared error is given by 0.0443. R^2 is 0.989, adjusted R^2 is 0.986. F-statistic vs. constant model is 317, and the p-value is approximately zero.
*** Statistically significant at 1% level.
** Statistically significant at 5% level.
* Statistically significant at 10% level.

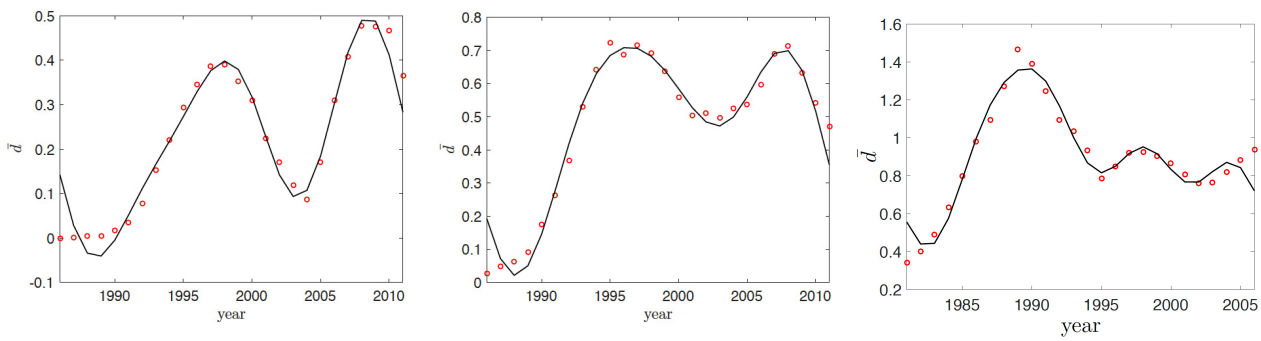


Figure 2: The top left panel shows the average degree $\bar{d}$ over the years 1985 to 2012. The circles indicate the empirical observations while the solid line indicates a Fourier regression fit with the parameter estimates shown in the top right panel. The bottom left and middle panels show $\bar{d}$ for the two largest sectors (computers and drugs development) in the data (cf. Appendix K.4). The bottom right panel shows $\bar{d}$ from the Cooperative Agreements and Technology Indicators (CATI) database [cf. Hagedoorn, 2002] (cf. Appendix K.5).

the evolution of a network of contacts of agents which are discovered randomly over time through already existing contacts. In these models, the link creation process is mechanistic, and does not depend on the marginal payoffs agents receive from forming links. In contrast, in our model link formation is based on a myopic profit maximizing rationale. We further improve on these models by allowing agents in a network to be characterized by an arbitrary number of characteristics instead of a single one. Moreover, none of these papers is applied to the current context of R&D collaborations, nor provides an empirical comparison or policy analysis.

Only a few studies analyze the interplay between knowledge creation, diffusion and network evolution. Most notably, the seminal works by Cowan et al. [2007], Cowan and Jonard [2008] and Baum et al. [2010] have taken into account the existence of ideas in an abstract "technology space" and how collaboration decisions are influenced and are influencing the innovation process. However, these studies are based on numerical simulations and do not provide an analytic framework for the study of innovation and technology diffusion in networks. Moreover, they do not provide an empirical application, and also do not explain the non-monotonic behavior of the collaboration activities of firms over time that we find in the empirical data.

2. Empirical R&D Networks

To motivate our model we consider a panel dataset of R&D alliances ranging over the years 1985 to 2011. The data stems from the Thomson SDC alliance database. Schilling [2009] shows that the Thomson SDC alliance database provides a consistent picture with alternative datasets in terms of the sectoral composition, the alliance activity over time and the geographical origin of the alliance participants. Similar to García-Canal et al. [2008] we take into account three types of alliances

reported in the SDC database: (i) alliances that imply the transmission of an existing technology from one partner to another or to the alliance; (ii) alliances that imply the cross-transfer of existing technologies between two or more partners or between these and the alliance, and (iii) alliances that include the undertaking of R&D activities. This gives us a total of $21,478$ firms across all years in our sample. We construct the R&D alliance network by assuming that an alliance lasts for 5 years similar to e.g. Rosenkopf and Padula [2008].[5,6]

Example networks over years from 1985 to 2010 can be seen in Figure 1, and the average number of collaborations, $\bar{d}_t$, per year t is shown in the top left panel of Figure 2.[7] These figures demonstrate that the evolution of the network is highly non-stationary, and the varying network density indicates a periodic rise and decline of the R&D network structure.

In order to investigate the non-stationary (oscillatory) pattern in the network at the aggregate level, we also perform an estimation procedure similar to Golbe and White [1993]. More precisely, we estimate a Fourier regression model for the average number of collaborations $\bar{d}_t$ in a given year t as indicated in the top panels of Figure 2, showing that all coefficients are statistically significant. Following Golbe and White [1993] we take this as an indicator for the presence of a cyclical pattern in the data.

From the fact that the amplitude of the cycle is decreasing over time we conclude that the average degree follows a "damped oscillation". The existence of cycles is also robust when breaking down the data into individual sectors or using alternative data on R&D collaborations (see the bottom panels of Figure 2 as well as Appendices K.4 and K.5). We also find that there exists a negligible correlation between business cycle movements and the cycle we observe in the R&D collaboration intensity (see also Appendix K.6), ruling out alternative exogenous explanations for the cyclicality in the R&D network data from exogenous business cycle variations. In the next section we develop a minimal model that can generate such a cyclical pattern in the R&D collaboration activities of the firms.

3. The Model

In the following sections we introduce our model for the coevolution of technological innovation, diffusion and the formation of the R&D network (Sections 3.2, 3.3 and 3.4), and we present a theorem which completely describes the evolution of this dynamic system (Section 4). The model has three essential components: (i) the dynamics of knowledge creation and obsolescence, (ii) the formation and removal of R&D collaborations, and (iii) the entry and exit of firms in the market.

3.1. Firms' Profits from Production

The dynamics of the model are driven by the decisions of the firms to form collaborations with others, which in turn are determined by the firms marginal profits from collaborations and their productivities. We will consider two alternative environments in which firms operate: one in which each firm is producing only a single good, and one in which firms can produce multiple goods and compete with each other over supplying these goods. We refer to the first as *single-product*

[5] Rosenkopf and Padula [2008] use a five-year moving window assuming that alliances have a five-year life span, and state that the choice of a five-year window is consistent with extant alliance studies [e.g. Gulati and Gargiulo, 1999; Stuart, 2000] and conforms to Kogut [1988], who finds that the normal life span of most alliances is no more than five years. Moreover, Harrigan [1988] studies 895 alliances over a time span of 60 and concludes that the average life-span of the alliance is relatively short, 3.5 years, with a standard deviation of 5.8 years and 85 % of these alliances last less than 10 years. Park and Russo [1996] focus on 204 joint ventures among firms in the electronic industry for the period 1979–1988. They show that less than half of these firms remain active beyond a period of five years and for those that last less than 10 years (2/3 of the total), the average lifetime turns out to be 3.9 years.

[6] However, this assumption turns out not to be crucial for our results. See also Footnote 7.

[7] Appendix K.3.1 shows the average degree, $\bar{d}_t$, with different alliance durations ranging from 3 to 7 years. Moreover, Appendix K.3.2 shows the average degree, $\bar{d}_t$, obtained when assuming that the duration of an alliances is random (based on the subset of observed alliances with known start and end date). In all cases the oscillatory pattern remains qualitatively unchanged and resembles Figure 2.

monopolies while the latter is referred to as *multi-product competition*.[8]

We consider a a population of n firms labelled by $\mathcal{N} = \{1, \ldots, n\}$. As we explain in detail in Appendix B, in the case of single-product monopolies, a firm i's profit, π_i, is simply proportional to its productivity, A_i [cf. Acemoglu et al., 2006]. In contrast, in the multi-product case the firm's profit, π_i, is proportional to the productivity A_i of the firm times the probability that the firm becomes the supplier of a specific product. We model this probability as the ratio of the productivity of the firm relative to the total productivities of the other firms in the market, $A_i / \sum_{j=1}^{n} A_j$.[9] This is consistent with U.S. census data, where it is found that more productive firms have a larger product range [cf. Bernard et al., 2010]. In both cases (i.e. with or without product competition) we can write profits net of R&D collaboration costs compactly as follows

$$\pi_i = \theta A_i + (1-\theta) A_i n \frac{A_i}{\sum_{j=1}^{n} A_j} - c d_i, \tag{1}$$

where $\theta \in \{0, 1\}$ is a (zero/one) parameter distinguishing between single-product monopolies ($\theta = 1$) and multi-product competition ($\theta = 0$), $c \in \mathbb{R}_+$ is a fixed R&D collaboration cost, d_i is the degree (i.e. the number of links/collaborations) of i in the network $G \in \mathcal{G}(n)$ and $\mathcal{G}(n)$ denotes the set of graphs/networks of size n.[10]

We next assume that the productivity A_i of firm i is a linearly increasing function of the number of technologies (size of the technology portfolio) owned by the firm [cf. Klette and Kortum, 2004].[11] Let $\mathbf{h}_i$ denote the knowledge vector (technology portfolio) of firm i, with $\mathbf{h}_i \in \mathcal{H}^N = \{0, 1\}^N$ and $N \in \mathbb{N}$ denoting the number of different technologies [cf. Baum et al., 2010; Cowan et al., 2007]. Then we assume that

$$A_i = a + b|\,\mathsf{S}(\mathbf{h}_i)|, \quad a, b \in \mathbb{R}_+, \tag{2}$$

where the support of $\mathbf{h}$ is $\mathsf{S}(\mathbf{h})$ and its cardinality is given by $|\,\mathsf{S}(\mathbf{h})| = \langle \mathbf{h}, \mathbf{u} \rangle$, counting the number of nonzero entries in $\mathbf{h}$. Here $\mathbf{u}$ is a vector of ones and $\langle \cdot, \cdot \rangle$ is the usual scalar product in $\mathbb{R}^N$. Normalizing $a = 1$, the profit function $\pi_i : \mathcal{H}^{n \times N} \times \mathcal{G}(n) \to \mathbb{R}$ from Equation (1) then becomes

$$\pi_i(\mathbf{h}, G) = \theta(1 + b|\,\mathsf{S}(\mathbf{h}_i)|) + (1-\theta) \frac{(1 + b|\,\mathsf{S}(\mathbf{h}_i)|)^2}{1 + b\bar{h}} - c d_i, \quad \theta \in \{0, 1\}, \tag{3}$$

with $\mathbf{h} \in \mathcal{H}^{n \times N} = \{0, 1\}^{n \times N}$ denoting the matrix of stacked vectors $\mathbf{h}_i$ for all $i \in \mathcal{N}$ and $\bar{h} \equiv \frac{1}{n} \sum_{j=1}^{n} |\,\mathsf{S}(\mathbf{h}_j)|$ denoting the average technology stock.

In the following we introduce the stochastic process governing innovation, diffusion and R&D network formation, where firms' profit flows are given by Equation (3). An illustration of the possible transition events in this stochastic model is shown in Figure 3. In the single-product monopolies case ($\theta = 1$) the identity of the firm producing a specific product does not change over time. In contrast, in the multi-product competition case ($\theta = 0$), the probability with which a firm becomes the producer of a variety depends on its productivity relative to the productivities of all other firms in the economy, and these productivities change stochastically over time through innovation or learning other firms's technologies in R&D collaborations.[12]

[8] As we will show in the subsequent sections, the single-product monopolies case is closely related to standard Schumpeterian models of monopolistic competition. To generate the fluctuations that we observe in the collaboration intensity in the data (cf. Section 2), however, it will be necessary to consider competition among firms from supplying various intermediate goods, and the interaction (feedback) effect among the firms that this entails, present in the multi-product competition case.

[9] This functional form is also known as a Tullock contest success function [cf. Baye and Hoppe, 2003; Tullock, 1980]. See Appendix B and Footnote 44 therein for further discussion and motivation.

[10] See Appendix E for further definitions and characterizations of networks.

[11] The linearity assumption not only keeps the model tractable, but can also be seen as a linearization (first order Taylor approximation) of a more complicated production technology. The same formulation can be found for example in Klette and Kortum [2004].

[12] The multi-product competition case is closely related to Futia [1980], where a discrete time model in which firms engage in an R&D race for a patent in each period is considered (see also the discussion in Section IV in Reinganum [1985]).

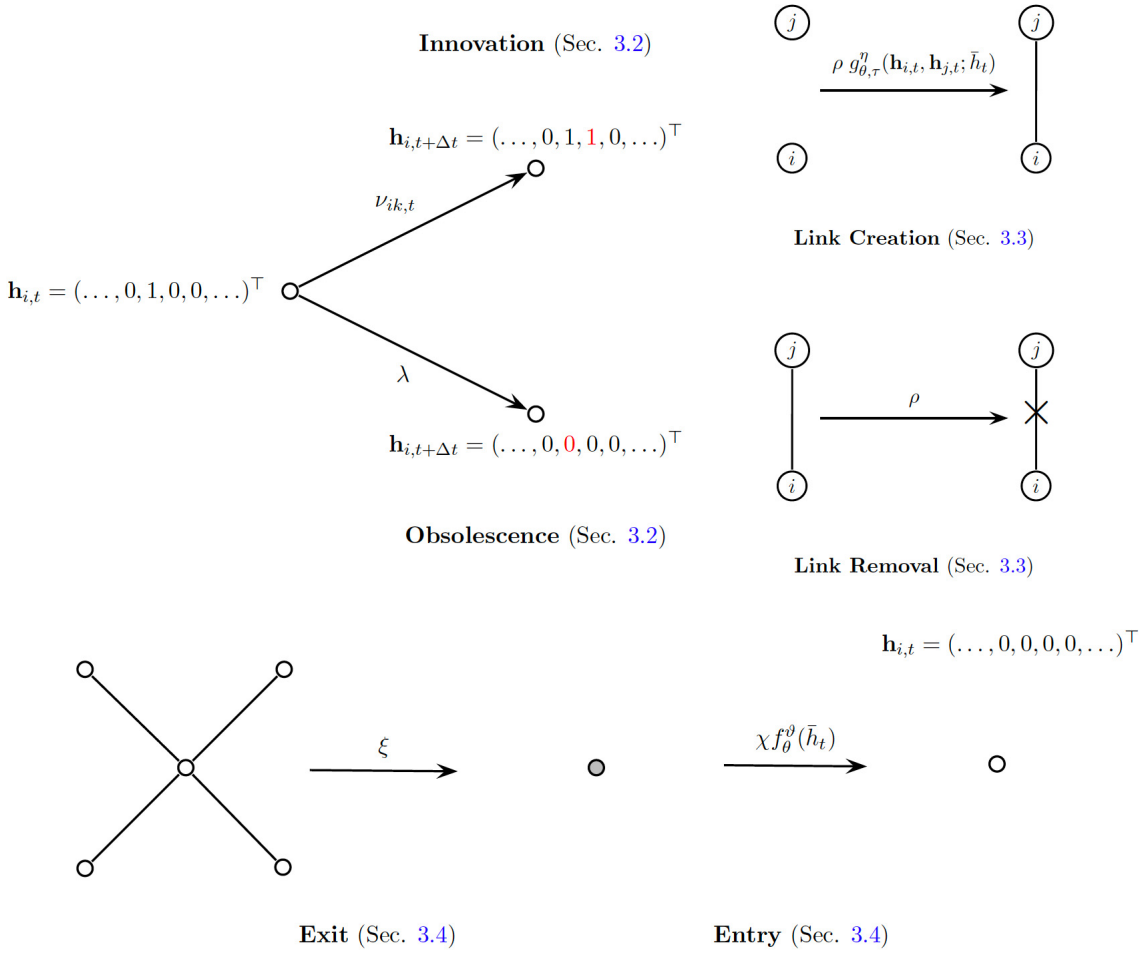


Figure 3: Illustration of the innovation, network formation and entry/exit process of Sections 3.2, 3.3 and 3.4.

3.2. Knowledge Creation, Diffusion and Obsolescence

The knowledge vectors $\mathbf{h}_{it} \in \mathcal{H}^N$ of the firms $i \in \mathcal{N}$ change over continuous time $t \in \mathbb{R}_+$. New ideas arrive as a Poisson process with an innovation rate that depends on the stock of knowledge of the firm [cf. Loury, 1979]. We also allow for spillovers between collaborating firms such that the rate with which a firm i makes an innovation in knowledge category k, with $k = 1, \ldots, N$, increases with the number of collaborating firms that know k [cf. Jackson and Rogers, 2007; Jovanovic and Rob, 1989]. In particular, we assume that a firm i discovers idea k at time t, if it does not know it already, at a rate

$$\nu_{ik,t} = \underbrace{\gamma + \alpha \sum_{l=1}^{N} h_{il,t}}_{\text{innovation}} + \underbrace{\beta \sum_{j=1}^{n} a_{ij,t} h_{jk,t}}_{\text{imitation}}, \tag{4}$$

where $\gamma, \alpha \geq 0$ are parameters related to innovations from in-house R&D, $\beta \geq 0$ is a technology spillover (imitation) parameter, and $a_{ij,t} \in \{0, 1\}$ indicates if there exists a collaboration between firms i and j at time t. The matrix $\mathbf{A}_t$, with components $a_{ij,t}$, $1 \leq i, j, \leq n$, then specifies the connectivity pattern of the collaboration network G_t. Some micro-level empirical evidence for Equation (4) can be found in Appendix A.

We further assume that with rate $\lambda \geq 0$ each knowledge category can also become obsolete.[13] A discussion of the ample literature on knowledge obsolescence can be found in Sørensen and Stuart [2000]. These authors also present empirical evidence from the semiconductor and biotechnology industries that the technologies of firms in changing environments will become obsolete over time. Knowledge obsolescence has also been introduced in Caballero and Jaffe [1993], Adams [1990],

[13] Klette and Kortum [2004] call this rate the "intensity of creative destruction".

Klette and Kortum [2004] and Jovanovic [2009] to describe the technological dynamics among heterogeneous firms.

More formally, each firm $i \in \mathcal{N}$ discovers knowledge category k at the rate $\nu_{ik,t}$ given by Equation (4), and the probability that firm i discovers idea k in the time interval $[t, t+\Delta t)$ is given by $\mathbb{P}(h_{ik,t+\Delta t} = 1 | h_{ik,t} = 0, \mathbf{h}_t, G_t) = 1 - e^{-\nu_{ik,t}\Delta t} = \nu_{ik,t}\Delta t + o(\Delta t)$.[14] In contrast, at rate λ, an idea k of a firm i becomes obsolescent, and the technology k is removed from the portfolio of the firm. That is, the probability that idea k of firm i becomes obsolescent in the time interval $[t, t+\Delta t)$ is given by $\mathbb{P}(h_{ik,t+\Delta t} = 0 | h_{ik,t} = 1, \mathbf{h}_t, G_t) = 1 - e^{-\lambda\Delta t} = \lambda\Delta t + o(\Delta t)$.

3.3. Link Creation and Removal

Opportunities for collaboration arrive as a Poisson process [cf. Sandholm, 2010], similar to Calvo models of pricing [cf. Calvo, 1983]. When a pair of firms receives such an opportunity to change, they decide to form a collaboration so as to maximize profits while taking the collaborations of the other firms as given. That is, we assume that firms are myopic (as e.g. in Watts [2001] and Jackson and Watts [2002]). We further introduce noise in marginal profits from collaborations, as the establishment of an R&D collaboration is fraught with ambiguity and uncertainty [cf. Czarnitzki et al., 2015; Fleming, 2001; Kelly et al., 2002], and many collaborations fail and are terminated early [Podolny and Page, 1998].

We assume that collaborative R&D agreements between firms have only a finite lifetime.[5] Let $\tau > 0$ denote the duration of a collaborative R&D agreement. When evaluating a potential collaboration, a firm computes its expected profit, taking the current network G_t as given [cf. Jackson and Watts, 2002].[15] Capturing the fact that there is a time lag between making an innovation and bringing it to the market [Ken et al., 2008], we further assume that profits accrue at the end of a collaboration when technology exchange has taken place.[16] Then we can derive the expected marginal profits of a firm from forming a collaboration as follows:

Proposition 1. *Denote by $V_i^\tau(\mathbf{h}_t, G_t) \equiv \mathbb{E}_t\left(\pi_i(\mathbf{h}_{t+\tau}, G_t) | \mathbf{h}_t, G_t\right)$ with profits from Equation (3). Then the marginal profit of firm i from forming the collabroation ij can be written as*

$$V_i^\tau(\mathbf{h}_t, G_t + ij) - V_i^\tau(\mathbf{h}_t, G_t) = \beta g_{\theta,\tau}(\bar{h}_t)(1 + b|\,\mathsf{S}(\mathbf{h}_{i,t})|)^{1-\theta}\langle \mathbf{h}_{it}^c, \mathbf{h}_{jt}\rangle - c, \tag{5}$$

where $\theta \in \{0, 1\}$, $\bar{h}_t \equiv \frac{1}{n}\sum_{i=1}^n |\,\mathsf{S}(\mathbf{h}_{i,t})|$ denotes the average stock of knowledge at time t, we have dropped terms of the order $o(\tau)$, and

$$g_{\theta,\tau}(\bar{h}_t) \equiv \tau b\left(\theta + 2\frac{1-\theta}{1+b\bar{h}_t}\right) = \begin{cases} \tau b, & \text{if } \theta = 1, \\ \frac{2\tau b}{1+b\bar{h}_t}, & \text{if } \theta = 0. \end{cases} \tag{6}$$

[14] Given a function $g(x)$ we write $f(x) = o(g(x))$ if and only if $\lim_{x\to 0}\left|\frac{f(x)}{g(x)}\right| = 0$. Further, we say that $f(x) = O(g(x))$ as $x \to 0$ if and only if $\limsup_{x\to 0}\left|\frac{f(x)}{g(x)}\right| < \infty$.

[15] This assumption is reminiscent of the myopic best response dynamics considered in Kandori et al. [1993], and can similarly be found in Hojman and Szeidl [2006], Cowan et al. [2007] and Baum et al. [2010], where agents assume that the network does not change when deciding about their links. The assumption of myopic behavior is common in the complex strategic environment that networks represent [cf. Jackson, 2008]. For example, Jackson and Watts [2002] state that "...in larger networks and networks where players' information might be local and limited, or in networks where players significantly discount the future, myopic behavior is a more natural assumption". Taking into account the strong uncertainty involved in R&D projects and R&D collaborations this assumption seems to be rather intuitive. Related theoretical models of networked agents with limited information sets can be found in DeMarzo et al. [2003], Jackson and Golub [2010] and Golub and Jackson [2012]. The assumption of myopic firms is also common in boundedly rational dynamic decision-making, as considered in Gabaix [2014].

[16] Alternatively, let the firm i's discounted profits over the duration τ of a collaboration be given by $V_i^\tau(\mathbf{h}_t, G_t) \equiv \int_0^\tau e^{-rs}\mathbb{E}_t\left(\pi_i(\mathbf{h}_{t+s}, G_t)|\mathbf{h}_t, G_t\right) ds$ with discount factor $r \geq 0$. Assume that the time τ of a collaboration is short compared to the dynamics of the generation and diffusion of knowledge in the entire industry. Using the Trapezoidal rule (i.e. by making a linear approximation), we can write $2/\tau\; V_i^\tau(\mathbf{h}_t, G_t) = \pi_i(\mathbf{h}_t, G_t) + e^{-r\tau}\mathbb{E}_t\left(\pi_i(\mathbf{h}_{t+\tau}, G_t)|\mathbf{h}_t, G_t\right) + o(1)$. The change in the expected profit of the firm i from forming the link ij is then proportional to the RHS of Equation (5) up to a factor $2/\tau\; e^{-r\tau}$.

From the above expression in Equation (5) we find that marginal profits for firm i from forming a link ij are increasing in the number of ideas that firm j has but firm i does not have, $\langle \mathbf{h}^c_{it}, \mathbf{h}_{jt} \rangle$, and – in the multi-product competitive case ($\theta = 0$) – the total stock of knowledge possessed by firm i, $|\,\mathsf{S}(\mathbf{h}_{i,t})|$, relative to the average stock of knowledge $\bar{h}_t$.[17] Thus, as in Jovanovic and Rob [1989], marginal profits from collaboration incorporate the fact that "...spillovers of knowledge depend not only on how hard people are trying, but also on the differences in what they know: if all of us know the same thing, we cannot learn from each other". The latter case (with multi-product competition, $\theta = 0$) shows that relatively more productive firms are better in reaping the gains from getting access to complementary knowledge than less productive ones (indicating economies of scale and scope). For the remaining sections we will assume that firms evaluate the marginal value of a collaboration on the basis of the change in profits of Equation (5).

Each (unordered) pair of firms $i, j \in \mathcal{N} \times \mathcal{N}$ receives an opportunity to create the link ij with rate $\rho \geq 0$.[18] If the pair i, j receives such an opportunity, then the link ij is created, if it is not present, with probability

$$\begin{aligned} \mathbb{P}(G_{t+\Delta t} &= G_t + ij | \mathbf{h}_t, G_t) \\ &= \rho\, \mathbb{P}(\{V_i^\tau(\mathbf{h}_t, G_t + ij) + \varepsilon_{it} > V_i^\tau(\mathbf{h}_t, G_t)\} \cap \{V_j^\tau(\mathbf{h}_t, G_t + ij) + \varepsilon_{jt} > V_j^\tau(\mathbf{h}_t, G_t)\}) \\ &= \rho\, g^\eta_{\theta,\tau}(\mathbf{h}_{it}, \mathbf{h}_{jt}; \bar{h}_t)\Delta t + o(\Delta t), \end{aligned} \tag{7}$$

where we have denoted by[19]

$$\begin{aligned} g^\eta_{\theta,\tau}(\mathbf{h}_{it}, \mathbf{h}_{jt}; \bar{h}_t) \equiv \Lambda^\eta &\left(\beta g_{\theta,\tau}(\bar{h}_t)(1 + b|\,\mathsf{S}(\mathbf{h}_{i,t})|)^{1-\theta}\langle \mathbf{h}^c_{it}, \mathbf{h}_{jt}\rangle - c\right) \\ &\times \Lambda^\eta \left(\beta g_{\theta,\tau}(\bar{h}_t)(1 + b|\,\mathsf{S}(\mathbf{h}_{j,t})|)^{1-\theta}\langle \mathbf{h}^c_{jt}, \mathbf{h}_{it}\rangle - c\right), \end{aligned} \tag{8}$$

for $\theta \in \{0, 1\}$, $\Lambda^\eta(x) \equiv e^{\eta x}/(1+e^{\eta x})$ is the logistic function, and we assumed that marginal profits in Equation (5) from forming a link are perturbed by identically, independently logistically distributed error terms $\varepsilon_{it}, \varepsilon_{jt}$ with parameter η, and $\bar{h}_t$ denotes the average technology stock (number of technologies) across firms at time t. The error terms, $\varepsilon_{it}, \varepsilon_{jt}$, in Equation (7) capture the inherent uncertainty in R&D collaborations [cf. Czarnitzki et al., 2015; Kelly et al., 2002; Podolny and Page, 1998] in a simple and tractable way, that is also common in the random choice literature [Anderson et al., 1992]. This uncertainty can stem, for example, from imperfectly observable technological capabilities of the firms, or from productivity shocks that affect the profitability of collaborations [cf. Bloom, 2009]. Equation (7) further states that a link between two firms is formed only if both firms find it profitable. Similarly, an existing link ij is removed when the collaboration between i and j expires. This happens at a rate ρ, so that $\mathbb{P}(G_{t+\Delta t} = G_t - ij|\mathbf{h}_t, G_t) = \rho\Delta t + o(\Delta t)$.[20,21]

3.4. Entry and Exit

At a rate $\xi \geq 0$ each firm $i \in \mathcal{N}$ is hit by a shock and exits. This happens independently and uniformly across firms [cf. e.g. Bilbiie et al., 2012].[22] That is, in a small time interval $[t, t + \Delta t)$,

[17] Note that $g_{0,\tau}(\bar{h}_t)$ is decreasing with the average knowledge stock $\bar{h}_t$. In contrast, $g_{1,\tau}(\bar{h}_t)$ is independent of $\bar{h}_t$ by definition.

[18] Observe that when $\rho = 0$ then the network does not change and when also $\alpha = \gamma = 0$ and $N = 1$ we are within the framework of the well known susceptible-infected-susceptible (SIS) model for epidemics spreading on a static network [cf. e.g. Jackson and Rogers, 2007].

[19] Some micro-level empirical evidence for Equation (7) can be found in Appendix A.

[20] Note that we have assumed that the link creation and link removal opportunity arrival rates are identical and given by ρ. This entails no loss of generality but helps us to simplify our notation.

[21] Note also that without link removal the network would eventually become complete. Moreover, without knowledge obsolescence ($\lambda = 0$) the firms' technology portfolios eventually become complete, and there would be no incentives to form collaborations so that the network would be empty. Both of these extreme cases are at odds with the R&D network structures that we observe in the data.

[22] As for example Bilbiie et al. [2012] we assume exogenous exit rates in order to keep our model simple and tractable. Bilbiie et al. [2012] also discuss recent empirical evidence indicating that this assumption is a reasonable starting point

$\Delta t \geq 0$, the probability that firm i exits is proportional to $\xi \Delta t$, or more succinctly: $\mathbb{P}(\text{exit}|\mathbf{h}_t, G_t) = \xi \Delta t + o(\Delta t)$.

Conversely, at a rate $\chi \geq 0$ an exited firm is replaced with a new firm i with an empty technology portfolio and no links, with probability $\mathbb{P}(\text{entry}|\mathbf{h}_t, G_t) = \chi\, \mathbb{P}(\pi_i(\mathbf{0}, G_t) - \kappa > \epsilon_{it}) = \chi f^{\vartheta}_{\theta,\kappa}(\bar{h}_t)\Delta t + o(\Delta t)$, where we have denoted by

$$f^{\vartheta}_{\theta,\kappa}(\bar{h}_t) \equiv \Lambda^{\vartheta}\left(\theta + \frac{1-\theta}{1+b\bar{h}_t} - \kappa\right), \tag{9}$$

for $\theta \in \{0,1\}$, $\Lambda^{\vartheta}(x) \equiv e^{\vartheta x}/(1+e^{\vartheta x})$ is the logistic function, $\bar{h}_t$ denotes the average technology stock across firms at time t, and $\kappa > 0$ is a fixed entry cost. Moreover ϵ_{it} is an identically, logistically distributed error term with parameter ϑ [Anderson et al., 1992], related to the uncertainty of profits in Equation (3) due to market entry. Similar to Section 3.3, the uncertainty from market entry can represent imperfectly observable technological capabilities of the incumbent firms, or productivity shocks that affect the entrant's profitability. In the following sections we will analyze the role of uncertainty in affecting the network structures that emerge and the corresponding technological output generated in the economy.[23]

4. Coevolution of Knowledge Portfolios and the Network

The dynamic process of network formation, innovation and technology diffusion from above can be described by a coupled system of ordinary differential equations. These equations completely determine the evolution of the average number of firms with a certain technology portfolio and the probability of a link between any pair of firms with given technology portfolios over time from any initial condition as the number of firms becomes large. This is shown in the next theorem.

Theorem 1. *Let the probability that a firm with technology vector* $\mathbf{h}$ *is connected to a firm with technology vector* $\mathbf{h}'$ *be denoted by* $y_t(\mathbf{h},\mathbf{h}') \equiv \mathbb{P}(a_{ij,t} = 1|\mathbf{h}_{it} = \mathbf{h}, \mathbf{h}_{jt} = \mathbf{h}')$*, and let the fraction of firms with knowledge vector* $\mathbf{h}$ *be* $x_t(\mathbf{h}) \equiv \mathbb{P}(\mathbf{h}_{it} = \mathbf{h})$*. Then at fixed* $t < T$*, in the limit of large* n*, the time derivate of* $x_t(\mathbf{h})$ *converges in probability with the limit expressed as*

$$\begin{aligned}
\frac{dx_t(\mathbf{h})}{dt} &= (\gamma + \alpha(|\,\mathsf{S}(\mathbf{h})| - 1)) \sum_{k\in\mathsf{S}(\mathbf{h})} x_t(\mathbf{h}-\mathbf{e}_k) - (\lambda|\,\mathsf{S}(\mathbf{h})| + \gamma|\,\mathsf{S}(\mathbf{h}^c)| + \alpha|\,\mathsf{S}(\mathbf{h})||\,\mathsf{S}(\mathbf{h}^c)|)\, x_t(\mathbf{h}) \\
&+ n\beta \sum_{k\in\mathsf{S}(\mathbf{h})} \sum_{\mathbf{h}'\in\mathcal{H}^N : h'_k = 1} y_t(\mathbf{h}-\mathbf{e}_k, \mathbf{h}') x_t(\mathbf{h}-\mathbf{e}_k) x_t(\mathbf{h}') - n\beta \sum_{k\in\mathsf{S}(\mathbf{h}^c)} \sum_{\mathbf{h}'\in\mathcal{H}^N : h'_k = 1} y_t(\mathbf{h},\mathbf{h}') x_t(\mathbf{h}) x_t(\mathbf{h}') \\
&+ \lambda \sum_{k\in\mathsf{S}(\mathbf{h}^c)} x_t(\mathbf{h}+\mathbf{e}_k) - \xi x_t(\mathbf{h}) + \chi f^{\vartheta}_{\theta,\kappa}(\bar{h}_t)\left(1 - \sum_{\mathbf{h}'\in\mathcal{H}^N} x_t(\mathbf{h}')\right) \mathbb{1}_{\{\mathbf{h}=\mathbf{0}\}},
\end{aligned} \tag{10}$$

where $\mathbf{e}_k$ *is the* k*-th unit basis vector in* $\mathcal{H}^N$*,* $\langle\cdot,\cdot\rangle$ *is the usual scalar product in* $\mathbb{R}^N$*,* $\mathsf{S}(\mathbf{h})$ *is the support of* $\mathbf{h}$*,* $f^{\vartheta}_{\theta,\kappa}(\bar{h}_t)$ *has been defined in Equation (9), and the derivative of* $y_t(\mathbf{h},\mathbf{h}')$ *converges in probability with the limit*

$$\frac{dy_t(\mathbf{h},\mathbf{h}')}{dt} = \rho\, g^{\eta}_{\theta,\tau}(\mathbf{h},\mathbf{h}';\bar{h}_t) - \rho\, \left(1 + g^{\eta}_{\theta,\tau}(\mathbf{h},\mathbf{h}';\bar{h}_t)\right) y_t(\mathbf{h},\mathbf{h}') + o(\rho), \tag{11}$$

where $g^{\eta}_{\theta,\tau}(\mathbf{h},\mathbf{h}';\bar{h}_t)$ *has been defined in Equation (8).*

Equation (10) keeps track of the changes in the fraction $x_t(\mathbf{h})$ of firms with technology portfolio

for theoretical analysis that is consistent with the data. In particular, using U.S. census data, Lee and Mukoyama [2008] find that, while plant entry is highly procyclical, annual exit rates are similar across booms and recessions over the business cycle. These authors also find that plants exiting in recessions are very similar in terms of employment or productivity to those exiting in booms.

[23] An overview of the definitions and notations used in this paper can be found in Appendix D.

$\mathbf{h}$. The first term represents contributions to the fraction of firms with technology portfolio $\mathbf{h}$ through in-house R&D by firms with technology portfolios $\mathbf{h}-\mathbf{e}_k$ discovering technology k. This happens at a rate $\gamma+\alpha(|\,\mathsf{S}(\mathbf{h})|-1)$. The second term accounts for a reduction in $x_t(\mathbf{h})$ if a firm with technology portfolio $\mathbf{h}$ discovers a new technology (that is, $h_k=0$) through in-house R&D at a rate $\gamma+\alpha|\,\mathsf{S}(\mathbf{h})|$, and there are $|\,\mathsf{S}(\mathbf{h}^c)|$ such new technologies to discover, or through obsolescence at the rate λ. The third and the fourth terms stem from contributions due to technology spillovers. Firms with technology portfolio $\mathbf{h}-\mathbf{e}_k$ can learn technology k from firms with technology portfolio $\mathbf{h}'$ when $h'_k=1$ at a rate $n\beta y_t(\mathbf{h}-\mathbf{e}_k,\mathbf{h}')x_t(\mathbf{h}-\mathbf{e}_k)x_t(\mathbf{h}')$. Similarly, $x_t(\mathbf{h})$ can decrease if a firm with technology portfolio $\mathbf{h}$ discovers a new technology (that is, $h_k=0$) through learning from collaborating firms with technology portfolio $\mathbf{h}'$ for all $\mathbf{h}'\in\mathcal{H}^N$ such that $h'_k=1$ at the rate $n\beta y_t(\mathbf{h},\mathbf{h}')x_t(\mathbf{h})x_t(\mathbf{h}')$. The fifth term captures an increase in $x_t(\mathbf{h})$ through obsolescence of firms with technology portfolios $\mathbf{h}+\mathbf{e}_k$. The last two terms capture contributions from entry and exit. Due to exit, $x_t(\mathbf{h})$ decreases at a rate ξ. Conversely, $x_t(\mathbf{0})$ grows at a rate $\chi f^{\vartheta}_{\theta,\kappa}(\bar{h}_t)(1-\sum_{\mathbf{h}\in\mathcal{H}^N}x_t(\mathbf{h}))$, where $1-\sum_{\mathbf{h}\in\mathcal{H}^N}x_t(\mathbf{h})$ counts the fraction of firms that have exited, and $\sum_{\mathbf{h}\in\mathcal{H}^N}x_t(\mathbf{h})$ is the fraction of firms that are active. In a similar way Equation (11) can be derived.[24]

With the dynamics of $x_t(\mathbf{h})$ and $y_t(\mathbf{h},\mathbf{h}')$ known from Theorem 1, the average stock of knowledge is given by $\bar{h}_t=\sum_{\mathbf{h}\in\mathcal{H}^N}|\,\mathsf{S}(\mathbf{h})|x_t(\mathbf{h})$, and the average degree is $\bar{d}_t=n\sum_{\mathbf{h},\mathbf{h}'\in\mathcal{H}^N}y_t(\mathbf{h},\mathbf{h}')x_t(\mathbf{h})x_t(\mathbf{h}')$.[25] Since the average stock of knowledge $\bar{h}_t$ is completely determined by $x_t(\mathbf{h})$ we will write $g_{\theta,\tau}(\bar{h}_t)$ as $g_{\theta,\tau}(x_t)$.

When we do not make the assumption that ρ is large, then we need to take into account the remainder term which is of the order of $o(\rho)$ in Equation (11). The differential equations governing the dynamics of the expected number of links becomes considerably more involved, and can be derived by making a *pair approximation*. Let $n_t(\mathbf{h})$ denote the expected number of firms with technology $\mathbf{h}$, $m_t(\mathbf{h},\mathbf{h}')$ the expected number of links between firms with technologies $\mathbf{h}$ and $\mathbf{h}'$. Moreover, let $\Delta_t(\mathbf{h},\mathbf{h}',\mathbf{h}'')$ denote the expected number of triplets with a firm with technology $\mathbf{h}$ being connected to a firm with technology $\mathbf{h}'$ and this firm being connected to a firm with technology $\mathbf{h}''$. Then the pair approximation is given by

$$\Delta_t(\mathbf{h},\mathbf{h}',\mathbf{h}'')\approx\frac{m_t(\mathbf{h},\mathbf{h}')m_t(\mathbf{h}',\mathbf{h}'')}{n_t(\mathbf{h}')}. \tag{12}$$

Appendix J provides a complete derivation of the dynamics for an arbitrary number N of technology categories using the approximation in Equation (12).[26] However, in the next sections we will be mostly concerned with the exact case in the limit of $\rho\to\infty$, as this also simplifies our analysis considerably.

5. Equilibrium Characterization

In this section we study the stationary states of the stochastic process introduced in Sections 3.2, 3.3 and 3.4 by computing the fixed points of the ODEs of Equations (10) and (11), respectively. We first analyze the case of single product monopolies in Section 5.1, while the more involved case of multi-product competition is investigated in Section 5.2. For the latter, to aid tractability, we also constrain the number of technologies to at most two.

The complexity of the dynamical system in Theorem 1 may be reduced drastically by observing the symmetry between the components in the knowledge portfolio (exchangeability) of a firm. In the following we consider the dynamics of the stocks of knowledge and the probability of a link between firms with given knowledge stocks. For this purpose, let the fraction of firms with a stock

[24] See the proof of Theorem 1 in Appendix C for a detailed derivation.

[25] We can also derive an ODE governing the fraction of exited firms, r_t, which is given by $\frac{dr_t}{dt}=-\chi f^{\vartheta}_{\theta,\kappa}(\bar{h}_t)r_t+\xi\sum_{\mathbf{h}\in\mathcal{H}^N}x_t(\mathbf{h})$.

[26] Numerical simulations of the ODEs in the case of the large ρ approximation and the pair approximation indicate that they yield similar results even for moderate values of ρ. See in particular Figure J.3 in Appendix J.

of knowledge of s, $0 \leq s \leq N$, be given by $\binom{N}{s}\bar{x}_t(s)$, where we have denoted by

$$\bar{x}_t(s) \equiv \frac{1}{\binom{N}{s}} \sum_{\mathbf{h}\in\mathcal{H}^N:|\,\mathsf{S}(\mathbf{h})|=s} x_t(\mathbf{h}), \tag{13}$$

with $x_t(\mathbf{h})$ being the solution of Equation (10). Further, let the probability of a link between a firm with knowledge stock s and a firm with s', with $0 \leq s, s' \leq N$, be given by $\binom{N}{s}\binom{N}{s'}\bar{y}_t(s,s')$, where we have introduced

$$\bar{y}_t(s,s') \equiv \frac{1}{\binom{N}{s}}\frac{1}{\binom{N}{s'}} \sum_{\substack{\mathbf{h}\in\mathcal{H}^N:|\,\mathsf{S}(\mathbf{h})|=s \\ \mathbf{h}'\in\mathcal{H}^N:|\,\mathsf{S}(\mathbf{h}')|=s'}} y_t(\mathbf{h},\mathbf{h}'), \tag{14}$$

and $y_t(\mathbf{h},\mathbf{h}')$ being the solution of Equation (11). Furthermore, define the symmetric matrix, $\bar{g}^{\eta}_{\theta,\tau}(s,s';\bar{h}_t) = \bar{g}^{\eta}_{\theta,\tau}(s',s;\bar{h}_t)$ for all $0 \leq s, s' \leq N$, given by

$$\bar{g}^{\eta}_{\theta,\tau}(s,s';\bar{h}_t) \equiv \frac{1}{\binom{N}{s}}\frac{1}{\binom{N}{s'}} \sum_{\substack{\mathbf{h}\in\mathcal{H}^N:|\,\mathsf{S}(\mathbf{h})|=s \\ \mathbf{h}'\in\mathcal{H}^N:|\,\mathsf{S}(\mathbf{h}')|=s'}} g^{\eta}_{\theta,\tau}(\mathbf{h},\mathbf{h}';\bar{h}_t), \tag{15}$$

with $g^{\eta}_{\theta,\tau}(\mathbf{h},\mathbf{h}';\bar{h}_t)$ as in Equation (8).[27] The average stock of knowledge is then given by $\bar{h}_t = \sum_{s=1}^{N} s\binom{N}{s}\bar{x}_t(s)$, and the average degree is $\bar{d}_t = n\sum_{s,s'=0}^{N}\binom{N}{s}\binom{N}{s'}\bar{x}_t(s)\bar{x}_t(s')\bar{y}_t(s,s')$. Then we observe that, while for the dynamics of the technology portfolios, $\mathbf{h} \in \mathcal{H}^N$, in Theorem 1 we needed to solve $2^N + 2^{2N}$ equations for $x_t(\mathbf{h})$ and $y_t(\mathbf{h},\mathbf{h}')$, for the dynamics of the knowledge stocks, $s = 0,\ldots,N$ we need to solve only $N + (N+1)^2$ equations for $\bar{x}_t(s)$ and $\bar{y}_t(s,s')$.[28]

5.1. Single-Product Monopolies

The general characterization of the equilibrium solution for an arbitrary parameter choice and arbitrary N is rather involved. However, further insights can be obtained by restricting our analysis to the case of independent markets ($\theta = 1$), when $\bar{g}^{\eta}_{1,\tau}(s,s')$ does not depend on $\bar{x}(s)$, respectively $\bar{h}_t$ (cf. Equation (8)), the absence of entry and exit ($\xi = \chi = 0$) and letting $\rho \to \infty$ (fast network adjustment) such that terms of the order $o(\rho)$ can be neglected in Equation (11). Under these assumptions, we can show that R&D collaborations indeed lead to higher average knowledge stocks, however, only when the R&D spillovers exceed a certain threshold. Moreover, both, the knowledge stocks and the degrees are power-law distributed, which is consistent with the data [cf. e.g. Gay and Dousset, 2005; König et al., 2014; Powell et al., 2005]. However, we find that the absence of product competition and the feedback mechanism on the firms' technology stocks this entails, fails to generate the oscillatory dynamics that we observe in the data (cf. Section 2).

The first part of the following proposition provides an explicit solution to the (conditional) linking probability of firms with different knowledge stocks, and a recursive characterization of the asymptotic stocks of knowledge. Moreover, we find that when there is no in-house R&D a stationary solution is always given by $\bar{x}(s) = \delta_{s,0}$,[29] where firms have empty technology portfolios and thus vanishing stocks of knowledge. However, this trivial stationary state is not the only stationary state if β exceeds a threshold β^c. Moreover, the trivial stationary state becomes unstable if β is higher

[27] We can compute Equation (15) more explicitly as follows

$$\bar{g}^{\eta}_{\theta,\tau}(s,s';\bar{h}_t) = \frac{1}{\binom{N}{s}}\frac{1}{\binom{N}{s'}} \sum_{k=\max\{0,s'-s\}}^{\min\{N-s,s'\}} \binom{N-s}{k}\binom{s}{s'-k}\binom{N}{N-s}$$
$$\times \Lambda^{\eta}\left(\beta g_{\theta,\tau}(x_t)(1+bs)^{1-\theta}k - c\right) \cdot \Lambda^{\eta}\left(\beta g_{\theta,\tau}(x_t)(1+bs')^{1-\theta}(s-s'+k) - c\right),$$

where $\Lambda^{\eta} : x \mapsto e^{\eta x}/(1+e^{\eta x})$ is the logistic function.

[28] See also Lemma 1 and Figure C.1 in Appendix C.

[29] The Kronecker delta function is defined as $\delta_{ij} = 0$ if $i \neq j$ and $\delta_{ij} = 1$ otherwise.

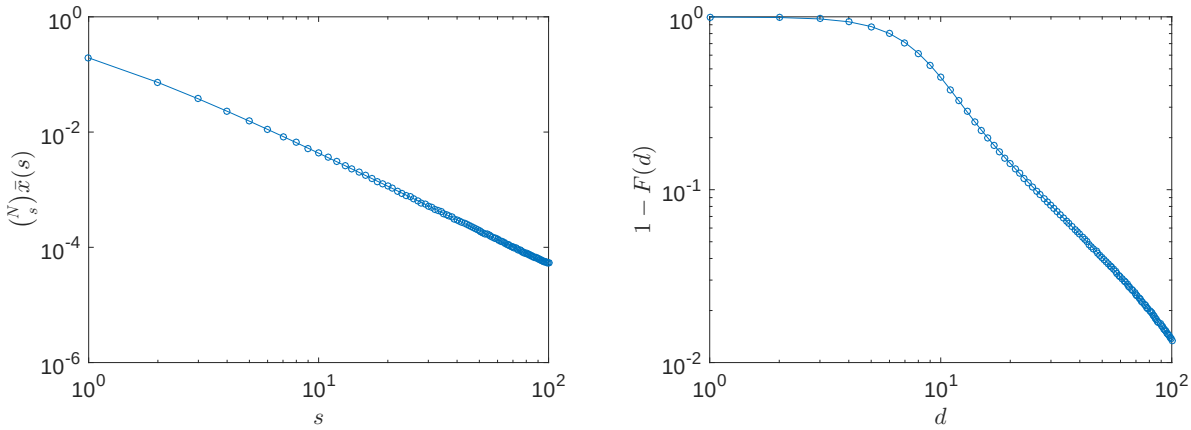


Figure 4: An illustration of the power-law distributed knowledge stocks (left panel) and the degrees (right panel) from numerical solution of Equations (10) and (11) for large N, weak spillovers (β) and no obsolescence ($\lambda = 0$), where one can show that the knowledge stocks distribution decays as a power-law with $\binom{N}{s}\bar{x}(s) \sim s^{-\left(1-\frac{\xi}{\alpha/N}\right)}$ as $s \to \infty$, while the degree distribution is a Mixed Poisson distribution with a power-law tail with the same exponent (see also Appendix F.2.

than β^c. This is shown in the second part of the proposition. The third and fourth parts of the proposition characterize the stationary state in the limit of weak spillovers from collaborations, and show that the knowledge stocks are distributed as a power-law, while their solution trajectories do not show any oscillatory behavior.

Proposition 2. *Consider the limit of large ρ such that terms of the order $o(\rho)$ can be neglected in Equation (11), assume $\theta = 1$ (single-product monopolies) and let $\xi = \chi = 0$ (no entry or exit).*

(i) The fixed points of Equation (11) with $\bar{y}(s, s')$ defined in Equation (14) are given by

$$\bar{y}(s, s') = \frac{\bar{g}^{\eta}_{1,\tau}(s, s')}{1 + \bar{g}^{\eta}_{1,\tau}(s, s')}, \tag{16}$$

while the fixed points of Equation (10) with $\bar{x}(s)$ defined in Equation (13) satisfy

$$\bar{x}(s+1) = \frac{\bar{x}(0)}{\lambda^{s+1}} \prod_{k=0}^{s} \left(\gamma + k\alpha + n\beta \sum_{s'=1}^{N} \binom{N-1}{s'-1} \frac{\bar{g}^{\eta}_{1,\tau}(k, s')}{1 + \bar{g}^{\eta}_{1,\tau}(k, s')} \bar{x}(s') \right),$$

for $0 \leq s \leq N-1$ and $\bar{x}(0) = 1 - \sum_{s=1}^{N} \binom{N}{s} \bar{x}(s)$, where $\bar{g}^{\eta}_{1,\tau}(s, s')$ has been defined in Equation (15).

(ii) Assume that $\gamma, \alpha \to 0$ (no in-house R&D). Then the unique, asymptotically stable[30] *stationary state is $\bar{x}(s) = \delta_{s,0}$ if $\beta < \beta^c$, with*

$$\beta^c = \frac{\lambda\left(2 + e^{c\eta}\right)}{n} + \frac{1}{b\eta\tau} W\left(\frac{b\eta\lambda\tau}{n} \left(1 + e^{c\eta}\right) e^{\frac{\eta\left(cn - b\lambda\tau\left(2+e^{c\eta}\right)\right)}{n}} \right), \tag{17}$$

where $W(x)$ is the Lambert W function (or product-log), which is implicitly defined by $W(x)e^{W(x)} = x$.

[30] Let $f : \mathbb{R}^n \to \mathbb{R}^n$ be locally Lipschitz, and consider the autonomous system $\frac{d\mathbf{x}}{dt} = f(\mathbf{x})$. Then $\mathbf{x} = \mathbf{0}$ is a stable fixed point if, for each $\epsilon > 0$, there exists a $\delta > 0$ such that $\|\mathbf{x}(\mathbf{0})\| < \delta$ implies that $\|\mathbf{x}(t)\| < \epsilon$, for all $t \geq 0$. The fixed point $\mathbf{x} = \mathbf{0}$ is asymptotically stable, if δ can be chosen such that if $\|\mathbf{x}(\mathbf{0})\| < \delta$ then $\lim_{t\to\infty} \mathbf{x}(t) = \mathbf{0}$. Further, define the Jacobian as $\mathbf{J}(\mathbf{x}) = \nabla f(\mathbf{x})$, then $\mathbf{x} = \mathbf{0}$ is asymptotically stable if all eigenvalues of $\mathbf{J}(\mathbf{0})$ are (strictly) negative [cf. Khalil, 2002, Theorem 4.7].

(iii) In the case of $\beta = 0$ (no spillovers) we have

$$\bar{x}(s) = \left(\sum_{s'=0}^{N} \binom{N}{s'} \prod_{k=0}^{s'-1} \frac{\gamma + \alpha k}{\lambda} \right)^{-1} \prod_{k=0}^{s-1} \frac{\gamma + \alpha k}{\lambda}, \quad s = 0, \ldots, N. \tag{18}$$

In particular, when $\alpha > 0$, in the limit of large N, the distribution of knowledge stocks decays as a power-law with $\binom{N}{s}\bar{x}(s) \sim \left(\frac{\alpha}{\lambda}\right)^{s-1} s^{-\left(1-\frac{\gamma}{\alpha}\right)}$ as $s \to \infty$, while when $\alpha = 0$ the knowledge stocks are Binom$\left(N, \frac{\gamma}{\lambda+\gamma}\right)$*–distributed.*

(iv) For $\beta/\gamma \to 0$ (weak spillovers), the stationary stocks of knowledge are distributed as

$$\bar{x}(s) = \bar{x}(s)|_{\beta=0} + \frac{n\beta}{\gamma} \frac{\Omega}{\Psi^2} \left(\frac{\omega(s)}{\Omega} - \bar{x}(s)|_{\beta=0} \right) + O\left(\frac{\beta}{\gamma}\right)^2, \quad s = 0, \ldots, N, \tag{19}$$

where $\bar{x}(s)|_{\beta=0}$ is given by Equation (18), a function $\omega : \{0, \ldots, N\} \to \mathbb{R}_+$ and constants Ω, Ψ whose expressions can be found in the proof of the proposition. Moreover, the eigenvalues of the Jacobian $\mathbf{J}$ corresponding to the dynamical system in Equations (10) and (11) are all real, and consequently their solution trajectories do not show any oscillatory behavior.

The first part of Proposition 2 shows that $\bar{x}(s)$ is higher for all $s > 0$ when $\beta > 0$. Hence, as expected, the presence of R&D collaborations leads to a higher average knowledge stock in the economy. It also shows that the knowledge stocks are increasing (not necessarily strictly) in β.

The second part of Proposition 2 illustrates that if the in-house R&D capabilities of firms are weak ($\gamma, \alpha \to 0$), then the presence of R&D collaborations can only lead to an economy with a significant fraction of firms having non-empty technology portfolios, $\bar{x}(0) < 1$, if the spillover parameter β exceeds a *threshold* β^c. Proposition 2 further states that when $\beta > \beta^c$ then the trivial stationary state becomes unstable and there exists a non-trivial stable stationary state. Moreover, one can show that the threshold β^c satisfies the following inequalities:[31]

$$\frac{\partial \beta^c}{\partial \lambda} > 0, \quad \frac{\partial \beta^c}{\partial c} > 0, \quad \frac{\partial \beta^c}{\partial b} < 0, \quad \frac{\partial \beta^c}{\partial \tau} < 0, \tag{20}$$

while β^c is a convex function of η as indicated in Figure 5 for different values of the linking cost c. The reason for this non-monotonicity is the following. When the noise is large (and η is small) then the threshold is decreasing with increasing values of η as lower levels of noise let firms form collaborations that are more profitable. The more profitable collaborations typically involve larger firms with larger technology portfolios, and this leads to more centralized networks that exhibit better technology diffusion properties [cf. Jackson and Rogers, 2007]. In such networks a technology can diffuse even when the spillover probability is small. In contrast, when the level of noise falls even further, and firms form only the most profitable links, then the selectivity of the firms in terms of their collaboration partners leads to an overall decline in the number of collaborations that are being formed. The resulting network is becoming increasingly sparse, and this weakens its technology diffusion properties. In such networks, a higher spillover probability is necessary to guarantee that a technology can diffuse through the economy.[32,33]

[31] See the proof of Proposition 2 in Appendix C.

[32] Additional comparative statics results with respect to the average technology stock, $\bar{h}$, can be found in Appendix F.1. There it is shown that the knowledge gains from R&D collaborations are higher in the presence of product competition across sectors, and are increasing with increasing in-house R&D capabilities, but only if these are below a threshold that depends on the knowledge obsolescence rate, λ, and decreasing otherwise.

[33] In Appendix H we also investigate the efficient (i.e. output maximizing) network structure, and compare it to the decentralized equilibrium. Our analysis indicates that equilibrium networks tend to be less centralized than the efficient structure. Moreover, we find that efficient networks have a core-periphery structure, and can be characterized as a "nested split graph" [cf. Belhaj et al., 2016; König et al., 2011]. A network is a nested split graph if the neighborhood of every node is contained in the neighborhoods of the nodes with higher degrees (see also Appendix E). Our analysis

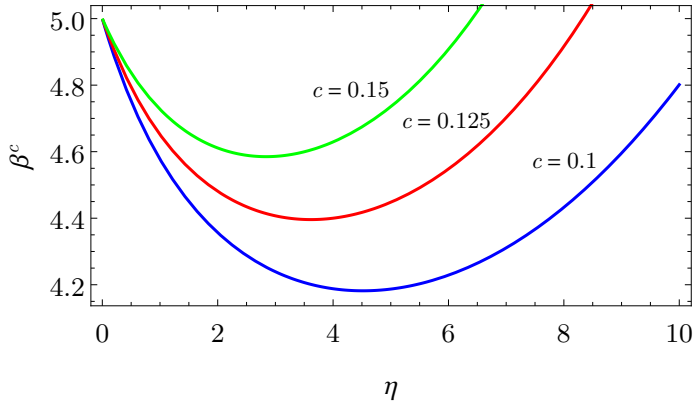


Figure 5: The threshold level β as a function of η for $b = 10$, $\tau = 0.01$, $\lambda = 1$, and $c \in \{0.1, 0.125, 0.15\}$.

The existence of a threshold β^c for the learning success probability between collaborating firms below which an economy with weak in-house R&D capabilities does not innovate even in the presence of R&D collaborations is a key finding of our model. In particular, it indicates that R&D collaborations can only benefit an economy if firms have developed sufficient absorptive capacities to learn and incorporate other firms' technologies [cf. Cohen and Levinthal, 1990; Griffith et al., 2003]. Moreover, the existence of a threshold also indicates that there exists a critical minimal size for innovation clusters (such that spillovers are high enough) to be successful [cf. Duranton et al., 2010; Saxenian, 1994].

An explicit solution for the technology stocks is provided in the third part of Proposition 2 for vanishing spillover effects ($\beta = 0$). There it is also shown that when N becomes large, the knowledge stocks decay as a *power-law*. This implies that the firms' productivities are power-law distributed [König et al., 2016]. The result also holds for more general parameter configurations, where we also find that the degree distribution decays as a power-law.[34] An illustration from a numerical solution of Equations (10) and (11) can be seen in Figure 4. The heavily skewed distributions originate from economies of scale/scope in the propensities of the firms to innovate (cf. Equation (4)) and the marginal profits from forming collaborations (cf. Equation (5)). Thus, our model is able to explain the joint occurrence of power-law distributed productivities and degrees, as one can find them in empirical studies of R&D networks [cf. e.g. König et al., 2014; Powell et al., 2005].

The last part of Proposition 2 shows that at least when spillover effects are not too strong, an economy without product competition cannot generate the *cyclical* dynamics of the average degree that we have documented in Section 2. However, as we will show in the following section, such oscillations can be obtained in a competitive environment.

5.2. Multi-Product Competition

In order to obtain more general results beyond the case of small spillover effects, in the following we analyze the case of at most two competing technologies ($N \in \{1, 2\}$),[35] considering both, single-product monopolies ($\theta = 1$) and a multi-product competitive environment ($\theta = 0$). A detailed derivation and analytic characterization can be found in Appendices G.1 and G.2, respectively, whereas here we describe the main results that we obtain.

further allows us to investigate the impact of product competition on output. We find that competition is output increasing as long as it does not deter entry. This is due to the fact that competition leads to reallocation and the replacement of less productive firms with more productive ones, which are characterized by a more central network position. Our paper thus provides a novel contribution to the discussion of whether competition has a conducive or detrimental effect on innovation [cf. Kretschmer et al., 2012], by focussing on R&D intensive industries where R&D collaborations between firms are commonly observed. Further, we show that both higher innovation rates, as well as increased technology spillovers, give rise to higher inequality; a fact consistent with the recent innovation literature [Jovanovic, 2009].

[34] See also Appendix F.2.

[35] This simplifying restriction is shared with various other works in a similar context such as Montanari and Saberi [2010], Young [2002], and Jovanovic and MacDonald [1994].

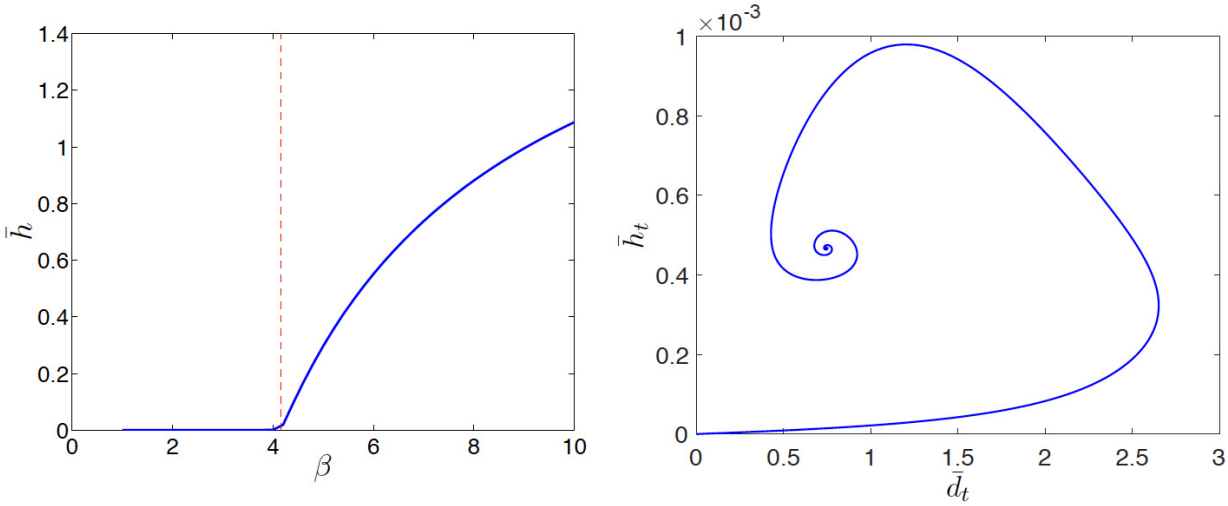


Figure 6: (Left panel) The asymptotic average stock of knowledge $\bar{h} = \lim_{t\to\infty} \bar{h}_t = \lim_{t\to\infty} 2(\bar{x}_t(1) + \bar{x}_t(2))$ for $N = 2$, $\eta = 0$, $\gamma = 0.01$, $b = 1$, $\rho = 10$, $n = 500$, $\lambda = 500$, $\xi = \chi = 0$ as a function of β with the threshold β^c indicated with a dashed line solving Equation (59). (Right panel) Trajectories of the average stock of knowledge $\bar{h}_t$, and the average degree $\bar{d}_t$ in the competitive case where $\theta = 0$ for $N = 2$, $\alpha = 0$, $\beta = 0.16$, $b = 1250$, $c = 1.3$, $\gamma = 0.0001$, $\eta = 1.5$, $\lambda = 2.1$, $n = 500$, $\rho = 30$, $\xi = 0.45$, $\chi = 1$, $\vartheta = 20$, $\kappa = 0.7$ and $\tau = 0.01$. The initial condition is $x_0(0) = 1$, $\bar{x}_0(1) = 0$, $\bar{x}_0(2) = 0$ and an empty network.

As in the previous section we can identify a threshold, β^c, that needs to be exceeded by the spillover probability β such that significant innovation activities can be observed in an industry with weak in-house R&D capabilities. The left panel in Figure 6 shows the asymptotic average stock of knowledge, $\lim_{t\to\infty} \bar{h}_t$, as a function of β with the threshold β^c indicated with a dashed line.[36] However, while we find that cyclical behavior did not occur when we assumed single-product monopolies ($\theta = 1$), we observe cyclical patterns in our numerical integration of the governing differential equations when multi-product competition was introduced ($\theta = 0$), as for example in the right panel of Figure 6. This indicates that competition across multiple products increases the variability of the network density over time, and is a necessary ingredient to explain the oscillatory pattern with have observed in Section 2. Moreover, the oscillations also crucially depend on the interplay between network formation and market adjustment through entry and exit. A phase diagram showing the region in the (χ, β)-parameter space where oscillations can be observed is shown in Figure 7. The figure indicates that only the combined effects of entry/exit $(\chi, \xi > 0)$ and spillovers $(\beta > 0)$ can give rise to oscillations. An intuitive explanation for this phenomenon can be given as follows. When a new technology is discovered, firms have strong incentives to enter the market and form collaborations which allow them to get access to the technology. However, once the technology has sufficiently diffused through the network, the average firm becomes more competitive, the market size shrinks, and so do the incentives for market entry and the formation of collaborations. The economy then experiences periods of high collaborative activity when competition is weak and there are large technological opportunities, followed by periods of low collaborative activity when competition is high and there are only few technological opportunities. The delay in the adjustment of the market through entry and exit to changes in the average technology stocks further leads to an over- and undershooting effect in the average number of collaborations as compared to the stationary state.

[36]The threshold β^c can be found in Equation (51) in Appendix G.1 for $N = 1$ and Equation (59) in Appendix G.2 for $N = 2$, respectively.

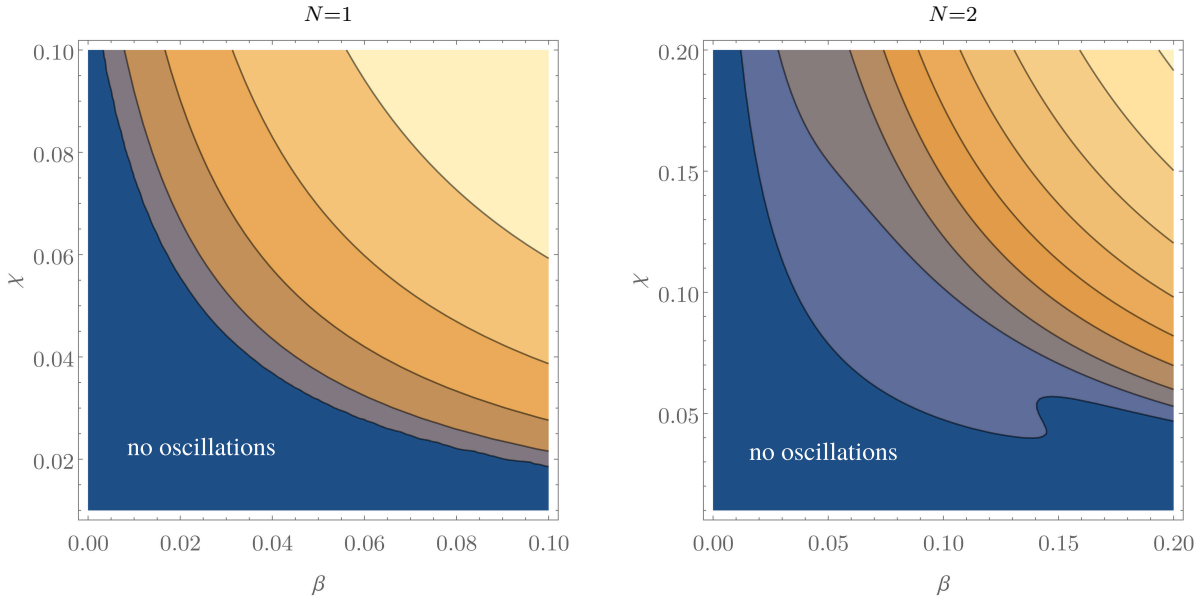


Figure 7: (Left panel) A phase diagram showing the imaginary part of the eigenvalues of the Jacobian associated with the system of ODEs of Equations (10) and (11) for varying values of χ and β in the case of $N = 1$. (Right panel) A phase diagram showing the imaginary part of the eigenvalues of the Jacobian associated with the system of ODEs of Equations (10) and (11) for varying values of χ and β with $\theta = 0$ in the case of $N = 2$. The dark blue area indicates the region in the parameter space for which the imaginary part is zero, and hence, there are no oscillations.

6. Estimation

We estimate the parameters of the model by targeting the temporal evolution of the average number of collaborations shown in Figure 2 in Section 2. We focus on this statistic not only because it is of primary interest for the analysis in this paper, but also because it captures the time varying pattern of the R&D collaboration network using all firms in the data sample in a concise way. As we have learned from the theoretical analysis of the model in Section 5, the empirically relevant case is the model specification with multi-product competition. Moreover, in order to reduce the computational burden for the estimation procedure, we set the number of technologies to two. We then estimate the parameters of the model using the Likelihood-Free Markov Chain Monte Carlo (LF-MCMC) algorithm suggested by Marjoram et al. [2003].[37] Figure 8 shows a comparison of the average degree from the estimated model for $N = 2$ and the empirical observation.[38] The figure illustrates that the model can reproduce well the oscillatory pattern that we observe in the data, when we allow for product competition across sectors.

7. Policy Implications

In the following sections we study counterfactual policy scenarios due to changes in parameters related to collaboration and entry costs. Our objective function will be net total output, $\overline{Y}$, given by total output minus the cost of production, R&D collaborations and entry.[39] More precisely, in Section 7.1 we study the impact on net output of an R&D collaboration costs reducing policy. Moreover, in Section 7.2 we investigate the impact of a reduction in the entry cost from policies aimed at deregulation or tax benefits for start-ups on net output. We will only consider the multi-

[37] A more detailed description of the estimation algorithm can be found in Appendix L, and the parameter estimates obtained from this procedure are shown in Appendix Table L.1.

[38] The parameters used are $\theta = 0$, $\alpha = 0$, $\beta = 0.2$, $b = 1250$, $c = 1.3$, $\gamma = 0.0001$, $\eta = 1.5$, $\lambda = 2.1$, $n = 500$, $\rho = 50$ and $\tau = 0.01$, $\xi = 0.45$, $\chi = 0.45$, $\vartheta = 15$ and $\kappa = 0.7$. The initial condition is $\bar{x}_0(1) = \bar{x}_0(2) = 0$ and an empty network.

[39] An explicit computation can be found in Equation (62) in Appendix H.1. Further, Appendix H.2 provides an analysis of the output maximizing (efficient) network structure.

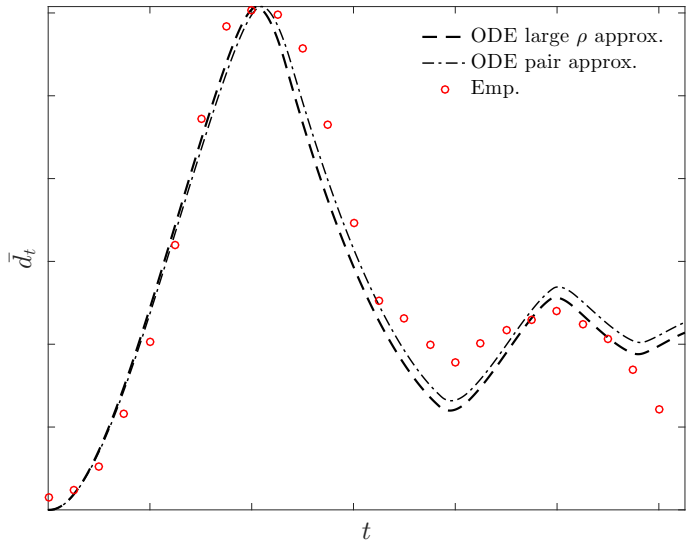


Figure 8: Comparison of the average degree $\bar{d}$ from the prediction of the theoretical model for $N = 2$ and the empirical observations indicated with circles. The parameters used are $\theta = 0$, $\alpha = 0$, $\beta = 0.2$, $b = 1250$, $c = 1.3$, $\gamma = 0.0001$, $\eta = 1.5$, $\lambda = 2.1$, $n = 500$, $\rho = 50$ and $\tau = 0.01$, $\xi = 0.45$, $\chi = 0.45$, $\vartheta = 15$ and $\kappa = 0.7$. The initial condition is $x_0(0) = \bar{x}_0(1) = \bar{x}_0(2) = 0$ and an empty network. Dashed lines correspond to the large ρ approximation of Appendix G.2, while dashed-dotted lines represent the pair approximation of Appendix J.2.

product competition case, as this is the empirically relevant case, and use the parameter estimates of Section 6 as our benchmark scenario.[40]

7.1. R&D Collaboration Subsidies

Many governments provide R&D subsidies to foster R&D collaborations between private firms [cf. e.g. Cohen, 1994; Czarnitzki et al., 2007]. One example is the Advanced Technology Program (ATP) which was administered by the National Institute of Standards and Technology (NIST) in the United States. In Europe, EUREKA is a Europe-wide network for industrial R&D. The main aim of this EU programme is to strengthen European competitiveness in the field of R&D by means of promoting market-driven collaborative research and technology development. Another example can be found in Germany, where the federal government provides R&D subsidies to stimulate collaboration activities between private organizations [Broekel and Graf, 2012].

In this section we analyze the impact of changes in the collaboration cost c due to R&D subsidies on net output. The top panels in Figure 9 show the number of active firms, n_a, and net output $\overline{Y}_0(\mathbf{h}, G)$ as a function of the linking cost, c, for different values of β (assuming that there are two technologies, $N = 2$). Comparing net output for different values of β with increasing collaboration costs c shows a non-monotonic behavior. For low values of β, net output is increasing with decreasing collaboration costs, c.[41] In contrast, when the spillover parameter β is high, then a decline in the collaboration cost c reduces net output. The reason for this effect is the large decline in the number of active firms, n_a (lower entry rates) that comes from a large increase in the average stock of knowledge $\bar{h}$ of the incumbents, which compensates for the direct effect of the decline in $\bar{h}$ on net output. The detrimental effect of having only few active firms with little entry is that if one of the active firms is hit by a shock and exits, then this has large negative consequences for the level of production and hence net output in the economy. We therefore find that the effectiveness of

[40] In Appendix H.3 we further analyze the effect of increased in-house R&D innovation rates and spillovers on inequality.

[41] This is because the negative effect of an increase in the cost c on the average knowledge stock $\bar{h}$ and the direct effect on the total collaboration cost exceed the increase in the coefficient of variation, CV_h, the moderate increase in the number of active firms, n_a, (from higher expected profits of potential entrants due to the lower average knowledge stock of incumbents) and the reduction in the average degree, $\bar{d}$ (and a lower total collaboration cost) in the net output function of Equation (64) derived in Appendix H.

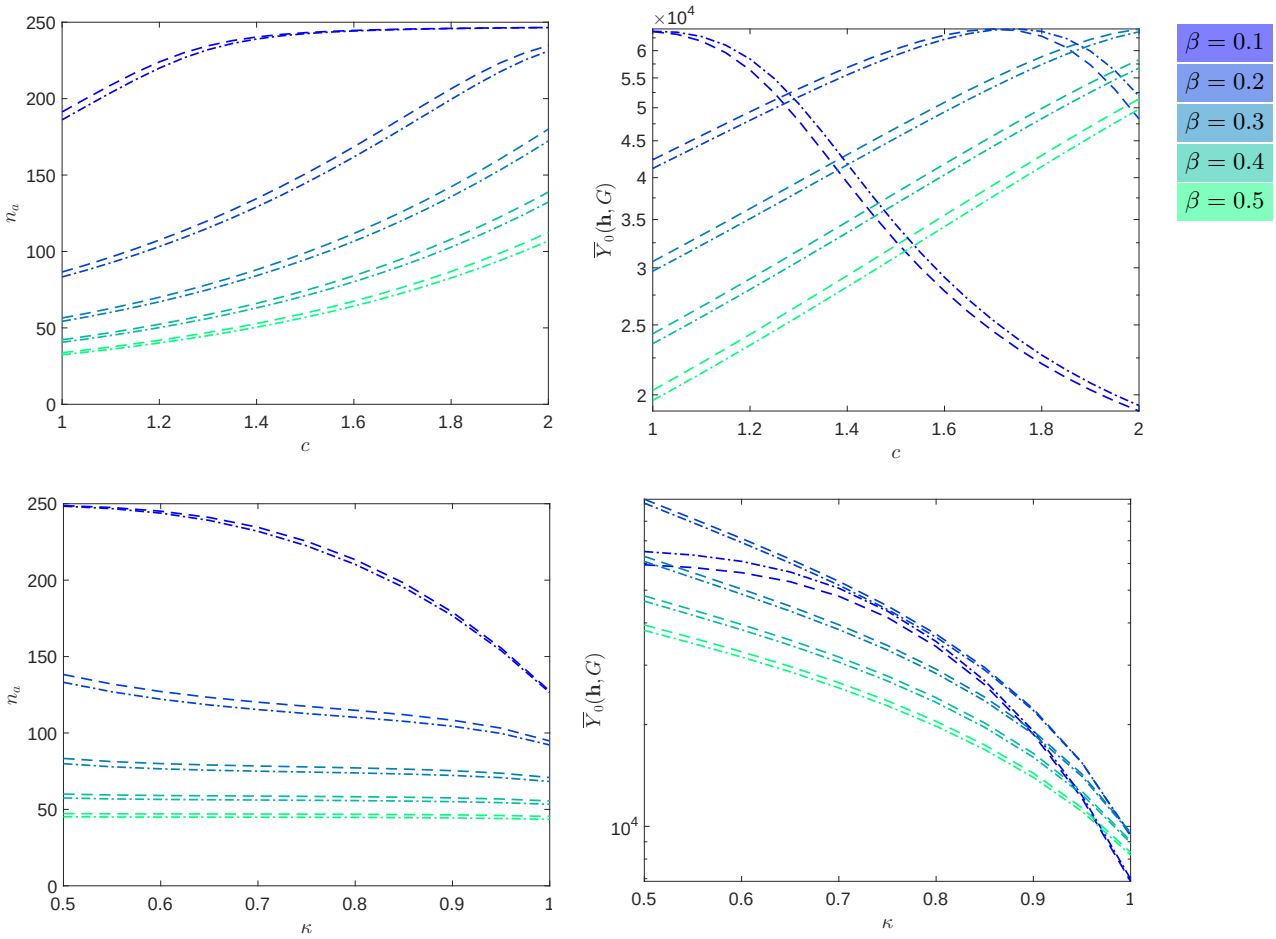


Figure 9: (Top left panel) the number of active firms, n_a, as a function of the linking cost, c, for different values of β. (Top right panel) Net output, $\overline{Y}_0(\mathbf{h}, G)$, as a function of the linking cost, c, for different values of β (cf. Equation (64) in Appendix H.1). The number of technologies is assumed to be $N = 2$. (Bottom left panel) the number of active firms, n_a, as a function of the entry cost, κ, for different values of β. (Bottom right panel) Net output, $\overline{Y}_0(\mathbf{h}, G)$, as a function of the entry cost, κ, for different values of β (cf. Equation (64) in Appendix H.1). The number of technologies is assumed to be $N = 2$. The parameters used correspond to Figure 8. Dashed lines correspond to the large ρ approximation of Appendix G.2, while dashed-dotted lines represent the pair approximation of Appendix J.2.

policies that reduce the cost of R&D collaborations depend on the extent of knowledge spillovers between collaborating firms, and how detrimental such policies might be to firm entry (because they typically benefit incumbent firms and not entrants).

7.2. Deregulation, Tax Benefits for Start-Ups and Cost of Entry

During the last two decades many European countries implemented reforms to deregulate their product markets. For example, Germany exhibited a decline of the OECD's index of product market regulation since the mid 1990s [cf. Felbermayr et al., 2014]. To examine the impact of such a reform, we consider a change in the entry costs [cf. Acemoglu and Cao, 2010]. Such a change in the entry cost could also result from tax benefits to start-up companies, a measure that is often used by policy makers to foster innovation and generate growth [cf. e.g. Lazear, 2005].

We analyze the impact of changes in the entry cost, κ, on net output. The bottom panels in Figure 9 show the number of active firms, n_a, and net output $\overline{Y}_0(\mathbf{h}, G)$ as a function of the entry cost, κ, for different values of β. When we compare net output for different values of β with increasing entry costs κ we find that net output is unanimously increased with decreasing κ. This is because an increase in the entry cost κ has a direct negative effect on net output, and further lowers the average stock of knowledge $\bar{h}$ as well as the number of active firms, n_a.[42] Thus, as compared to policies reducing the cost of R&D collaborations (cf. Section 7.1), policies that target lowering the entry cost unambiguously raise net output. We therefore conclude that a sensible R&D policy

[42] This effect is not compensated for by a higher coefficient of variation CV_h or a slightly lower average degree $\bar{d}$ (and hence a lower total collaboration cost) in the net output function $\overline{Y}_0(\mathbf{h}, G)$. See Equation (64) in Appendix H.1.

should balance subsidies for entrants as well as R&D collaborations between incumbent firms.

8. Conclusion

In this paper we have analyzed the co-evolutionary dynamics of firms' technology portfolios and the formation of R&D collaborations that influence – and are influenced by – the technology portfolios. We investigated the stationary states of this dynamic process, and showed that there exists a critical level for the technology spillover parameter below which no significant innovation takes place in the economy. Moreover, we have analyzed the impact of competition across multiple products on innovation and R&D network formation, and find that it is output increasing as long as it does not deter entry. This is due to the fact that product competition leads to reallocation and the replacement of less productive firms with more productive ones in the economy. We have further identified the efficient network structure as a nested split graph, which is characterized by a core-periphery structure, that facilitates the diffusion of technologies through the network while economizing on the number of links [cf. König et al., 2011]. The stability analysis of our model indicates that the R&D collaboration intensity can exhibit a cyclical pattern, which can be described as a damped oscillation. This novel observation is confirmed by analysis of an empirical sample of a large R&D collaboration network over the years 1985 to 2012. We provide a formal explanation for this novel empirical observation, and our results indicate that the cyclicality in the data is a competition effect.

References

Acemoglu, D., Aghion, P., and Zilibotti, F. (2006). Distance to frontier, selection, and economic growth. *Journal of the European Economic Association*, 4(1):37–74.

Acemoglu, D., Akcigit, U., and Kerr, W. R. (2016). Innovation network. *Proceedings of the National Academy of Sciences*, page 201613559.

Acemoglu, D. and Cao, D. (2010). Innovation by entrants and incumbents. *NBER Working Paper No. 16411*.

Acemoglu, D., Ozdaglar, A., and Yildiz, E. (2011). Diffusion of innovations in social networks. In *European Control Conference (CDC-ECC)*, pages 2329–2334. IEEE.

Adams, J. D. (1990). Fundamental stocks of knowledge and productivity growth. *Journal of Political Economy*, 98(4):673–702.

Akerlof, G. A. (1970). The market for "lemons": Quality uncertainty and the market mechanism. *The quarterly journal of economics*, pages 488–500.

Anderson, S. P., De Palma, A., and Thisse, J. F. (1992). *Discrete choice theory of product differentiation*. MIT Press.

Baum, J. A., Cowan, R., and Jonard, N. (2010). Network-independent partner selection and the evolution of innovation networks. *Management Science*, 56(11):2094–2110.

Baye, M. and Hoppe, H. (2003). The strategic equivalence of rent-seeking, innovation, and patent-race games. *Games and Economic Behavior*, 44(2):217–226.

Berliant, M. and Fujita, M. (2008). Knowledge creation as a square dance on the hilbert cube. *International Economic Review*, 49(4):1251–1295.

Bernard, A. B., Redding, S. J., and Schott, P. K. (2010). Multiple-product firms and product switching. *The American Economic Review*, 100(1):70–97.

Bilbiie, F. O., Ghironi, F., and Melitz, M. J. (2012). Endogenous entry, product variety, and business cycles. *Journal of Political Economy*, 120(2):304–345.

Bloch, F. (1995). Endogenous structures of association in oligopolies. *The RAND Journal of Economics*, 26(3):537–556.

Bloom, N. (2009). The impact of uncertainty shocks. *Econometrica*, 77(3):623–685.

Bloom, N. and Van Reenen, J. (2010). Why do management practices differ across firms and countries? *The Journal of Economic Perspectives*, pages 203–224.

Blume, L., Easley, D., Kleinberg, J., Kleinberg, R., and Tardos, E. (2011). Network formation in the presence of contagious risk. In *Proceedings of the 12th ACM conference on Electronic commerce*, pages 1–10. ACM.

Broekel, T. and Graf, H. (2012). Public research intensity and the structure of german R&D networks: a comparison of 10 technologies. *Economics of Innovation and New Technology*, 21(4):345–372.

Caballero, R. and Jaffe, A. (1993). How high are the giants' shoulders: An empirical assessment of knowledge spillovers and creative destruction in a model of economic growth. *NBER Macroeconomics Annual 1993*, 8:11.

Cabrales, A., Calvó-Armengol, A., and Zenou, Y. (2011). Social interactions and spillovers. *Games and Economic Behavior*, 72(2):339–360.

Calvo, G. A. (1983). Staggered prices in a utility-maximizing framework. *Journal of Monetary Economics*, 12(3):383–398.

Cohen, L. (1994). When can government subsidize research joint ventures? politics, economics, and limits to technology policy. *The American Economic Review*, 84(2):159–163.

Cohen, W. M. and Levinthal, D. A. (1990). Absorptive capacity: A new perspective on learning and innovation. *Administrative Science Quarterly*, 35(1):128–152.

Cowan, R. and Jonard, N. (2008). If the alliance fits...: Innovation and network dynamics. *Network Strategy: Advances in Strategic Management*, 25.

Cowan, R. and Jonard, N. (2009). Knowledge portfolios and the organization of innovation networks. *The Academy of Management Review*, 34(2):320–342.

Cowan, R., Jonard, N., and Zimmermann, J. (2007). Bilateral Collaboration and the Emergence of Innovation Networks. *Management Science*, 53(7):1051.

Czarnitzki, D., Ebersberger, B., and Fier, A. (2007). The relationship between R&D collaboration, subsidies and R&D performance: empirical evidence from Finland and Germany. *Journal of applied econometrics*, 22(7):1347–1366.

Czarnitzki, D., Hussinger, K., and Schneider, C. (2015). R&D collaboration with uncertain intellectual property rights. *Review of Industrial Organization*, 46(2):183–204.

D'Aspremont, C. and Jacquemin, A. (1988). Cooperative and noncooperative R&D in duopoly with spillovers. *The American Economic Review*, 78(5):1133–1137.

Dawid, H. and Hellmann, T. (2014). The evolution of R&D networks. *Journal of Economic Behavior & Organization*, 105:158–172.

DeMarzo, P., Vayanos, D., and Zwiebel, J. (2003). Persuasion Bias, Social Influence, and Unidimensional Opinions. *Quarterly Journal of Economics*, 118(3):909–968.

Dixit, A. (1987). Strategic behavior in contests. *The American Economic Review*, pages 891–898.

Duranton, G., Martin, P., Mayer, T., Mayneris, F., et al. (2010). *The economics of clusters: Lessons from the French experience*. Oxford University Press.

Elaydi, S. (2005). *An introduction to difference equations*. Springer.

Felbermayr, G., Impullitti, G., and Prat, J. (2014). Firm dynamics and residual inequality in open economies. *University of Nottingham, Working Paper No. 2014/05*.

Fleming, L. (2001). Recombinant uncertainty in technological search. *Management science*, 47(1):117–132.

Fosco, C., Vega-Redondo, F., and Marsili, M. (2010). Peer effects and peer avoidance: The diffusion of behavior in coevolving networks. *Journal of the European Economic Association*, 8(1):169–202.

Futia, C. (1980). Schumpeterian competition. *The Quarterly Journal of Economics*, 94(4):675–695.

Gabaix, X. (2014). A sparsity-based model of bounded rationality. *The Quarterly Journal of Economics*, 1661:1710.

García-Canal, E., Valdés-Llaneza, A., and Sánchez-Lorda, P. (2008). Technological flows and choice of joint ventures in technology alliances. *Research Policy*, 37(1):97–114.

Gay, B. and Dousset, B. (2005). Innovation and network structural dynamics: Study of the alliance network of a major sector of the biotechnology industry. *Research Policy*, 34(10):1457–1475.

Golbe, D. and White, L. (1993). Catch a wave: The time series behavior of mergers. *The Review of Economics and Statistics*, pages 493–499.

Golub, B. and Jackson, M. (2012). How homophily affects the speed of learning and best-response dynamics. *The Quarterly Journal of Economics*, 127(3):1287–1338.

Gomes-Casseres, B. (1986). Joint venture cycles: The evolution of ownership strategies of US MNEs, 1945–75. Vol. 87. Division of Research, Harvard Business School, 1986.

Goyal, S. and Joshi, S. (2003). Networks of collaboration in oligopoly. *Games and Economic Behavior*, 43(1):57–85.

Griffith, R., Redding, S., and Van Reenen, J. (2003). R&D and absorptive capacity: Theory and empirical evidence. *Scandinavian Journal of Economics*, 105(1):99–118.

Gross, T., D'Lima, C. J. D., and Blasius, B. (2006). Epidemic dynamics on an adaptive network. *Physical review letters*, 96(20):208701.

Gulati, R. and Gargiulo, M. (1999). Where do interorganizational networks come from? *The American Journal of Sociology*, 104(5):1439–1493.

Gulati, R., Sytch, M., and Tatarynowicz, A. (2010). The rise and fall of small worlds: Exploring the dynamics of social structure. *Organization Science, Articles in Advance*, pages 1–23.

Hagedoorn, J. (2002). Inter-firm R&D partnerships: an overview of major trends and patterns since

1960. *Research Policy*, 31(4):477–492.
Hagedoorn, J. and van Kranenburg, H. (2003). Growth patterns in R&D partnerships: an exploratory statistical study. *International Journal of Industrial Organization*, 21(4):517 – 531.
Hanaki, N., Nakajima, R., and Ogura, Y. (2010). The dynamics of R&D network in the IT industry. *Research Policy*, 39(3):386–399.
Hanemann, W. M. (1984). Discrete/continuous models of consumer demand. *Econometrica*, 52(3):541–561.
Harrigan, K. (1988). Strategic alliances and partner asymmetries. *Management international review*, 28(4):53–72.
Hartwick, J. M. (1982). Efficient prizes in prototype development contests. *Economics Letters*, 10(3):375–379.
Hojman, D. A. and Szeidl, A. (2006). Endogenous networks, social games, and evolution. *Games and Economic Behavior*, 55(1):112–130.
Horn, R. A. and Johnson, C. R. (1990). *Matrix Analysis*. Cambridge University Press.
Hurd, T. R. (2015). *Contagion! The Spread of Systemic Risk in Financial Networks*. Springer.
Jackson, M. (2008). *Social and Economic Networks*. Princeton University Press.
Jackson, M. and Golub, B. (2010). Naive learning in social networks: Convergence, influence and wisdom of crowds. *American Economic Journal: Microeconomics*, 2(1):112–149.
Jackson, M. and Rogers, B. (2007a). Relating network structure to diffusion properties through stochastic dominance. *The BE Journal of Theoretical Economics*, 7(1):1–13.
Jackson, M. O. and Rogers, B. W. (2007b). Meeting strangers and friends of friends: How random are social networks? *American Economic Review*, 97(3):890–915.
Jackson, M. O. and Watts, A. (2002). The evolution of social and economic networks. *Journal of Economic Theory*, 106(2):265–295.
Jia, H. (2008). A stochastic derivation of the ratio form of contest success functions. *Public Choice*, 135(3):125–130.
Jones, C. I. (1995). R & D-based models of economic growth. *The Journal of Political Economy*, 103:759–784.
Jones, C. I. (2005). Growth and ideas. *Handbook of economic growth*, 1:1063–1111.
Jovanovic, B. (2009). The technology cycle and inequality. *The Review of Economic Studies*, 76(2):707–729.
Jovanovic, B. and MacDonald, G. M. (1994). Competitive diffusion. *Journal of Political Economy*, 102(1):24–52.
Jovanovic, B. and Rob, R. (1989). The growth and diffusion of knowledge. *The Review of Economic Studies*, 56(4):569–582.
Kandori, M., Mailath, G., and Rob, R. (1993). Learning, mutation, and long run equilibria in games. *Econometrica*, 61(1):29–56.
Karrer, B., Newman, M. E., and Zdeborová, L. (2014). Percolation on sparse networks. *Physical review letters*, 113(20):208702.
Kelly, M. J., Schaan, J.-L., and Joncas, H. (2002). Managing alliance relationships: key challenges in the early stages of collaboration. *R&D Management*, 32(1):11–22.
Ken, Y., Tsai, T.-Y., and Ou, Y.-K. (2008). Study of the time lag effect of patent impact on profitability of us pharmaceutical industry from innovation to profit. In *Management of Engineering & Technology, 2008. PICMET 2008.*, pages 2588–2596. IEEE.
Khalil, H. K. (2002). *Nonlinear Systems*. Prentice Hall.
Klette, T. and Kortum, S. (2004). Innovating firms and aggregate innovation. *Journal of Political Economy*, 112(5):986–1018.
Kogut, B. (1988). Joint ventures: Theoretical and empirical perspectives. *Strategic Management Journal*, 9(4):319–332.
König, M. D., Battiston, S., Napoletano, M., and Schweitzer, F. (2011). The efficiency and stability of R&D networks. *Games and Economic Behavior*, 75:694 –713.
König, M. D., Liu, X., and Zenou, Y. (2014). R&D networks: Theory, empirics and policy implications. *CEPR Discussion Paper No. 9872*.
König, M. D., Lorenz, J., and Zilibotti, F. (2016). Innovation vs. imitation and the evolution of productivity distributions. *Theoretical Economics*, 11:1053–1102.
Kretschmer, T., Miravete, E. J., and Pernías, J. C. (2012). Competitive pressure and the adoption of complementary innovations. *The American Economic Review*, 102(4):1540–1570.
Lazear, E. (2005). Entrepreneurship. *Journal of Labor Economics*, 23(4):649–680.
Lee, Y. and Mukoyama, T. (2008). Entry, exit, and plant-level dynamics over the business cycle. *Working Paper No. 07–18, Fed. Reserve Bank Cleveland*.
Loury, G. C. (1979). Market structure and innovation. *The Quarterly Journal of Economics*, 93(3):395–410.

Mahadev, N. and Peled, U. (1995). *Threshold Graphs and Related Topics*. North Holland.
Marjoram, P., Molitor, J., Plagnol, V., and Tavaré, S. (2003). Markov chain Monte Carlo without likelihoods. *Proceedings of the National Academy of Sciences*, 100(26):15324.
Matsuyama, K. (1999). Growing through cycles. *Econometrica*, 67(2):335–347.
Meagher, K. and Rogers, M. (2004). Network density and R&D spillovers. *Journal of Economic Behavior & Organization*, 53(2):237–260.
Menezes, F. M. and Quiggin, J. (2010). Markets for influence. *International Journal of Industrial Organization*, 28(3):307–310.
Montanari, A. and Saberi, A. (2010). The spread of innovations in social networks. *Proceedings of the National Academy of Sciences*, 107(47):20196.
Morris, S. (2000). Contagion. *The Review of Economic Studies*, 67(1):57–78.
Nocke, V. and Schutz, N. (2016). Multiproduct-firm oligopoly: An aggregative games approach. *Mimeo, University of California Los Angeles*.
Novshek, W. and Sonnenschein, H. (1979). Marginal consumers and neoclassical demand theory. *Journal of Political Economy*, 87(6):1368–1376.
Paolella, M. S. (2007). *Intermediate Probability*. Wiley.
Park, S. and Russo, M. (1996). When competition eclipses cooperation: An event history analysis of joint venture failure. *Management science*, pages 875–890.
Podolny, J. M. and Page, K. L. (1998). Network forms of organization. *Annual review of sociology*, 24(1):57–76.
Powell, W. W., White, D. R., Koput, K. W., and Owen-Smith, J. (2005). Network dynamics and field evolution: The growth of interorganizational collaboration in the life sciences. *American Journal of Sociology*, 110(4):1132–1205.
Reinganum, J. (1985). Innovation and industry evolution. *The Quarterly Journal of Economics*, 100(1):81–99.
Rogers, T. (2015). Assessing node risk and vulnerability in epidemics on networks. *EPL (Europhysics Letters)*, 109(2):28005.
Rogers, T., Clifford-Brown, W., Mills, C., and Galla, T. (2012). Stochastic oscillations of adaptive networks: application to epidemic modelling. *Journal of Statistical Mechanics: Theory and Experiment*, 2012(08):P08018.
Rosenkopf, L. and Padula, G. (2008). Investigating the Microstructure of Network Evolution: Alliance Formation in the Mobile Communications Industry. *Organization Science*, 19(5):669–687.
Sandholm, W. (2010). *Population games and evolutionary dynamics*. MIT Press.
Saxenian, A. L. (1994). *Regional Advantage: Culture and Competition in Silicon Valley and Route 128*. Harvard University Press, Cambridge.
Schilling, M. (2009). Understanding the alliance data. *Strategic Management Journal*, 30(3):233–260.
Schumpeter, J. (1934). *Theory of Economic Development: An Inquiry Into Profits, Capital, Credit, Interest, and the Business Cycle*. Harvard University Press. first published in 1911.
Sørensen, J. B. and Stuart, T. E. (2000). Aging, obsolescence, and organizational innovation. *Administrative science Quarterly*, 45(1):81–112.
Sørenson, O. and Rivkin, J. W. and Fleming, L. (2006). Complexity, networks and knowledge flow. *Research policy*, 35(7):994–1017.
Stuart, T. E. (2000). Interorganizational alliances and the performance of firms: A study of growth and innovation rates in a high-technology industry. *Strategic management journal*, 21(8):791–811.
Szidarovszky, F. and Okuguchi, K. (1997). On the existence and uniqueness of pure nash equilibrium in rent-seeking games. *Games and Economic Behavior*, 18(1):135–140.
Tullock, G. (1980). *Toward a theory of the rent-seeking society*, volume 97, chapter Efficient rent seeking, page 112. Texas A & M University Press.
Veselić, K. (1979). On real eigenvalues of real tridiagonal matrices. *Linear Algebra and its Applications*, 27:167–171.
Watts, A. (2001). A dynamic model of network formation. *Games and Economic Behavior*, 34(2):331–341.
Westbrock, B. (2010). Natural concentration in industrial research collaboration. *The RAND Journal of Economics*, 41(2):351–371.
Young, H. P. (2002). The diffusion of innovations in social networks. *The Economy as an Evolving Complex System. Proceedings volume in the Santa Fe Institute studies in the sciences of complexity*, 3:267–282.

Appendix

A. Micro-level Empirical Evidence

From our sample of 21,478 firms we could obtain patent information for 4,223 of them (19.66%) from the European Patent Office (EPO). We matched the firms in our alliance data with the assignees in the EPO Worldwide Patent Statistical Database (PATSTAT). We classified the patents according to the USPTO 3-digit classification system as of 2008 [see also Hall et al., 2001]. This allowed us to construct the technology portfolios for the subset of the firms for which patent data was available, and we obtained 261 unique patent classes for the matched firms. We take the year 2010 as our reference case.[43] With the patent data and the R&D collaboration data we can illustrate the two-way influence of patent portfolios and R&D collaborations at the firm level.

Do R&D collaborations affect the diffusion of technologies? On the one hand, R&D collaborations facilitate the diffusion of technologies across firms. On the other hand, technological opportunities through learning and imitation from other firms' patent portfolios determine the creation of R&D collaborations. To illustrate the first effect, i.e. the impact of R&D collaborations on the technology portfolios, we estimate the following Poisson regression model [cf. Cameron and Trivedi, 2005]

$$\mathbb{P}(h_{ik,t} = 1) = 1 - e^{-\left(\alpha_0 + \alpha_1 |\mathsf{S}(\mathbf{h}_{it})| + \alpha_2 \sum_{j=1}^{n} a_{ij,t} h_{jk,t}\right)}. \tag{21}$$

We obtain the estimates $\hat{\alpha}_0 = 0.0042^{***}$ (0.0001) and $\hat{\alpha}_1 = 0.0557^{***}$ (0.0013) with standard errors reported in parenthesis for the year $t = 2010$. In particular, we find that the estimate for the technology spillover coefficient, $\hat{\alpha}_1$, is positive and highly significant. This illustrates the importance of R&D collaborations for the diffusion of ideas across firms [cf. Jovanovic and MacDonald, 1994], and provides empirical support the innovation arrival rate that we have introduced in Equation (4).

Do technological opportunities affect the formation of R&D collaborations? In order to illustrate the second effect, i.e. the impact of the technology portfolios on R&D collaborations, we estimate the following the logistic regression model [cf. Cameron and Trivedi, 2005]

$$\mathbb{P}\left(a_{ij,t} = 1\right) = \Lambda\left(\beta_0 + \beta_1 \langle \mathbf{h}^c_{it}, \mathbf{h}_{jt} \rangle\right) \cdot \Lambda\left(\beta_0 + \beta_1 \langle \mathbf{h}^c_{jt}, \mathbf{h}_{it} \rangle\right), \tag{22}$$

where $\Lambda(x) \equiv e^x/(1+e^x)$ is the logistic function, interpolating smoothly between zero and one. We obtain the estimates $\hat{\beta}_0 = -4.7240^{***}$ (0.0200) and $\hat{\beta}_1 = 0.0123^{***}$ (0.0009) with bootstrapped standard errors in parenthesis for the year $t = 2010$. That is, we obtain a positive and significant coefficient $\hat{\beta}_1$ for the effect of the ideas j possesses but i does not, $\langle \mathbf{h}^c_{it}, \mathbf{h}_{jt} \rangle$, and i possesses but j does not, $\langle \mathbf{h}^c_{jt}, \mathbf{h}_{it} \rangle$, on their propensity to form a collaboration. This indicates the importance of complementarities in the technology portfolios for the creation of R&D collaborations [cf. Jovanovic and Rob, 1989], and is consistent with the link formation probability introduced in Equation (8).

The above exploratory empirical results serve as a motivation for our general model of network formation and technology diffusion introduced in Sections 3.2 and 3.3.

B. Single- vs. Multi-Product Competition

In the following we introduce an intermediate and final goods producing sector and derive the corresponding firms' profits from production of Section 3.1. We will consider two alternative environments, one in which each firm is producing only a single intermediate good, and one in which firms can produce multiple intermediate goods and compete with each other over supplying these goods. We refer to the first as *single-product monopolies* while the latter is referred to as *multi-product*

[43] The results remain robust, however, when choosing different years for the estimation.

competition. For notational simplicity we omit the time dependency of the firms' productivities and output levels.

Single-product monopolies. We consider a Schumpeterian model of monopolistic competition as in e.g. Acemoglu et al. [2006]. A unique final good, denoted by Y, is produced by a representative competitive firm using labor and a set of intermediate goods x_i, $i \in \mathcal{N} = \{1, \ldots, n\}$, according to the production function

$$Y = \frac{1}{\phi} L^{1-\phi} \sum_{i=1}^{n} A_i^{1-\phi} x_i^{\phi}, \quad \phi \in (0,1),$$

where x_i is the economy's input of intermediate good i and A_i is the productivity of the firm in sector i. We further normalize the labor force to unity, $L = 1$. The final good Y is used for consumption and also as an input to the production of intermediate goods. The profit maximization program yields the following inverse demand function for intermediate goods, $p_i = \left(\frac{A_i}{x_i}\right)^{1-\phi}$, where the price of the final good is set to be the numeraire. Each intermediate good i is produced by a firm i with constant marginal cost φ. The firm sets the price equal to $p_i = \frac{\varphi}{\phi}$, and sells at that price the equilibrium quantity $x_i = \phi^{\frac{1}{1-\phi}} \varphi^{-\frac{1}{1-\phi}} A_i$. The gross profit earned by a firm i in an intermediate sector i, not taking into account any R&D collaboration or start-up costs, will then be a linear function of its productivity

$$\tilde{\pi}_i = (p_i - \varphi)\, x_i = \psi A_i, \tag{23}$$

where $\psi \equiv \frac{1-\phi}{\phi} \phi^{\frac{1}{1-\phi}} \varphi^{-\frac{1}{1-\phi}}$, which is monotonically decreasing in ϕ and φ. Total output is then proportional to aggregate productivity as $Y = \phi^{\frac{2\phi-1}{1-\phi}} \varphi^{-\frac{\phi}{1-\phi}} \sum_{i=1}^{n} A_i$. The economic environment discussed so far in which each firm is producing only one intermediate good will be referred to as *single-product monopolies*.

Multi-product competition. We also consider an environment in which we assume that a firm can produce more than one intermediate good. We refer to this scenario as the case with *multi-product competition*. In this setup firms can produce multiple varieties in different sectors by applying their technological knowledge in all sectors, similar to variety expanding models [cf. Jones, 1995, 2005]. We assume that the probability that a firm i becomes the supplier in sector j by winning a production contract in that sector, and to produce the intermediate good j, is given by a *contest success function* [cf. Tullock, 1980]. This assumption follows the discrete/continuous choice literature pioneered by Novshek and Sonnenschein [1979] and Hanemann [1984], and more recently used in Nocke and Schutz [2016], where a consumer first decides which product to consume, and then how much of that product to consume.

In particular, we assume that the quality of the intermediate good produced by firm i is given by $\epsilon_i A_i$, where A_i is the productivity of firm i and ϵ_i is a firm specific productivity shock [cf. Bloom, 2009].[44] For each input j, the final good producing firm chooses the firm with the highest quality as a supplier. Assuming that the productivity shocks ϵ_i are independent and follow an inverse exponential distribution [cf. Jia, 2008], one can show that the probability that firm i has the highest

[44] The error term, ϵ_i, can also represent imperfect buyer information [cf. Akerlof, 1970], where the quality of the intermediate good produced by firm i with productivity A_i is observed by the final good producing firm with a multiplicative error term ϵ_i. Alternative interpretations include a mismatch between the input requirements of the final good producing firm and the characteristics of the intermediate goods [Bernard et al., 2010], an intermediate goods producing firm specific random effect affecting the characteristics of the intermediate good, or can represent different managerial practices that can affect the output produced [cf. Bloom and Van Reenen, 2010].

quality and thus supplies input j to the final good producing firm is given by[45]

$$\mathbb{P}(\text{firm } i \text{ produces variety } j) = \mathbb{P}(\epsilon_i A_i > \epsilon_k A_k, \forall k \neq i) = \frac{A_i}{\sum_{k=1}^{n} A_k}. \tag{24}$$

Equation (24) states that more productive firms are more likely to become the suppliers for every intermediate goods variety. Hence, as in the model by Bernard et al. [2010], more productive firms have a larger product range. This is consistent with the data. In particular, using U.S. census data, Bernard et al. [2010] find that multiple-product firms have higher measured revenue-based productivity than single-product firms.

If firm i becomes the producer, it earns a gross profit of ψA_i. The contests in each sector are assumed to be independent. Firm i's expected gross profit from all n sectors is then given by[46]

$$\tilde{\pi}_i = \psi A_i n \frac{A_i}{\sum_{j=1}^{n} A_j} = \psi \frac{A_i^2}{\frac{1}{n}\sum_{j=1}^{n} A_j}. \tag{25}$$

Total gross output is then given by $Y = n\phi^{\frac{2\phi-1}{1-\phi}} \varphi^{-\frac{\phi}{1-\phi}} \sum_{i=1}^{n} \frac{A_i^2}{\sum_{k=1}^{n} A_k}$.[47]

For both cases, with or without product competition, we can write profits by combining Equations (23) and (25) compactly as follows

$$\tilde{\pi} = \theta A_i + (1-\theta)\frac{A_i^2}{\frac{1}{n}\sum_{j=1}^{n} A_j}, \tag{26}$$

where $\theta \in \{0,1\}$ is a (zero/one) parameter distinguishing between single-product monopolies and multi-product competition, and we have normalized $\psi = 1$ Net total output, $\overline{Y}$, is given by gross final output minus the cost of intermediate goods production [cf. König et al., 2016]. This yields

$$\overline{Y}_\theta = \Gamma \sum_{i=1}^{n} \left(\theta + (1-\theta)\frac{A_i}{\frac{1}{n_a}\sum_{j=1}^{n_a} A_j} \right) A_i, \tag{27}$$

for $\theta \in \{0,1\}$, where we have denoted by $\Gamma \equiv (1-\phi)\phi^{\frac{1}{1-\phi}-2}\varphi^{-\frac{\phi}{1-\phi}}$.

C. Proofs

Proof of Proposition 1. When the time τ of a collaboration is short compared to the dynamics of the generation and diffusion of knowledge in the entire industry, we can write the expected stock of knowledge of firm i at time $t+\tau$, given the current knowledge portfolios $\mathbf{h}_t$ and network G_t, as follows $\mathbb{E}_t\left(|\mathsf{S}(\mathbf{h}_{i,t+\tau})|\middle|\mathbf{h}_t, G_t\right) = |\mathsf{S}(\mathbf{h}_{it})| + (\gamma\tau + \alpha\tau|\mathsf{S}(\mathbf{h}_{it})|)|\mathsf{S}(\mathbf{h}_{it}^c)| - \lambda\tau|\mathsf{S}(\mathbf{h}_{it})| + \beta\tau\sum_{j=1}^{n} a_{ij,t}\langle \mathbf{h}_{it}^c, \mathbf{h}_{jt}\rangle + o(\tau)$. The product $\langle \mathbf{h}_{it}^c, \mathbf{h}_{jt}\rangle$ counts the number of ideas i does not know but j knows. Denoting by $f_{it} \equiv |\mathsf{S}(\mathbf{h}_{it})|$ and $\Delta f_{it} \equiv (\gamma + \alpha|\mathsf{S}(\mathbf{h}_{it})|)|\mathsf{S}(\mathbf{h}_{it}^c)| - \lambda|\mathsf{S}(\mathbf{h}_{it})| + \beta\tau\sum_{j=1}^{n} a_{ij,t}\langle \mathbf{h}_{it}^c, \mathbf{h}_{jt}\rangle$, we can write this as $\mathbb{E}_t\left(|\mathsf{S}(\mathbf{h}_{i,t+\tau})|\middle|\mathbf{h}_t, G_t\right) = f_{it} + \Delta f_{it}\tau + o(\tau)$. Considering the network $G_t + ij$ obtained from G_t by adding the link ij we then can write $\mathbb{E}_t\left(|\mathsf{S}(\mathbf{h}_{i,t+\tau})|\middle|\mathbf{h}_t, G_t + ij\right) = f_{it} + \Delta f_{it}\tau + \beta f_{ij,t}\tau + o(\tau)$, where we have also denoted by $f_{ij,t} \equiv \langle \mathbf{h}_{it}^c, \mathbf{h}_{jt}\rangle$.

From Equation (3) it follows that in the case of independent markets when $\theta = 1$, a firm's gross profit is a linear function of its stock of knowledge. We then can write the expected profit of firm i at time $t+\tau$ as $\mathbb{E}_t\left(\pi_i(\mathbf{h}_{t+\tau})|\mathbf{h}_t, G_t\right) = 1 + \mathbb{E}_t\left(|\mathsf{S}(\mathbf{h}_{i,t+\tau t})|\middle|\mathbf{h}_t, G_t\right) - cd_{it} = 1 + f_{it} + \Delta f_{it}\tau + \beta f_{ij,t}\tau - cd_{it} + o(\tau)$. Then the change in the present discounted profit of a firm i from forming the link ij is given by $V_i(\mathbf{h}_t, G_t + ij) - V_i(\mathbf{h}_t, G_t) =$

[45] Alternatively, we could assume that a firm can win a patent race within in a prototype development contest that allows it to produce the intermediate good variety [cf. Futia, 1980; Hartwick, 1982; Reinganum, 1985]. Moreover, following Dixit [1987], we can interpret Equation (24) as the market share of firm i in market j in a homogeneous product oligopoly with a unit-elastic demand. See also Szidarovszky and Okuguchi [1997] and Menezes and Quiggin [2010] for applications of the contest success function in the industrial organization literature.

[46] Let $p_i = \frac{A_i}{\sum_{k=1}^{n} A_k}$ denote the probability that firm i becomes the producer in sector j. Then the expected number of sectors in which firm i is producing is given by $\sum_{j=1}^{n}\binom{n}{j}p_i^j(1-p_i)^{n-j} = np_i$.

[47] In the special case of all firms being equally productive, i.e. $A_i = A$ for all $i = 1, \ldots, n$, total output in the non-competitive and the competitive case is identical and given by $Y = n\phi^{\frac{2\phi-1}{1-\phi}}\varphi^{-\frac{\phi}{1-\phi}}A$, and there is no output gain from allowing more productive firms to produce more intermediate goods (see also Appendix H).

$\delta\left(\beta\tau\langle\mathbf{h}_{it}^c,\mathbf{h}_{jt}\rangle - c\right) + o(\tau)$. To simplify our notation we have assumed in the previous expression that the per period cost for an additional link needs to be paid at the end of a collaboration period.

Next we consider the case of competitive markets when $\theta = 1$. Note that $\mathbb{E}_t\left(|\,\mathsf{S}(\mathbf{h}_{i,t+\tau})|^2|\mathbf{h}_t, G_t\right) = |\,\mathsf{S}(\mathbf{h}_{it})|^2 + 2|\,\mathsf{S}(\mathbf{h}_{it})|((\gamma+\alpha|\,\mathsf{S}(\mathbf{h}_{it})|)|\,\mathsf{S}(\mathbf{h}_{it}^c)| - \lambda|\,\mathsf{S}(\mathbf{h}_{it})| + \beta\sum_{j=1}^n a_{ij,t}\langle\mathbf{h}_{it}^c,\mathbf{h}_{jt}\rangle)\tau + o(\tau)$, which can be written as $\mathbb{E}_t\left(|\,\mathsf{S}(\mathbf{h}_{i,t+\tau})|^2|\mathbf{h}_t, G_t\right) = f_{it}^2 + 2f_{it}\Delta f_{it}\tau + o(\tau)$. Adding the link ij yields $\mathbb{E}_t\left(|\,\mathsf{S}(\mathbf{h}_{i,t+\tau})|^2|\mathbf{h}_t, G_t + ij\right) = f_{it}^2 + 2f_{it}\Delta f_{it}\tau + 2\beta f_{it}f_{ij,t}\tau + o(\tau)$. Moreover, we have that $\bar{h}_{t+\tau}(\mathbf{h}_t, G_t) \equiv \mathbb{E}_t\left(\frac{1}{n}\sum_{i=1}^n|\,\mathsf{S}(\mathbf{h}_{i,t+\tau})|\,|\,\mathbf{h}_t, G_t\right) = \bar{h}_t + \Delta\bar{h}_t\tau + o(\tau)$, where we have denoted by $\bar{h}_t \equiv \frac{1}{n}\sum_{i=1}^n|\,\mathsf{S}(\mathbf{h}_{it})|$ and $\Delta\bar{h}_t \equiv \gamma N - \frac{\lambda+\gamma}{n}\sum_{i=1}^n|\,\mathsf{S}(\mathbf{h}_{it})| + \frac{\alpha}{n}\sum_{i=1}^n|\,\mathsf{S}(\mathbf{h}_{it})||\,\mathsf{S}(\mathbf{h}_{it}^c)| + \frac{\beta}{n}\sum_{i=1}^n\sum_{j=1}^n a_{ij,t}\langle\mathbf{h}_{it}^c,\mathbf{h}_{jt}\rangle$. Similarly we get $\mathbb{E}_t\left(\frac{1}{n}\sum_{i=1}^n|\,\mathsf{S}(\mathbf{h}_{i,t+\tau})|\,|\,\mathbf{h}_t, G_t + ij\right) = \bar{h}_t + \Delta\bar{h}_t\tau + \frac{1}{n}\beta\tau(f_{ij,t} + f_{ji,t}) + o(\tau)$. Using a Taylor expansion around the mean (see e.g. Chap. 2.3 in Paolella [2007]),[48] we then can write[49]

$$\mathbb{E}_t\left(\frac{1}{1+b\frac{1}{n}\sum_{j=1}^n|\,\mathsf{S}(\mathbf{h}_{j,t+\tau})|}\middle|\mathbf{h}_t, G_t + ij\right) = \frac{1}{1+b\mathbb{E}_t\left(\frac{1}{n}\sum_{j=1}^n|\,\mathsf{S}(\mathbf{h}_{j,t+\tau})|\middle|\mathbf{h}_t, G_t + ij\right)} + O\left(\frac{1}{n}\right),$$

and

$$\frac{1}{1+b\mathbb{E}_t\left(\frac{1}{n}\sum_{j=1}^n|\,\mathsf{S}(\mathbf{h}_{j,t+\tau})|\middle|\mathbf{h}_t, G_t + ij\right)} = \frac{1}{1+b(\bar{h}_t + \Delta\bar{h}_t\tau)} + O\left(\frac{\tau}{n}\right).$$

It then follows that $\mathbb{E}_t\left(\frac{1}{1+b\frac{1}{n}\sum_{j=1}^n|\,\mathsf{S}(\mathbf{h}_{j,t+\tau})|}\middle|\mathbf{h}_t, G_t + ij\right) - \mathbb{E}_t\left(\frac{1}{1+b\frac{1}{n}\sum_{j=1}^n|\,\mathsf{S}(\mathbf{h}_{j,t+\tau})|}\middle|\mathbf{h}_t, G_t\right) = O\left(\frac{\tau}{n}\right)$. Moreover, similar to above using a Taylor approximation around the mean for large n we have that

$$\mathbb{E}_t\left(\frac{|\,\mathsf{S}(\mathbf{h}_{i,t+\tau})|}{1+b\frac{1}{n}\sum_{j=1}^n|\,\mathsf{S}(\mathbf{h}_{j,t+\tau})|}\middle|\mathbf{h}_t, G_t + ij\right) = \frac{\mathbb{E}_t\left(|\,\mathsf{S}(\mathbf{h}_{i,t+\tau})|\,|\,\mathbf{h}_t, G_t + ij\right)}{1+b\mathbb{E}_t\left(\frac{1}{n}\sum_{j=1}^n|\,\mathsf{S}(\mathbf{h}_{j,t+\tau})|\middle|\mathbf{h}_t, G_t + ij\right)} + O\left(\frac{1}{n}\right),$$

and

$$\frac{\mathbb{E}_t\left(|\,\mathsf{S}(\mathbf{h}_{i,t+\tau})|\,|\,\mathbf{h}_t, G_t + ij\right)}{1+b\mathbb{E}_t\left(\frac{1}{n}\sum_{j=1}^n|\,\mathsf{S}(\mathbf{h}_{j,t+\tau})|\middle|\mathbf{h}_t, G_t + ij\right)} = \frac{f_{it} + \Delta f_{it}\tau + \beta f_{ij,t}\tau + o(\tau)}{1+b(\bar{h}_t + \Delta\bar{h}_t\tau)} + O\left(\frac{\tau}{n}\right).$$

We then have that

$$\begin{aligned}\mathbb{E}_t\left(\frac{|\,\mathsf{S}(\mathbf{h}_{i,t+\tau})|}{1+b\frac{1}{n}\sum_{j=1}^n|\,\mathsf{S}(\mathbf{h}_{j,t+\tau})|}\middle|\mathbf{h}_t, G_t + ij\right) - \mathbb{E}_t\left(\frac{|\,\mathsf{S}(\mathbf{h}_{i,t+\tau})|}{1+b\frac{1}{n}\sum_{j=1}^n|\,\mathsf{S}(\mathbf{h}_{j,t+\tau})|}\middle|\mathbf{h}_t, G_t\right)\\ = \frac{\beta f_{ij,t}\tau}{1+b(\bar{h}_t + \Delta\bar{h}_t\tau)} + O\left(\frac{\tau}{n}\right) = \frac{\beta f_{ij,t}\tau}{1+b\bar{h}_t} + O\left(\frac{\tau}{n}\right)\end{aligned}$$

Next, note that similar to above we an write due to a Taylor approximation around the mean for large n

$$\mathbb{E}_t\left(\frac{|\,\mathsf{S}(\mathbf{h}_{i,t+\tau})|^2}{1+b\frac{1}{n}\sum_{j=1}^n|\,\mathsf{S}(\mathbf{h}_{j,t+\tau})|}\middle|\mathbf{h}_t, G_t + ij\right) = \frac{\mathbb{E}_t\left(|\,\mathsf{S}(\mathbf{h}_{i,t+\tau})|^2\middle|\mathbf{h}_t, G_t + ij\right)}{1+b\mathbb{E}_t\left(\frac{1}{n}\sum_{j=1}^n|\,\mathsf{S}(\mathbf{h}_{j,t+\tau})|\middle|\mathbf{h}_t, G_t + ij\right)} + O\left(\frac{1}{n}\right),$$

and

$$\frac{\mathbb{E}_t\left(|\,\mathsf{S}(\mathbf{h}_{i,t+\tau})|^2\middle|\mathbf{h}_t, G_t + ij\right)}{1+b\mathbb{E}_t\left(\frac{1}{n}\sum_{j=1}^n|\,\mathsf{S}(\mathbf{h}_{j,t+\tau})|\middle|\mathbf{h}_t, G_t + ij\right)} = \frac{f_{it}^2 + 2f_{it}\Delta f_{it}\tau + 2\beta f_{it}f_{ij,t}\tau + o(\tau)}{1+b(\bar{h}_t + \Delta\bar{h}_t\tau)} + O\left(\frac{\tau}{n}\right).$$

[48] We have used the fact that for any two random variables X, Y we can write $\mathbb{E}\left(\frac{X}{Y}\right) = \frac{\mathbb{E}(X)}{\mathbb{E}(Y)} + O\left(\frac{1}{\mathbb{E}(Y)^2}\right)$.

[49] This approximation also becomes more accurate the higher is the average stock of knowledge and the larger is b.

It then follows that

$$\mathbb{E}_t\left(\frac{|\,\mathsf{S}(\mathbf{h}_{i,t+\tau})|^2}{1+b\frac{1}{n}\sum_{j=1}^n|\,\mathsf{S}(\mathbf{h}_{j,t+\tau})|}\middle|\mathbf{h}_t,G_t+ij\right)-\mathbb{E}_t\left(\frac{|\,\mathsf{S}(\mathbf{h}_{i,t+\tau})|^2}{1+b\frac{1}{n}\sum_{j=1}^n|\,\mathsf{S}(\mathbf{h}_{j,t+\tau})|}\middle|\mathbf{h}_t,G_t\right)$$
$$=\frac{2\beta f_{it}f_{ij,t}\tau}{1+b(\bar{h}_t+\Delta\bar{h}_t\tau)}+O\left(\frac{\tau}{n}\right)=\frac{2\beta f_{it}f_{ij,t}\tau}{1+b\bar{h}_t}+O\left(\frac{\tau}{n}\right).$$

The change in the present discounted profit of a firm i from forming the link ij in the competitive case of $\theta=0$ is given by

$$\begin{aligned}
&V_i^\tau(\mathbf{h}_t,G_t+ij)-V_i^\tau(\mathbf{h}_t,G_t)=\\
&\quad\mathbb{E}_t\left(\frac{1}{1+b\frac{1}{n}\sum_{j=1}^n|\,\mathsf{S}(\mathbf{h}_{j,t+\tau})|}\middle|\mathbf{h}_t,G_t+ij\right)-\mathbb{E}_t\left(\frac{1}{1+b\frac{1}{n}\sum_{j=1}^n|\,\mathsf{S}(\mathbf{h}_{j,t+\tau})|}\middle|\mathbf{h}_t,G_t\right)\\
&\quad+\mathbb{E}_t\left(\frac{2b|\,\mathsf{S}(\mathbf{h}_{i,t+\tau})|}{1+b\frac{1}{n}\sum_{j=1}^n|\,\mathsf{S}(\mathbf{h}_{j,t+\tau})|}\middle|\mathbf{h}_t,G_t+ij\right)-\mathbb{E}_t\left(\frac{2b|\,\mathsf{S}(\mathbf{h}_{i,t+\tau})|}{1+b\frac{1}{n}\sum_{j=1}^n|\,\mathsf{S}(\mathbf{h}_{j,t+\tau})|}\middle|\mathbf{h}_t,G_t\right)\\
&\quad+\mathbb{E}_t\left(\frac{b^2|\,\mathsf{S}(\mathbf{h}_{i,t+\tau})|^2}{1+b\frac{1}{n}\sum_{j=1}^n|\,\mathsf{S}(\mathbf{h}_{j,t+\tau})|}\middle|\mathbf{h}_t,G_t+ij\right)-\mathbb{E}_t\left(\frac{b^2|\,\mathsf{S}(\mathbf{h}_{i,t+\tau})|^2}{1+b\frac{1}{n}\sum_{j=1}^n|\,\mathsf{S}(\mathbf{h}_{j,t+\tau})|}\middle|\mathbf{h}_t,G_t\right)+\ o(\tau)-c.
\end{aligned}$$

From the above calculations it follows that for $\theta=0$ the change in the present discounted profit can be written as $V_i^\tau(\mathbf{h}_t,G_t+ij)-V_i^\tau(\mathbf{h}_t,G_t)=\frac{2b\beta(1+bf_{it})f_{ij,t}\tau}{1+b\bar{h}_t}-c+O\left(\frac{\tau}{n}\right)$. Therefore, we can write the change in the present discounted profit of the firm i from forming the link ij for the general case of $\theta\in\{0,1\}$ (for independent or competitive sectors) as follows $V_i^\tau(\mathbf{h}_t,G_t+ij)-V_i^\tau(\mathbf{h}_t,G_t)=\beta\tau\left(\theta+(1-\theta)\frac{2b(1+bf_{it})}{1+b\bar{h}_t}\right)f_{ij,t}-c+O\left(\frac{\tau}{n}\right)$. □

Proof of Theorem 1. Let $n_t(\mathbf{h})$ denote the expected number of firms with technology $\mathbf{h}$, $m_t(\mathbf{h},\mathbf{h}')$ the expected number of links between firms with technologies $\mathbf{h}$ and $\mathbf{h}'$ for $\mathbf{h}\neq\mathbf{h}'$ and $m(\mathbf{h},\mathbf{h})$ being equal to twice the number of links between firms with technology $\mathbf{h}$. Further, let $x_t(\mathbf{h})=\frac{n_t(\mathbf{h})}{n}$ and $z_t(\mathbf{h},\mathbf{h}')=\frac{m_t(\mathbf{h},\mathbf{h}')}{n^2}$. The average number of technologies per firm is given by $\bar{h}_t=\sum_{\mathbf{h}\in\mathcal{H}^N}x_t(\mathbf{h})|\,\mathsf{S}(\mathbf{h})|$.

The expected number $n_t(\mathbf{h})$ of firms with technology $\mathbf{h}$ can increase by in-house R&D through discovering idea k of firms with technology $\mathbf{h}-\mathbf{e}_k$. This happens at a rate $\gamma+\alpha\langle\mathbf{h}-\mathbf{e}_k,\mathbf{u}\rangle$. Moreover, firms with technology $\mathbf{h}-\mathbf{e}_k$ can learn idea k from firms with technology $\mathbf{h}'$ when $h'_k=1$ at a rate $\beta m_t(\mathbf{h}-\mathbf{e}_k,\mathbf{h}')$. Similarly, $n_t(\mathbf{h})$ can decrease if a firm with technology $\mathbf{h}$ discovers a new idea ($h_k=0$) either through in-house R&D at a rate $\gamma+\alpha\langle\mathbf{h},\mathbf{u}\rangle$ or through learning from collaborating firms with technology $\mathbf{h}'$ at a rate $\beta m_t(\mathbf{h},\mathbf{h}')$ for all $\mathbf{h}'\in\mathcal{H}^N$ such that $h'_k=1$. Further, due to exit, $n_t(\mathbf{h})$ decreases at a rate ξ. Conversely, $n_t(\mathbf{0})$ grows at a rate $\chi f^{\vartheta}_{\theta,\kappa}(\bar{h}_t)(n-\sum_{\mathbf{h}\in\mathcal{H}^N}n_t(\mathbf{h}))$, where $n-\sum_{\mathbf{h}\in\mathcal{H}^N}n_t(\mathbf{h})$ counts the number of firms that have exited, and $\sum_{\mathbf{h}\in\mathcal{H}^N}n_t(\mathbf{h})$ is the number of firms that are active.

The expected change in the number $n_t(\mathbf{h})$ of firms with technology $\mathbf{h}$ is then given by

$$\begin{aligned}
n\mathsf{F}_x(\mathbf{h})\equiv&\sum_{k=1}^N\mathbb{1}_{\{h_k=1\}}\left((\gamma+\alpha\langle\mathbf{h}-\mathbf{e}_k,\mathbf{u}\rangle)n_t(\mathbf{h}-\mathbf{e}_k)+\beta\sum_{\mathbf{h}'\in\mathcal{H}^N:h'_k=1}m_t(\mathbf{h}-\mathbf{e}_k,\mathbf{h}')\right)\\
&-\sum_{k=1}^N\mathbb{1}_{\{h_k=0\}}\left((\gamma+\alpha\langle\mathbf{h},\mathbf{u}\rangle)n_t(\mathbf{h})+\beta\sum_{\mathbf{h}'\in\mathcal{H}^N:h'_k=1}m_t(\mathbf{h},\mathbf{h}')\right)\\
&+\lambda\sum_{k=1}^N\mathbb{1}_{\{h_k=0\}}n_t(\mathbf{h}+\mathbf{e}_k)-\lambda\sum_{k=1}^N\mathbb{1}_{\{h_k=1\}}n_t(\mathbf{h})-\xi n_t(\mathbf{h})+\chi f^{\vartheta}_{\theta,\kappa}(\bar{h}_t)\left(n-\sum_{\mathbf{h}'\in\mathcal{H}^N}n_t(\mathbf{h}')\right)\mathbb{1}_{\{\mathbf{h}=\mathbf{0}\}}.
\end{aligned}\tag{28}$$

Dividing by n and using the fact that $x_t(\mathbf{h})=\frac{n_t(\mathbf{h})}{n}$ and $z_t(\mathbf{h},\mathbf{h}')=\frac{m_t(\mathbf{h},\mathbf{h}')}{n^2}$ yields the expected change for

the fraction $x_t(\mathbf{h})$ of firms with technology $\mathbf{h}$ given by

$$\begin{aligned}\mathsf{F}_x(\mathbf{h}) = &\sum_{k=1}^{N} \mathbb{1}_{\{h_k=1\}} \left(\gamma + \alpha\langle \mathbf{h}-\mathbf{e}_k, \mathbf{u}\rangle + n\beta \sum_{\mathbf{h}'\in\mathcal{H}^N : h'_k=1} \frac{z_t(\mathbf{h}-\mathbf{e}_k,\mathbf{h}')}{x_t(\mathbf{h}-\mathbf{e}_k)} \mathbb{1}_{\{x_t(\mathbf{h}-\mathbf{e}_k)>0\}} \right) x_t(\mathbf{h}-\mathbf{e}_k) \\ &- \sum_{k=1}^{N} \mathbb{1}_{\{h_k=0\}} \left(\gamma + \alpha\langle \mathbf{h}, \mathbf{u}\rangle + n\beta \sum_{\mathbf{h}'\in\mathcal{H}^N : h'_k=1} \frac{z_t(\mathbf{h},\mathbf{h}')}{x_t(\mathbf{h})} \mathbb{1}_{\{x_t(\mathbf{h})>0\}} \right) x_t(\mathbf{h}) \\ &+ \lambda \sum_{k=1}^{N} \mathbb{1}_{\{h_k=0\}} x_t(\mathbf{h}+\mathbf{e}_k) - \lambda \sum_{k=1}^{N} \mathbb{1}_{\{h_k=1\}} x_t(\mathbf{h}) - \xi x_t(\mathbf{h}) + \chi f^{\vartheta}_{\theta,\kappa}(\bar{h}_t) \left(1 - \sum_{\mathbf{h}'\in\mathcal{H}^N} x_t(\mathbf{h}') \right) \mathbb{1}_{\{\mathbf{h}=\mathbf{0}\}}. \end{aligned} \tag{29}$$

Note that the probability that a firm with technology vector $\mathbf{h}$ is connected to a firm with technology vector $\mathbf{h}'$ is given by $\mathbb{P}(a_{ij,t}=1|\mathbf{h}_{it}=\mathbf{h},\mathbf{h}_{jt}=\mathbf{h}') = \frac{\mathbb{P}(a_{ij,t}=1,\mathbf{h}_{it}=\mathbf{h},\mathbf{h}_{jt}=\mathbf{h}')}{\mathbb{P}(\mathbf{h}_{it}=\mathbf{h})\mathbb{P}(\mathbf{h}_{jt}=\mathbf{h}')} = \frac{m_t(\mathbf{h},\mathbf{h}')}{n^2}\frac{1}{\frac{n_t(\mathbf{h})}{n}}\frac{1}{\frac{n_t(\mathbf{h}')}{n}} = \frac{m_t(\mathbf{h},\mathbf{h}')}{n_t(\mathbf{h})n_t(\mathbf{h}')}$ $= \frac{z_t(\mathbf{h},\mathbf{h}')}{x_t(\mathbf{h})x_t(\mathbf{h}')} \equiv y_t(\mathbf{h},\mathbf{h}')$, where $z_t(\mathbf{h},\mathbf{h}') = \mathbb{P}(a_{ij,t}=1,\mathbf{h}_{it}=\mathbf{h},\mathbf{h}_{jt}=\mathbf{h}') = \frac{m_t(\mathbf{h},\mathbf{h}')}{n^2}$ and $x_t(\mathbf{h}) \equiv \mathbb{P}(\mathbf{h}_{it}=\mathbf{h}) = \frac{n_t(\mathbf{h})}{n}$. The probability that a randomly selected firm j has technology $\mathbf{h}'$, given that it is connected to a firm i with technology $\mathbf{h}$, is $\mathbb{P}(a_{ij,t}=1,\mathbf{h}_{jt}=\mathbf{h}'|\mathbf{h}_{it}=\mathbf{h}) = \mathbb{P}(a_{ij,t}=1|\mathbf{h}_{it}=\mathbf{h},\mathbf{h}_{jt}=\mathbf{h}')\mathbb{P}(\mathbf{h}_{jt}=\mathbf{h}')$ $= \frac{\mathbb{P}(a_{ij,t}=1,\mathbf{h}_{jt}=\mathbf{h}',\mathbf{h}_{it}=\mathbf{h})}{\mathbb{P}(\mathbf{h}_{it}=\mathbf{h})\mathbb{P}(\mathbf{h}_{jt}=\mathbf{h}')}\mathbb{P}(\mathbf{h}_{jt}=\mathbf{h}') = \frac{\mathbb{P}(a_{ij,t}=1,\mathbf{h}_{jt}=\mathbf{h}',\mathbf{h}_{it}=\mathbf{h})}{\mathbb{P}(\mathbf{h}_{it}=\mathbf{h})} = \frac{z_t(\mathbf{h},\mathbf{h}')}{x_t(\mathbf{h})}$. Let the support of $\mathbf{h}$ be $\mathsf{S}(\mathbf{h})$ and its cardinality $|\,\mathsf{S}(\mathbf{h})| = \langle \mathbf{h},\mathbf{u}\rangle$, counting the number of nonzero entries in $\mathbf{h}$, with $\mathbf{u}$ being a vector of ones. Then we can write Equation (29) as follows

$$\begin{aligned}\mathsf{F}_x(\mathbf{h}) = &\gamma \sum_{k\in\mathsf{S}(\mathbf{h})} x_t(\mathbf{h}-\mathbf{e}_k) - \gamma|\,\mathsf{S}(\mathbf{h}^c)|x_t(\mathbf{h}) + \lambda \sum_{k\in\mathsf{S}(\mathbf{h}^c)} x_t(\mathbf{h}+\mathbf{e}_k) - \lambda|\,\mathsf{S}(\mathbf{h})|x_t(\mathbf{h}) \\ &+ \alpha(|\,\mathsf{S}(\mathbf{h})|-1) \sum_{k\in\mathsf{S}(\mathbf{h})} x_t(\mathbf{h}-\mathbf{e}_k) - \alpha|\,\mathsf{S}(\mathbf{h})||\,\mathsf{S}(\mathbf{h}^c)|x_t(\mathbf{h}) - \xi x_t(\mathbf{h}) + \chi f^{\vartheta}_{\theta,\kappa}(\bar{h}_t) \left(1 - \sum_{\mathbf{h}'\in\mathcal{H}^N} x_t(\mathbf{h}') \right) \mathbb{1}_{\{\mathbf{h}=\mathbf{0}\}} \\ &+ n\beta \sum_{k\in\mathsf{S}(\mathbf{h})} \sum_{\mathbf{h}'\in\mathcal{H}^N : h'_k=1} z_t(\mathbf{h}-\mathbf{e}_k,\mathbf{h}')\mathbb{1}_{\{x_t(\mathbf{h}-\mathbf{e}_k)>0\}} - n\beta \sum_{k\in\mathsf{S}(\mathbf{h}^c)} \sum_{\mathbf{h}'\in\mathcal{H}^N : h'_k=1} z_t(\mathbf{h},\mathbf{h}')\mathbb{1}_{\{x_t(\mathbf{h})>0\}}. \end{aligned} \tag{30}$$

In the following let $n^2\mathsf{F}_z(\mathbf{h},\mathbf{h}')$ denote the expected increment in the number $m_t(\mathbf{h},\mathbf{h}')$ of links between firms with technologies $\mathbf{h}$ and $\mathbf{h}'$. The rate at which links between firms with technologies $\mathbf{h}$ and $\mathbf{h}'$ decay is given by $\rho n_t(\mathbf{h})n_t(\mathbf{h}')\frac{m_t(\mathbf{h},\mathbf{h}')}{n_t(\mathbf{h})n_t(\mathbf{h}')}$, where $n_t(\mathbf{h})n_t(\mathbf{h}')$ is the expected number of pairs of firms with technologies $\mathbf{h}$ and $\mathbf{h}'$ that are selected, and $\frac{m_t(\mathbf{h},\mathbf{h}')}{n_t(\mathbf{h})n_t(\mathbf{h}')}$ is the probability that a link exists between them. Similarly, the rate at which such links are created is given by $\rho n_t(\mathbf{h})n_t(\mathbf{h}')g^{\eta}_{\theta,\tau}(\mathbf{h},\mathbf{h}';\bar{h}_t)\left(1-\frac{m_t(\mathbf{h},\mathbf{h}')}{n_t(\mathbf{h})n_t(\mathbf{h}')}\right)$, where $1-\frac{m_t(\mathbf{h},\mathbf{h}')}{n_t(\mathbf{h})n_t(\mathbf{h}')}$ is the probability that a link does not exist between the firms with technologies $\mathbf{h}$ and $\mathbf{h}'$, and $g^{\eta}_{\theta,\tau}(\mathbf{h},\mathbf{h}';\bar{h}_t)$ is the probability that they want to form a link when they have the opportunity. Collecting these terms, and noting that when the network adjustment is very fast (high ρ) contributions stemming from changes in the technologies $\mathbf{h}$ and $\mathbf{h}'$ or firm exit happen at a rate $o(\rho)$, we get

$$n^2\mathsf{F}_z(\mathbf{h},\mathbf{h}') \equiv \rho\, n_t(\mathbf{h})n_t(\mathbf{h}')g^{\eta}_{\theta,\tau}(\mathbf{h},\mathbf{h}';\bar{h}_t)\left(1-\frac{m_t(\mathbf{h},\mathbf{h}')}{n_t(\mathbf{h})n_t(\mathbf{h}')}\right) - \rho\, m_t(\mathbf{h},\mathbf{h}') + o(\rho). \tag{31}$$

Dividing Equation (31) by n^2 gives

$$\mathsf{F}_z(\mathbf{h},\mathbf{h}') = \rho\, x_t(\mathbf{h})x_t(\mathbf{h}')\left[g^{\eta}_{\theta,\tau}(\mathbf{h},\mathbf{h}';\bar{h}_t)\left(1-\frac{z_t(\mathbf{h},\mathbf{h}')}{x_t(\mathbf{h})x_t(\mathbf{h}')}\right) - \frac{z_t(\mathbf{h},\mathbf{h}')}{x_t(\mathbf{h})x_t(\mathbf{h}')}\right] + o(\rho). \tag{32}$$

We next introduce the vector $P^n(t) = ((x_t(\mathbf{h}))_{\mathbf{h}\in\mathcal{H}^N}, (z_t(\mathbf{h},\mathbf{h}')_{\mathbf{h},\mathbf{h}'\in\mathcal{H}^N})) \in \mathbb{R}_+^{2^N+2^N\times 2^N}$. Moreover, we introduce the random variable $\zeta_P^n = (\zeta_x^n, \zeta_z^n)$, whose distribution describes the stochastic increments of $(P^n(t))_{t\in T}$ from the state P to state z given by $\mathbb{P}\left(\zeta_P^n = z\right) = \mathbb{P}\left(P^n\left(t+\Delta t\right) = P + z|\, P^n\left(t\right) = P\right)$. The increments ζ_x^n describe the change due to innovation or obsolescence, while the increments ζ_z^n correspond to the change due to link formation or decay.

Let $\mathsf{F}_x(\mathbf{h})$ be defined as in Equation (30) and $\mathsf{F}_z(\mathbf{h},\mathbf{h}')$ as in Equation (32). Further, we introduce the functions V_y^N, A_y^N and $A_{y,\delta}^N$ defined by $V_y^n(P) \equiv \lambda_y^n\mathbb{E}[\zeta_y^n]$, $A_y^n(P) \equiv \lambda_y^n\mathbb{E}[|\zeta_y^n|]$, and $A_{y,\delta}^n(P) \equiv \lambda_y^n\mathbb{E}[\left|\zeta_y^n I_{\{|\zeta_y^n|>\delta\}}\right|]$, with $y\in\{x,z\}$. The jump rate at which innovations happen is given by $\lambda_x^n = n$ while the jump rate of link changes is given by $\lambda_z^n = n^2$. Observe that $V^n(P) = (V_x^n(P), V_z^n(P))$ is the expected increment of $\zeta_{y\in\{x,z\}}^n$

for a short time interval $[t, t+\Delta t)$. Consider some sequence $(\delta^n)_{n=n_0}^{\infty}$ with $\lim_{n\to\infty}\delta^n = 0$. In the following we want to show that the following three conditions hold: (i) $\lim_{n\to\infty}\sup_{P\in P^n}\left|V_y^n(P) - V_y(P)\right| = 0$, (ii) $\sup_n \sup_{P\in P^n} A_y^n(P) < \infty$, and (iii) $\lim_{n\to\infty}\sup_{P\in P^n} A_{y,\delta^n}^n(P) = 0$. First, consider $y = x$. Let $e_{\mathbf{h}}$ be the standard unit basis vector corresponding to technology $\mathbf{h}$. Observe that $V_x^n(P) = n\mathbb{E}[\zeta_x^n] = n\sum_{\mathbf{h},\mathbf{h}'}\frac{1}{n}(e_{\mathbf{h}'} - e_{\mathbf{h}})\mathbb{P}\left(\zeta_x^n = \frac{1}{n}(e_{\mathbf{h}'} - e_{\mathbf{h}})\right) = \sum_{\mathbf{h}} e_{\mathbf{h}}\mathsf{F}_x(\mathbf{h}) = V_x(P)$, which is independent of n assuming that β is propositional to $1/n$. This implies that condition (i) is satisfied. Further, observe that since $|e_{\mathbf{h}'} - e_{\mathbf{h}}| = \sqrt{2}$ for $\mathbf{h} \neq \mathbf{h}'$ and 0 otherwise, $(P^n(t))_{t\in T}$ has jumps of at most $\sqrt{2}/n$. Hence, for $\delta^n = \sqrt{2}/n$ it follows that $A_{x,\delta^n}^n(P) = n\mathbb{E}\left[\left|\zeta_x^n I_{\{|\zeta_x^n|>\sqrt{2}/n\}}\right|\right] = 0$, and condition (iii) holds. Finally, we find that $A_x^n(P) = n\mathbb{E}[|\zeta_x^n|] \leq n\frac{\sqrt{2}}{n} = \sqrt{2} < \infty$, and also condition (ii) is satisfied. Next, consider the case of $y = z$. Let $e_{\mathbf{h},\mathbf{h}'}$ be the standard unit basis vector indicating a link between a firm with technology $\mathbf{h}$ and a firm with technology $\mathbf{h}'$. First, observe that $V_z^n(P) = n^2\mathbb{E}[\zeta_z^n] = n^2\sum_{\mathbf{h}'',\mathbf{h}'''}\frac{1}{n^2}(e_{\mathbf{h}'',\mathbf{h}'''} - e_{\mathbf{h},\mathbf{h}'})\mathbb{P}\left(\zeta_z^n = \frac{1}{n^2}(e_{\mathbf{h}'',\mathbf{h}'''} - e_{\mathbf{h},\mathbf{h}'})\right) = \sum_{\mathbf{h},\mathbf{h}'} e_{\mathbf{h},\mathbf{h}'}\mathsf{F}_z(\mathbf{h},\mathbf{h}) = V_z(P)$, which is independent of n. This implies that condition (i) is satisfied. Further, observe that since $|e_{\mathbf{h}'',\mathbf{h}'''} - e_{\mathbf{h},\mathbf{h}'}| = \sqrt{2}$ for $(\mathbf{h}'',\mathbf{h}''') \neq (\mathbf{h},\mathbf{h}')$ and 0 otherwise, $(P^n(t))_{t\in T}$ has jumps of at most $\sqrt{2}/n^2$. Hence, for $\delta^n = \sqrt{2}/n^2$ we have that $A_{z,\delta^n}^n(P) = n^2\mathbb{E}\left[\left|\zeta_z^n I_{\{|\zeta_z^n|>\sqrt{2}/n^2\}}\right|\right] = 0$, and condition (iii) holds. Moreover, we find that $A_z^n(P) = n\mathbb{E}[|\zeta_z^n|] \leq n^2\frac{\sqrt{2}}{n^2} = \sqrt{2} < \infty$, and also condition (ii) is satisfied. Finally, observe that $V(P)$ is a Lipschitz continuous vector field in P, as both $\mathsf{F}_x(\cdot)$ and $\mathsf{F}_z(\cdot,\cdot)$ are linear functions of x and z in the limit of large ρ, and hence have bounded derivatives. Together with conditions (i), (ii) and (iii), we then can apply Kurtz's Theorem [cf. Sandholm, 2010, Chap.10.2], which states that for any solution $\{P(t)\}_{t\in T}$ of the mean-field dynamics

$$\frac{dP}{dt} = \frac{d}{dt}\begin{bmatrix}\vdots\\ x_t(\mathbf{h})\\ \vdots\\ z_t(\mathbf{h},\mathbf{h}')\\ \vdots\end{bmatrix} = V(P) = \begin{bmatrix}V_x(P)\\ V_z(P)\end{bmatrix} = \begin{bmatrix}\vdots\\ \mathsf{F}_x(\mathbf{h})\\ \vdots\\ \mathsf{F}_z(\mathbf{h},\mathbf{h}')\\ \vdots\end{bmatrix} \tag{33}$$

starting from P_0 we have that $\lim_{n\to\infty}\mathbb{P}\left(\sup_{t\in[0,T]}|P^n(t) - P(t)| \geq \epsilon\right) = 0$, for any $T < \infty$ and $\epsilon > 0$. In particular, in the limit of large n we then can write $\operatorname*{plim}_{n\to\infty}\frac{dx_t(\mathbf{h})}{dt} = \mathsf{F}_x(\mathbf{h})$, and $\operatorname*{plim}_{n\to\infty}\frac{dz_t(\mathbf{h},\mathbf{h}')}{dt} = \mathsf{F}_z(\mathbf{h},\mathbf{h}')$. Next, if we introduce the variable $y_t(\mathbf{h},\mathbf{h}') \equiv \mathbb{P}(a_{ij,t} = 1|\mathbf{h}_{it} = \mathbf{h}, \mathbf{h}_{jt} = \mathbf{h}') = \frac{z_t(\mathbf{h},\mathbf{h}')}{x_t(\mathbf{h})x_t(\mathbf{h}')}$, then we have that

$$\frac{dy_t(\mathbf{h},\mathbf{h}')}{dt} = \frac{1}{x_t(\mathbf{h})x_t(\mathbf{h}')}\frac{dz_t(\mathbf{h},\mathbf{h}')}{dt} - \frac{z_t(\mathbf{h},\mathbf{h}')}{x_t(\mathbf{h})x_t(\mathbf{h}')}\left(\frac{1}{x_t(\mathbf{h})}\frac{dx_t(\mathbf{h})}{dt} + \frac{1}{x_t(\mathbf{h}')}\frac{dx_t(\mathbf{h}')}{dt}\right).$$

Inserting Equation (32) gives

$$\begin{aligned}\frac{dy_t(\mathbf{h},\mathbf{h}')}{dt} &= \rho g_{\theta,\tau}^{\eta}(\mathbf{h},\mathbf{h}';\bar{h}_t)\left(1 - y_t(\mathbf{h},\mathbf{h}')\right) - \rho\, y_t(\mathbf{h},\mathbf{h}')\\ &\quad + y_t(\mathbf{h},\mathbf{h}')\left(\frac{1}{x_t(\mathbf{h})}\frac{dx_t(\mathbf{h})}{dt} + \frac{1}{x_t(\mathbf{h}')}\frac{dx_t(\mathbf{h}')}{dt}\right)\\ &\quad - y_t(\mathbf{h},\mathbf{h}')\left(\frac{1}{x_t(\mathbf{h})}\frac{dx_t(\mathbf{h})}{dt} + \frac{1}{x_t(\mathbf{h}')}\frac{dx_t(\mathbf{h}')}{dt}\right) + o(\rho)\\ &= \rho\, g_{\theta,\tau}^{\eta}(\mathbf{h},\mathbf{h}';\bar{h}_t) - \rho\left(1 + g_{\theta,\tau}^{\eta}(\mathbf{h},\mathbf{h}';\bar{h}_t)\right)y_t(\mathbf{h},\mathbf{h}') + o(\rho).\end{aligned} \tag{34}$$

Moreover, inserting the definition of $y_t(\mathbf{h},\mathbf{h}') \equiv \frac{z_t(\mathbf{h},\mathbf{h}')}{x_t(\mathbf{h})x_t(\mathbf{h}')}$ into Equation (30) we get

$$\begin{aligned}\frac{dx_t(\mathbf{h})}{dt} &= \gamma\sum_{k\in\mathsf{S}(\mathbf{h})} x_t(\mathbf{h}-\mathbf{e}_k) - (\lambda|\,\mathsf{S}(\mathbf{h})| + \gamma|\,\mathsf{S}(\mathbf{h}^c)|)\,x_t(\mathbf{h}) + \lambda\sum_{k\in\mathsf{S}(\mathbf{h}^c)} x_t(\mathbf{h}+\mathbf{e}_k) - \alpha|\,\mathsf{S}(\mathbf{h})||\,\mathsf{S}(\mathbf{h}^c)|x_t(\mathbf{h})\\ &\quad + \alpha(|\,\mathsf{S}(\mathbf{h})| - 1)\sum_{k\in\mathsf{S}(\mathbf{h})} x_t(\mathbf{h}-\mathbf{e}_k) - \xi x_t(\mathbf{h}) + \chi f_{\theta,\kappa}^{\vartheta}(\bar{h}_t)\left(1 - \sum_{\mathbf{h}'\in\mathcal{H}^N} x_t(\mathbf{h}')\right)\mathbb{1}_{\{\mathbf{h}=\mathbf{0}\}}\\ &\quad + n\beta\sum_{k\in\mathsf{S}(\mathbf{h})}\sum_{\mathbf{h}'\in\mathcal{H}^N:h'_k=1} y_t(\mathbf{h}-\mathbf{e}_k,\mathbf{h}')x_t(\mathbf{h}-\mathbf{e}_k)x_t(\mathbf{h}') - n\beta\sum_{k\in\mathsf{S}(\mathbf{h}^c)}\sum_{\mathbf{h}'\in\mathcal{H}^N:h'_k=1} y_t(\mathbf{h},\mathbf{h}')x_t(\mathbf{h})x_t(\mathbf{h}').\end{aligned} \tag{35}$$

This concludes the proof. □

The following lemma derives the dynamics for the fraction $\bar{x}_t(s)$ of firms with knowledge stock s defined in Equation (13), and the probability $\bar{y}_t(s,s')$ of a link between firms with knowledge stocks s and s', respectively, defined in Equation (14). This lemma will be crucial for the proof of Proposition 2 that follows.

Lemma 1. *Let the fraction of firms with a stock of knowledge of s be denoted by $\bar{x}_t(s)$ and let the probability of a link between a firm with knowledge stock s and a firm with s' be $\bar{y}_t(s,s')$ for any $0 \leq s, s' \leq N$ defined as in Equations (13) and (14). Then $\bar{x}_t(s)$ is the solution of the system of ODEs*

$$\begin{aligned}\frac{d\bar{x}_t(s)}{dt} &= (\gamma s + \alpha(s-1)s)\bar{x}_t(s-1) - (\lambda s + \gamma(N-s) + \alpha s(N-s))\bar{x}_t(s) + \lambda(N-s)\bar{x}_t(s+1) \\ &+ n\beta \sum_{s'=1}^{N} \binom{N-1}{s'-1} (s\bar{y}_t(s-1,s')\bar{x}_t(s-1)\bar{x}_t(s') - (N-s)\bar{y}_t(s,s')\bar{x}_t(s)\bar{x}_t(s')) \\ &- \xi\bar{x}_t(s) + \chi f^{\vartheta}_{\theta,\kappa}(\bar{h}_t)\left(1 - \sum_{s'=0}^{N}\binom{N}{s'}\bar{x}_t(s')\right)\mathbb{1}_{\{s=0\}},\end{aligned} \tag{36}$$

where $f^{\vartheta}_{\theta,\kappa}(\bar{h}_t)$ has been defined in Equation (9), and $\bar{y}_t(s,s')$ is the solution of the system of ODEs

$$\frac{d\bar{y}_t(s,s')}{dt} = \rho\bar{g}^{\eta}_{\theta,\tau}(s,s';\bar{h}_t) - \rho(1+\bar{g}^{\eta}_{\theta,\tau}(s,s';\bar{h}_t))\bar{y}_t(s,s') + o(\rho), \tag{37}$$

where $\bar{g}^{\eta}_{\theta,\tau}(s,s';\bar{h}_t))$ has been defined in Equation (15).

Proof of Lemma 1. Let us define $\tilde{x}_t(s) \equiv \sum_{\mathbf{h}\in\mathcal{H}^N:|\mathsf{S}(\mathbf{h})|=s} x_t(\mathbf{h})$, for $0 \leq s \leq N$. Summation over all $\mathbf{h} \in \mathcal{H}^N$ with the property that $|\mathsf{S}(\mathbf{h})| = s$ in Equation (30) gives

$$\begin{aligned}\frac{d\tilde{x}_t(s)}{dt} &= \gamma(N-s+1)\tilde{x}_t(s-1) - (\lambda s + \gamma(N-s))\tilde{x}_t(s) + \lambda(s+1)\tilde{x}_t(s+1) - \alpha s(N-s)\tilde{x}_t(s) \\ &+ \alpha(s-1)(N-s+1)\tilde{x}_t(s-1) - \xi\tilde{x}_t(s) + \chi f^{\vartheta}_{\theta,\kappa}(\bar{h}_t)\left(1 - \sum_{s'=0}^{N}\tilde{x}_t(s')\right)\mathbb{1}_{\{s=0\}} \\ &+ n\beta \sum_{\mathbf{h}\in\mathcal{H}^N:|\mathsf{S}(\mathbf{h})|=s}\ \sum_{k\in\mathsf{S}(\mathbf{h})}\ \sum_{\mathbf{h}'\in\mathcal{H}^N:h'_k=1} z_t(\mathbf{h}-\mathbf{e}_k,\mathbf{h}') - n\beta \sum_{\mathbf{h}\in\mathcal{H}^N:|\mathsf{S}(\mathbf{h})|=s}\ \sum_{k\in\mathsf{S}(\mathbf{h}^c)}\ \sum_{\mathbf{h}'\in\mathcal{H}^N:h'_k=1} z_t(\mathbf{h},\mathbf{h}'),\end{aligned} \tag{38}$$

where we have used the fact that $\sum_{\mathbf{h}\in\mathcal{H}^N} x_t(\mathbf{h}) = \sum_{s=0}^{N}\sum_{\mathbf{h}\in\mathcal{H}^N:|\mathsf{S}(\mathbf{h})|=s} x_t(\mathbf{h}) = \sum_{s=0}^{N}\tilde{x}_t(s)$, and $\sum_{\mathbf{h}\in\mathcal{H}^N:|\mathsf{S}(\mathbf{h})|=s}$ $\sum_{k\in\mathsf{S}(\mathbf{h})} x_t(\mathbf{h}-\mathbf{e}_k) = \sum_{\mathbf{h}\in\mathcal{H}^N:|\mathsf{S}(\mathbf{h})|=s-1}\sum_{k\in\mathsf{S}(\mathbf{h}^c)} x_t(\mathbf{h}) = (N-s+1)\sum_{\mathbf{h}\in\mathcal{H}^N:|\mathsf{S}(\mathbf{h})|=s-1} x_t(\mathbf{h}) = (N-s+1)\tilde{x}_t(s-1)$, and $\sum_{\mathbf{h}\in\mathcal{H}^N:|\mathsf{S}(\mathbf{h})|=s}\sum_{k\in\mathsf{S}(\mathbf{h}^c)} x_t(\mathbf{h}+\mathbf{e}_k) = \sum_{\mathbf{h}\in\mathcal{H}^N:|\mathsf{S}(\mathbf{h})|=s+1}\sum_{k\in\mathsf{S}(\mathbf{h})} x_t(\mathbf{h}) = (s+1)$ $\sum_{\mathbf{h}\in\mathcal{H}^N:|\mathsf{S}(\mathbf{h})|=s+1} x_t(\mathbf{h}) = (s+1)\tilde{x}_t(s+1)$. Further, define $\tilde{z}_t(s,s') \equiv \sum_{\mathbf{h}\in\mathcal{H}^N:|\mathsf{S}(\mathbf{h})|=s,\mathbf{h}'\in\mathcal{H}^N:|\mathsf{S}(\mathbf{h}')|=s'} z_t(\mathbf{h},\mathbf{h}')$, for $0 \leq s, s' \leq N$. With $\sum_{\mathbf{h}\in\mathcal{H}^N:|\mathsf{S}(\mathbf{h})|=s}\sum_{k\in\mathsf{S}(\mathbf{h})}\sum_{\mathbf{h}'\in\mathcal{H}^N:h'_k=1} z_t(\mathbf{h}-\mathbf{e}_k,\mathbf{h}') = (N-s+1)\sum_{s'=1}^{N}\frac{s'}{N-s'+1}\tilde{z}_t(s-1,s')$, and similarly $\sum_{\mathbf{h}\in\mathcal{H}^N:|\mathsf{S}(\mathbf{h})|=s}\sum_{k\in\mathsf{S}(\mathbf{h}^c)}\sum_{\mathbf{h}'\in\mathcal{H}^N:h'_k=1} z_t(\mathbf{h},\mathbf{h}') = (N-s+1)\sum_{s'=1}^{N}\frac{s'}{N-s'+1}\tilde{z}_t(s,s')$,[50] we obtain

$$\begin{aligned}\frac{d\tilde{x}_t(s)}{dt} &= \gamma(N-s+1)\tilde{x}_t(s-1) - (\lambda s + \gamma(N-s))\tilde{x}_t(s) + \lambda(s+1)\tilde{x}_t(s+1) - \alpha s(N-s)\tilde{x}_t(s) \\ &+ \alpha(s-1)(N-s+1)\tilde{x}_t(s-1) - \xi\tilde{x}_t(s) + \chi f^{\vartheta}_{\theta,\kappa}(\bar{h}_t)\left(1 - \sum_{s'=0}^{N}\tilde{x}_t(s')\right)\mathbb{1}_{\{s=0\}} \\ &+ n\beta(N-s+1)\sum_{s'=1}^{N}\frac{s'}{N-s'+1}\tilde{z}_t(s-1,s')\mathbb{1}_{\{\tilde{x}_t(s-1)>0\}} - n\beta(N-s+1)\sum_{s'=1}^{N}\frac{s'}{N-s'+1}\tilde{z}_t(s,s')\mathbb{1}_{\{\tilde{x}_t(s)>0\}}.\end{aligned} \tag{39}$$

[50] This is because of the exchangeability (symmetry) of the coordinates in the technology vector $\mathbf{h}$. If the initial conditions are such that $x_0(\mathbf{h})$, $z_0(\mathbf{h},\mathbf{h}')$ and $y_0(\mathbf{h},\mathbf{h}')$ depend only on the knowledge stocks $|\mathsf{S}(\mathbf{h})|$ and $|\mathsf{S}(\mathbf{h}')|$, respectively, then this holds for all later times $t > 0$.

We can write Equation (39) more compactly as follows

$$\begin{aligned}\frac{d\tilde{x}_t(s)}{dt} &= (N-s+1)(\gamma+\alpha(s-1))\tilde{x}_t(s-1) - (\lambda s+\gamma(N-s)+\alpha s(N-s))\tilde{x}_t(s) + \lambda(s+1)\tilde{x}_t(s+1)\\ &+ n\beta(N-s+1)\sum_{s'=1}^{N}\frac{s'}{N-s'+1}(\tilde{z}_t(s-1,s')\mathbb{1}_{\{\tilde{x}_t(s-1)>0\}} - \tilde{z}_t(s,s')\mathbb{1}_{\{\tilde{x}_t(s)>0\}})\\ &- \xi\tilde{x}_t(s) + \chi f^{\vartheta}_{\theta,\kappa}(\bar{h}_t)\left(1-\sum_{s'=0}^{N}\tilde{x}_t(s')\right)\mathbb{1}_{\{s=0\}}. \end{aligned} \tag{40}$$

We next introduce the average $\bar{x}_t(s) \equiv \frac{1}{\binom{N}{s}}\tilde{x}_t(s) = \frac{1}{\binom{N}{s}}\sum_{\mathbf{h}\in\mathcal{H}^N:|\mathsf{S}(\mathbf{h})|=s} x_t(\mathbf{h})$, for $0 \le s \le N$, and similarly let $\bar{z}_t(s,s') \equiv \frac{1}{\binom{N}{s}}\frac{1}{\binom{N}{s'}}\sum_{\mathbf{h}\in\mathcal{H}^N:|\mathsf{S}(\mathbf{h})|=s,\mathbf{h}'\in\mathcal{H}^N:|\mathsf{S}(\mathbf{h}')|=s'} z_t(\mathbf{h},\mathbf{h}') = \frac{1}{\binom{N}{s}}\frac{1}{\binom{N}{s'}}\tilde{z}_t(s,s')$, for $0 \le s,s' \le N$. Then, using the fact that $\frac{d\bar{x}_t(s)}{dt} = \frac{1}{\binom{N}{s}}\frac{d\bar{x}_t(s)}{dt} = \frac{1}{\binom{N}{s}}\sum_{\mathbf{h}\in\mathcal{H}^N:|\mathsf{S}(\mathbf{h})|=s}\frac{dx_t(\mathbf{h})}{dt}$, where $\binom{N}{s}$ counts the number of vectors $\mathbf{h}\in\mathcal{H}^N$ with $|\mathsf{S}(\mathbf{h})| = s$, we obtain from Equation (40)

$$\begin{aligned}\frac{d\bar{x}_t(s)}{dt} &= (\gamma s+\alpha(s-1)s)\bar{x}_t(s-1) - (\lambda s+\gamma(N-s)+\alpha s(N-s))\bar{x}_t(s) + \lambda(N-s)\bar{x}_t(s+1)\\ &+ n\beta\sum_{s'=1}^{N}\binom{N-1}{s'-1}(s\bar{z}_t(s-1,s')\mathbb{1}_{\{\bar{x}_t(s-1)>0\}} - (N-s)\bar{z}_t(s,s')\mathbb{1}_{\{\bar{x}_t(s)>0\}})\\ &- \xi\bar{x}_t(s) + \chi f^{\vartheta}_{\theta,\kappa}(\bar{h}_t)\left(1-\sum_{s'=0}^{N}\binom{N}{s'}\bar{x}_t(s')\right)\mathbb{1}_{\{s=0\}}. \end{aligned} \tag{41}$$

We further define

$$\bar{y}_t(s,s') \equiv \frac{1}{\binom{N}{s}}\frac{1}{\binom{N}{s'}}\sum_{\substack{\mathbf{h}\in\mathcal{H}^N:|\mathsf{S}(\mathbf{h})|=s\\ \mathbf{h}'\in\mathcal{H}^N:|\mathsf{S}(\mathbf{h}')|=s'}}\frac{z_t(\mathbf{h},\mathbf{h}')}{x_t(\mathbf{h})x_t(\mathbf{h}')} = \frac{\bar{z}_t(s,s')}{\bar{x}_t(s)\bar{x}_t(s')}, \tag{42}$$

where we have used the exchangeability (symmetry) of the coordinates in the technology vector $\mathbf{h}$, we can write Equation (41) as follows

$$\begin{aligned}\frac{d\bar{x}_t(s)}{dt} &= (\gamma s+\alpha(s-1)s)\bar{x}_t(s-1) - (\lambda s+\gamma(N-s)+\alpha s(N-s))\bar{x}_t(s) + \lambda(N-s)\bar{x}_t(s+1)\\ &+ n\beta\sum_{s'=1}^{N}\binom{N-1}{s'-1}(s\bar{y}_t(s-1,s')\bar{x}_t(s-1)\bar{x}_t(s') - (N-s)\bar{y}_t(s,s')\bar{x}_t(s)\bar{x}_t(s'))\\ &- \xi\bar{x}_t(s) + \chi f^{\vartheta}_{\theta,\kappa}(\bar{h}_t)\left(1-\sum_{s'=0}^{N}\binom{N}{s'}\bar{x}_t(s')\right)\mathbb{1}_{\{s=0\}}. \end{aligned} \tag{43}$$

Summation over all $\mathbf{h}\in\mathcal{H}^N$ with the property that $|\mathsf{S}(\mathbf{h})| = s$ and $\mathbf{h}'\in\mathcal{H}^N$ with $\mathsf{S}(\mathbf{h}')| = s'$ and inserting the definition in Equation (42) into Equation (34) gives

$$\frac{d\bar{y}_t(s,s')}{dt} = \rho\bar{g}^{\eta}_{\theta,\tau}(s,s';\bar{h}_t) - \rho(1+\bar{g}^{\eta}_{\theta,\tau}(s,s';\bar{h}_t))\bar{y}_t(s,s') + o(\rho), \tag{44}$$

where we have denoted by

$$\begin{aligned}\bar{g}^{\eta}_{\theta,\tau}(s,s';\bar{h}_t) &\equiv \frac{1}{\binom{N}{s}}\frac{1}{\binom{N}{s'}}\tilde{g}^{\eta}_{\theta,\tau}(s,s';\bar{h}_t) = \frac{1}{\binom{N}{s}}\frac{1}{\binom{N}{s'}}\sum_{\substack{\mathbf{h}\in\mathcal{H}^N:|\mathsf{S}(\mathbf{h})|=s\\ \mathbf{h}'\in\mathcal{H}^N:|\mathsf{S}(\mathbf{h}')|=s'}} g^{\eta}_{\theta,\tau}(\mathbf{h},\mathbf{h}';\bar{h}_t)\\ &= \frac{1}{\binom{N}{s}}\frac{1}{\binom{N}{s'}}\sum_{k=\max\{0,s'-s\}}^{\min\{N-s,s'\}}\binom{N-s}{k}\binom{s}{s'-k}\binom{N}{N-s}\\ &\times\frac{e^{\eta(\beta g_{\theta,\tau}(\bar{h}_t)(1+bs)^{1-\theta}k-c)}}{1+e^{\eta(\beta g_{\theta,\tau}(\bar{h}_t)(1+bs)^{1-\theta}k-c)}}\frac{e^{\eta(\beta g_{\theta,\tau}(\bar{h}_t)(1+bs')^{1-\theta}(s-s'+k)-c)}}{1+e^{\eta(\beta g_{\theta,\tau}(\bar{h}_t)(1+bs')^{1-\theta}(s-s'+k)-c)}}., \end{aligned} \tag{45}$$

which is a symmetric matrix, $\bar{g}^{\eta}_{\theta,\tau}(s,s';\bar{h}_t) = \bar{g}^{\eta}_{\theta,\tau}(s',s;\bar{h}_t)$ for all $0 \le s,s' \le N$. This completes the proof. □

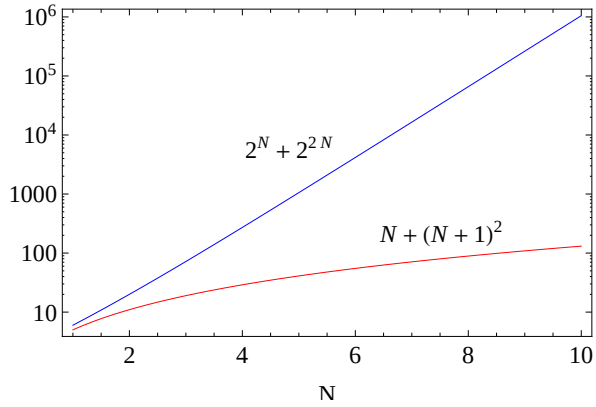


Figure C.1: A comparison of the number of equations that need to be considered for the technology portfolios, $\mathbf{h} \in \mathcal{H}^N$ and the knowledge stocks, $|\mathsf{S}(\mathbf{h})| \in \{0, \ldots, N\}$.

Note that for the dynamics of the technology portfolios, $\mathbf{h} \in \mathcal{H}^N$, in Theorem 1 we needed to solve $2^N + 2^{2N}$ equations for $x_t(\mathbf{h})$ and $y_t(\mathbf{h}, \mathbf{h}')$. In contrast, when we consider the dynamics of the knowledge stocks, $s = 0, \ldots, N$, as in Lemma 1, we need to solve only $N + (N+1)^2$ equations for $\bar{x}_t(s)$ and $\bar{y}_t(s, s')$. A comparison of the number of equations that need to be considered is shown in Figure C.1 for increasing values of N.

In what follows we study the stationary states of the dynamic system introduced in Lemma 1 and their stability properties.

Proof of Proposition 2. We first consider part (i) of the propostion. It follows immediately that the fixed points for $\bar{y}(s, s')$ of Equation (37) are given by $\bar{y}(s, s') = \frac{\bar{g}(s,s')}{1+\bar{g}(s,s')}$. Moreover, the fixed points from Equation (36) satisfy the following recursive equation

$$\left(\lambda s + (N-s)(\gamma + \alpha s) + n\beta(N-s)\sum_{s'=1}^{N}\binom{N-1}{s'-1}\bar{y}(s,s')\bar{x}(s')\right)\bar{x}(s)$$
$$= \left(s(\gamma + \alpha(s-1)) + n\beta s\sum_{s'=1}^{N}\binom{N-1}{s'-1}\bar{y}(s-1,s')\bar{x}(s')\right)\bar{x}(s-1) + \lambda(N-s)\bar{x}(s+1).$$

For $s = 0$ we obtain

$$\left(\gamma + n\beta\sum_{s'=1}^{N}\binom{N-1}{s'-1}\bar{y}(0,s')\bar{x}(s')\right)\bar{x}(0) = \lambda\bar{x}(1). \tag{46}$$

Similarly, for $s = 1$ we get

$$\left(\lambda + (N-1)(\gamma + \alpha) + n\beta(N-1)\sum_{s'=1}^{N}\binom{N-1}{s'-1}\bar{y}(1,s')\bar{x}(s')\right)\bar{x}(1)$$
$$= \left(\gamma + n\beta\sum_{s'=1}^{N}\binom{N-1}{s'-1}\bar{y}(0,s')\bar{x}(s')\right)\bar{x}(0) + \lambda(N-1)\bar{x}(2).$$

Using Equation (46), this can be written as

$$\left(\gamma + \alpha + n\beta\sum_{s'=1}^{N}\binom{N-1}{s'-1}\bar{y}(1,s')\bar{x}(s')\right)\bar{x}(1) = \lambda\bar{x}(2).$$

One can show by induction that for general $s \geq 0$ the following recursive equation holds

$$\left(\gamma + s\alpha + n\beta\sum_{s'=1}^{N}\binom{N-1}{s'-1}\bar{y}(s,s')\bar{x}(s')\right)\bar{x}(s) = \lambda\bar{x}(s+1).$$

Hence, we get

$$\bar{x}(s+1) = \frac{\bar{x}(0)}{\lambda^{s+1}} \prod_{k=0}^{s} \left(\gamma + k\alpha + n\beta \sum_{s'=1}^{N} \binom{N-1}{s'-1} \bar{y}(k,s')\bar{x}(s') \right). \tag{47}$$

We next consider part (iii) of the proposition. In the case of $\beta = 0$ we obtain from Equation (47) $\bar{x}(s+1) = \frac{\gamma+\alpha s}{\lambda}\bar{x}(s) = \bar{x}(0)\prod_{k=0}^{s}\frac{\gamma+\alpha k}{\lambda}$. From the normalization condition $\sum_{s=0}^{N}\binom{N}{s}\bar{x}(s) = 1$, we find that $\bar{x}(0) = \left(\sum_{s=0}^{N}\binom{N}{s}\prod_{k=0}^{s-1}\frac{\gamma+\alpha k}{\lambda}\right)^{-1}$, and inserting yields $\bar{x}(s) = \left(\sum_{s'=0}^{N}\binom{N}{s'}\prod_{k=0}^{s'-1}\frac{\gamma+\alpha k}{\lambda}\right)^{-1}\prod_{k=0}^{s-1}\frac{\gamma+\alpha k}{\lambda}$. When also $\alpha = 0$ then we get from above $\bar{x}(s) = \left(\sum_{s'=0}^{N}\binom{N}{s'}\left(\frac{\gamma}{\lambda}\right)^{s'}\right)^{-1}\left(\frac{\gamma}{\lambda}\right)^{s}$. Using the fact that $\sum_{s=0}^{N}\binom{N}{s}\left(\frac{\gamma}{\lambda}\right)^{s} = \left(\frac{\gamma+\lambda}{\lambda}\right)^{N}$, we get $\bar{x}(s) = \left(\frac{\lambda}{\gamma+\lambda}\right)^{N}\left(\frac{\gamma}{\lambda}\right)^{s} = \left(\frac{\gamma}{\lambda+\gamma}\right)^{s}\left(\frac{\lambda}{\lambda+\gamma}\right)^{N-s}$. This is a Binomial distribution with success probability $\frac{\gamma}{\lambda+\gamma}$. Next, let $\tilde{x}(s) = \binom{N}{s}\bar{x}(s)$. Then we can write $\lambda\tilde{x}(s+1) = (\gamma+\alpha s)\frac{\binom{N}{s+1}}{\binom{N}{s}}\tilde{x}(s)$. Using the fact that for large N, $\frac{\binom{N}{s+1}}{\binom{N}{s}} \sim \frac{1}{s+1}$, we then get $\lambda(s+1)\tilde{x}(s+1) = (\gamma+\alpha s)\,\tilde{x}(s)$. Up to a multiplicative constant, the solution is given by $\tilde{x}(s+1) \propto \left(\frac{\alpha}{\lambda}\right)^{s-1}\frac{\Gamma\left(\frac{\gamma}{\alpha}+s\right)}{1+s} \sim \left(\frac{\alpha}{\lambda}\right)^{s-1} s^{-\left(1-\frac{\gamma}{\alpha}\right)}$, where we have taken Stirling's approximation to the Gamma functions in the last step.[51]

Next we consider part (iv) of the proposition. From Equation (47) we have that in the stationary state

$$\bar{x}(s+1) = \bar{x}(0)\left(\frac{\gamma}{\lambda}\right)^{s+1}\prod_{k=0}^{s}\left(1 + k\frac{\alpha}{\gamma} + \frac{n\beta}{\gamma}\sum_{s'=1}^{N}\binom{N-1}{s'-1}\bar{y}(k,s')\bar{x}(s')\right).$$

In the following we denote by $\tilde{\alpha} \equiv \frac{\alpha}{\gamma}$, $\tilde{\beta} \equiv \frac{n\beta}{\gamma}$, $\tilde{\gamma} \equiv \frac{\gamma}{\lambda}$ and $\hat{y}(s,s') \equiv \binom{N-1}{s'-1}\bar{y}(s,s')$, so that we can write

$$\bar{x}(s+1) = \bar{x}(0)\tilde{\gamma}^{s+1}\prod_{k=0}^{s}\left(1 + k\tilde{\alpha} + \tilde{\beta}\sum_{s'=1}^{N}\hat{y}(k,s')\bar{x}(s')\right).$$

Using the Taylor expansion $\prod_{k=0}^{s}(a_k + \beta b_k) = \prod_{k=0}^{s} a_k + \beta\sum_{k=0}^{s}\prod_{l\neq k}^{s} a_l b_k + O(\tilde{\beta}^2)$, we get

$$\prod_{k=0}^{s}\left(1 + k\tilde{\alpha} + \tilde{\beta}\sum_{s'=1}^{N}\binom{N-1}{s'-1}\bar{y}(k,s')\bar{x}(s')\right) = \prod_{k=0}^{s}(1+\tilde{\alpha}k) + \tilde{\beta}\sum_{k=0}^{s}\prod_{l\neq k}^{s}(1+\tilde{\alpha}l)\sum_{s'=1}^{N}\hat{y}(k,s')\bar{x}(s') + O(\tilde{\beta}^2).$$

We then obtain

$$\bar{x}(s+1) = \bar{x}(0)\tilde{\gamma}^{s+1}\left(\prod_{k=0}^{s}(1+\tilde{\alpha}k) + \tilde{\beta}\sum_{k=0}^{s}\prod_{l\neq k}^{s}(1+\tilde{\alpha}l)\sum_{s'=1}^{N}\hat{y}(k,s')\bar{x}(s')\right) + O(\tilde{\beta}^2).$$

From which we get

$$\bar{x}(s+1) = \bar{x}(0)\tilde{\gamma}^{s+1}\left(\prod_{k=0}^{s}(1+\tilde{\alpha}k) + \tilde{\beta}\sum_{k=0}^{s}\prod_{l\neq k}^{s}(1+\tilde{\alpha}l)\sum_{s'=1}^{N}\hat{y}(k,s')\bar{x}(0)\tilde{\gamma}^{s'}\prod_{k'=0}^{s'-1}(1+\tilde{\alpha}k')\right) + O(\tilde{\beta}^2).$$

Further, denoting by $\psi_s \equiv \tilde{\gamma}^s\prod_{k=0}^{s-1}(1+\tilde{\alpha}k)$, and $\omega_s \equiv \tilde{\gamma}^s\sum_{k=0}^{s-1}\prod_{l\neq k}^{s-1}(1+\tilde{\alpha}l)\sum_{s'=1}^{N}\hat{y}(k,s')\tilde{\gamma}^{s'}\prod_{k'=0}^{s'-1}(1+\tilde{\alpha}k')$ we can write $\bar{x}(s) = \bar{x}(0)\psi_s + \tilde{\beta}\bar{x}(0)^2\omega_s$. Moreover, from the normalization condition $1 = \sum_{s=0}^{N}\binom{N}{s}\bar{x}(s)$ (when there is no entry and exit) we obtain $1 = \sum_{s=0}^{N}\binom{N}{s}\bar{x}(s) = \bar{x}(0)\sum_{s=0}^{N}\binom{N}{s}\psi_s + \tilde{\beta}\bar{x}(0)^2\sum_{s=0}^{N}\binom{N}{s}\omega_s$. Denoting by $\Psi \equiv \sum_{s=0}^{N}\binom{N}{s}\psi_s$ and $\Omega \equiv \sum_{s=0}^{N}\binom{N}{s}\omega_s$ we find that $\bar{x}(0) = \frac{2}{\Psi+\sqrt{\Psi^2+4\tilde{\beta}\Omega}}$. From a first order Taylor expansion we then find $\bar{x}(0) = \frac{1}{\Psi} - \tilde{\beta}\frac{\Omega}{\Psi^3} + O(\beta^2)$, and $\tilde{\beta}\bar{x}(0)^2 = \frac{\tilde{\beta}}{\Psi^2} + O(\tilde{\beta}^2)$, so that $\bar{x}(s) = \frac{1}{\Psi}\psi_s +$

[51] By Stirling's formula we can approximate the Gamma function for large k as $\Gamma(k) = \sqrt{\frac{2\pi}{k}}\left(\frac{k}{e}\right)^k\left(1 + O\left(\frac{1}{k}\right)\right)$. Hence, $\frac{\Gamma(k)}{\Gamma(k+a)} = \left(1 + O\left(\frac{1}{k}\right)\right)\sqrt{(1+a/k)}(1+a/k)^{-k}\left(\frac{k}{k+a}\right)^k\left(\frac{k+a}{e}\right)^{-a}$. Since $\sqrt{(1+a/k)} \to 1$ for $k \to \infty$ this term is asymptotically negligible. Additionally $(1+a/k)^{-k} \to e^{-a}$ for $k \to \infty$, and $(k+a)^{-a} \sim k^{-a}$ for $k \to \infty$. Therefore, the leading order approximation of the ratio of Gamma functions is given by $\frac{\Gamma(k)}{\Gamma(k+a)} = k^{-a}\left(1 + O\left(\frac{1}{k}\right)\right)$.

$\tilde{\beta}\frac{1}{\Psi^2}\left(\omega_s - \frac{\Omega}{\Psi}\psi_s\right)$. Further, observe that $\frac{\psi_s}{\Psi} = \frac{\prod_{k=0}^{s-1}(1+\tilde{\alpha}k)}{\sum_{s'=0}^{N}\binom{N}{s'}\prod_{k=0}^{s'-1}(1+\tilde{\alpha}k)} = \frac{\prod_{k=0}^{s-1}(\gamma+\alpha k)}{\sum_{s'=0}^{N}\binom{N}{s'}\prod_{k=0}^{s'-1}(\gamma+\alpha k)} = \bar{x}(s)|_{\beta=0} \equiv \bar{x}_0(s)$. Hence, we can write $\bar{x}(s) = \bar{x}_0(s) + \tilde{\beta}\frac{\omega_s - \Omega\bar{x}_0(s)}{\Psi^2}$. This is $\bar{x}(s) = \bar{x}_0(s) + \frac{n\beta}{\gamma}\frac{\Omega}{\Psi^2}\left(\frac{\omega_s}{\Omega} - \bar{x}_0(s)\right) + O\left(\frac{\beta}{\gamma}\right)^2$, where $\bar{x}_0(s) \equiv \bar{x}(s)|_{\beta=0}$ is given in Equation (18), $\omega_s = \sum_{k=0}^{s-1}\prod_{l\neq k}^{s-1}\frac{\gamma+\alpha l}{\lambda}\sum_{s'=1}^{N}\binom{N-1}{s'-1}\frac{\bar{g}(k,s')}{1+\bar{g}(k,s')}\prod_{k''=0}^{s'-1}\frac{\gamma+\alpha k''}{\lambda}$, $\Omega = \sum_{s'=0}^{N}\binom{N}{s'}\omega_{s'}$ and $\Psi = \sum_{k=0}^{N}\binom{N}{k}\prod_{l=0}^{k-1}\frac{\gamma+\alpha l}{\lambda}$ for all $s = 0, \ldots, N$.

We next analyze the stability of the stationary state. In the following we denote by $F_s \equiv \frac{d\bar{x}(s)}{dt}$ and $Z_{s,s'} \equiv \frac{d\bar{y}(s,s')}{dt}$. In the case of independent markets, $\theta = 1$, when $\bar{g}(s,s')$ does not depend on $\bar{x}(s)$, we can write the Jacobian as follows

$$\mathbf{J} = \begin{bmatrix} \frac{\partial F_0}{\partial \bar{x}(0)} & \cdots & \frac{\partial F_0}{\partial \bar{x}(N)} & \frac{\partial F_0}{\partial \bar{y}(0,0)} & \cdots & \frac{\partial F_0}{\partial \bar{y}(N,N)} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \frac{\partial F_N}{\partial \bar{x}(0)} & \cdots & \frac{\partial F_N}{\partial \bar{x}(N)} & \frac{\partial F_N}{\partial \bar{y}(0,0)} & \cdots & \frac{\partial F_N}{\partial \bar{y}(N,N)} \\ 0 & \cdots & 0 & \frac{\partial Z_{0,0}}{\partial \bar{y}(0,0)} & 0 & 0 \\ \vdots & \ddots & \vdots & 0 & \ddots & 0 \\ 0 & \cdots & 0 & 0 & 0 & \frac{\partial Z_{N,N}}{\partial \bar{y}(N,N)} \end{bmatrix}. \tag{48}$$

The Jacobian is thus a block matrix of the form $\mathbf{J} = \begin{bmatrix} \mathbf{J}_{11} & \mathbf{J}_{12} \\ \mathbf{0} & \mathbf{J}_{22} \end{bmatrix}$, whose eigenvalues are the combined eigenvalues of $\mathbf{J}_{11}$ and $\mathbf{J}_{22}$. The latter is a diagonal matrix and has eigenvalues given by $\mu_s = \frac{\partial Z_{s,s}}{\partial \bar{y}(s,s)} = -\rho(1+\bar{g}(s,s))$. In order to compute the eigenvalues of the first, observe that

$$\begin{aligned} \frac{\partial F_s}{\partial \bar{x}(k)} &= s\left(\gamma + \alpha(s-1)\right)\delta_{k,s-1} + \lambda(N-s)\delta_{k,s+1} - (\lambda s + (N-s)(\gamma+\alpha s))\delta_{k,s} \\ &+ n\beta\left(s\sum_{s'=1}^{N}\binom{N-1}{s'-1}\bar{y}_{s-1,s'}\bar{x}(s')\delta_{k,s-1} - (N-s)\sum_{s'=1}^{N}\binom{N-1}{s'-1}\bar{y}_{s,s'}\bar{x}(s')\delta_{k,s}\right. \\ &\left. + s\bar{x}(s-1)\binom{N-1}{k-1}\bar{y}_{s-1,k} - (N-s)\bar{x}(s)\binom{N-1}{k-1}\bar{y}_{s,k}\right) \end{aligned}$$

This can be written as $\frac{\partial F_s}{\partial \bar{x}(k)} = F_{sk}^{(0)} + \beta F_{sk}^{(1)}$, where the matrix $\mathbf{F}^{(0)}$ is given by

$$\mathbf{F}^{(0)} = \begin{bmatrix} -\gamma N & \lambda N & 0 & 0 & 0 & \cdots \\ \gamma & -(\lambda + (N-1)(\gamma+\alpha)) & \lambda(N-1) & 0 & 0 & \cdots \\ 0 & 2(\gamma+\alpha) & -(2\lambda + (N-2)(\gamma+2\alpha)) & \lambda(N-2) & 0 & \cdots \\ 0 & 0 & 3(\gamma+2\alpha) & -(3\lambda + (N-3)(\gamma+3\alpha)) & \lambda(N-3) & \\ \vdots & \vdots & \ddots & \ddots & \ddots & \end{bmatrix}.$$

Observe that $\mathbf{F}^{(0)}$ is a real tridiagonal matrix of the form

$$\mathbf{F}^{(0)} = \begin{bmatrix} a_1 & b_2 & & \\ c_2 & a_2 & b_3 & \\ & \ddots & \ddots & \ddots \end{bmatrix}$$

where $a_i = -(\lambda(i-1) + (N-i+1)(\gamma+\alpha(i-1)))$, $b_i = \lambda(N-i+2)$ and $c_i = (i-1)(\gamma+\alpha(i-2))$. It is known that such a real tridiagonal matrix has only real and simple eigenvalues if $c_i b_i > 0$ [cf. e.g. Veselić, 1979]. We have that $c_i b_i = \lambda(N-i+2)(i-1)(\gamma+\alpha(i-2)) > 0$ and so this condition is satisfied. We thus conclude that $\mathbf{F}^{(0)}$ has only real and simple eigenvalues. Moreover, we have that [cf. Horn and Johnson, 1990, Theorem 6.3.12]

$$\mu_i(\mathbf{J}_{11}) = \mu_i(\mathbf{F}^{(0)}) + \beta\frac{\mathbf{w}_i^\top \mathbf{F}^{(1)}\mathbf{v}_i}{\mathbf{w}_i^\top \mathbf{v}_i} + O(\beta^2), \tag{49}$$

where $\mathbf{w}_i$ and $\mathbf{v}_i$ are the left and right eigenvectors corresponding to the eigenvalue $\mu_i(\mathbf{F}^{(0)})$. The characteristic equation $0 = (\mathbf{F}^{(0)} - \mu\mathbf{I}_n)\mathbf{v}$ for the right eigenvector $\mathbf{v}$ is given by the following second order difference equation with variable coefficients $k(\gamma+\alpha(k-1))v_k - (\lambda k + (N-k)(\gamma+\alpha k) + \mu)v_{k+1} + \lambda(N-k)v_{k+2} = 0$,

with the boundary conditions $v_0 = v_{N+2} = 0$ for $k = 0, \ldots, N-1$ [cf. Elaydi, 2005]. Similarly, the difference equation for the left eigenvector reads as $\lambda(N-k)w_k - (\lambda k + (N-k)(\gamma + \alpha k) + \mu)w_{k+1} + k(\gamma + \alpha(k-1))w_{k+2} = 0$, with the boundary conditions $w_0 = w_{N+2} = 0$ for $k = 0, \ldots, N-1$. We then find that all terms on the right hand side of Equation (49) are real up to the first order terms in β, showing that the eigenvalues of the Jacobian are real.

Finally, we give a proof of part (ii) of the proposition. In the limit of $\alpha, \gamma \to 0$ (no in-house R&D) from Equation (47) we find that $\bar{x}(s) = \bar{x}(0)\left(\frac{n\beta}{\lambda}\right)^s \prod_{k=0}^{s-1} \sum_{s'=1}^{N} \hat{y}(k,s')\bar{x}(s')$, where we have denoted by $\hat{y}(k,s') = \binom{N-1}{s'-1}\bar{y}(k,s')$. From this equation we see that $\bar{x}(s) = \delta_{s,0}$ is always a solution. We next compute a threshold β^c such that for all $\beta < \beta^c$ this is the unique solution. For $s = 1$ we obtain $\bar{x}(1) = \bar{x}(0)\frac{n\beta}{\lambda}\hat{y}(0,1)\bar{x}(1)$. When $\bar{x}(1) > 0$ (and consequently $\bar{x}(0) < 1$) we must have that $\bar{x}(0) = \frac{\lambda}{n\beta\hat{y}(0,1)}$. As $\bar{x}(0) < 1$ it must hold that $\beta > \beta^c \equiv \frac{\lambda}{n\hat{y}(0,1)} = \frac{\lambda(1+\bar{g}(0,1))}{n\bar{g}(0,1)}$. Note that $\bar{g}(0,1) = \frac{e^{\eta(\beta b\tau - 2c)}}{(1+e^{\eta(\beta b\tau - c)})(1+e^{-\eta c})}$, from which we get $\beta = \frac{1}{n}\lambda e^{-b\beta\eta\tau}\left(2e^{b\beta\eta\tau} + e^{\eta(b\beta\tau + c)} + e^{c\eta} + e^{2c\eta}\right)$, and solving for β yields $\beta^c = \frac{\lambda(2+e^{c\eta})}{n} + \frac{1}{b\eta\tau}W\left(\frac{1}{n}b\eta\lambda\tau\left(e^{c\eta}+1\right)e^{\frac{\eta(cn - b\lambda\tau(e^{c\eta}+2))}{n}}\right)$, and $W(x)$ is the Lambert W function (or product-log), which is implicitly defined by $W(x)e^{W(x)} = x$.

We next analyze the stability of the trivial solution. Denote by $F_s \equiv \frac{d\bar{x}(s)}{dt}$ and $Z_{s,s'} \equiv \frac{d\bar{y}(s,s')}{dt}$. In the case of independent markets, $\theta = 1$, when $\bar{g}(s,s')$ does not depend on $\bar{x}(s)$, we can write the Jacobian as in Equation (48), and the Jacobian is thus a block matrix of the form $\mathbf{J} = \begin{bmatrix} \mathbf{J}_{11} & \mathbf{J}_{12} \\ \mathbf{0} & \mathbf{J}_{22} \end{bmatrix}$, whose eigenvalues are the combined eigenvalues of $\mathbf{J}_{11}$ and $\mathbf{J}_{22}$. The latter is a diagonal matrix and has eigenvalues given by $\mu_s = \frac{\partial Z_{s,s}}{\partial \bar{y}(s,s)} = -\rho(1 + \bar{g}(s,s))$. In order to compute the eigenvalues of the first, $\mathbf{J}_{11}$, observe that in the limit of $\gamma, \alpha \to 0$ we have that

$$(\mathbf{J}_{11})_{sk} = \frac{\partial F_s}{\partial \bar{x}(k)} = \lambda(N-s)\delta_{k,s+1} - \lambda s\delta_{k,s} + n\beta\left(s\sum_{s'=1}^{N}\binom{N-1}{s'-1}\bar{y}_{s-1,s'}\bar{x}(s')\delta_{k,s-1}\right.$$
$$\left. -(N-s)\sum_{s'=1}^{N}\binom{N-1}{s'-1}\bar{y}_{s,s'}\bar{x}(s')\delta_{k,s} + s\bar{x}(s-1)\binom{N-1}{k-1}\bar{y}_{s-1,k} - (N-s)\bar{x}(s)\binom{N-1}{k-1}\bar{y}_{s,k}\right).$$

This can be written as $\dfrac{\partial F_s}{\partial \bar{x}(k)} = F_{sk}^{(0)} + \beta F_{sk}^{(1)}$, where the matrix $\mathbf{F}^{(0)}$ is given by

$$\mathbf{F}^{(0)} = \begin{bmatrix} 0 & \lambda N & 0 & 0 & 0 & \cdots \\ 0 & -\lambda & \lambda(N-1) & 0 & 0 & \cdots \\ 0 & 0 & -2\lambda & \lambda(N-2) & 0 & \cdots \\ 0 & 0 & 0 & -3\lambda & \lambda(N-3) & \\ \vdots & \vdots & \ddots & \ddots & \ddots & \end{bmatrix}.$$

Observe that $\mathbf{F}^{(0)}$ is a upper triangular matrix whose eigenvalues are given by the entries on the diagonal, that is, $0, -\lambda, -2\lambda, \ldots, -N\lambda$. Similarly, we get

$$\mathbf{F}^{(1)} = \begin{bmatrix} 0 & -nN\bar{y}(0,1)\bar{x}(0) & -nN(N-1)\bar{y}(0,2)\bar{x}(0) & \cdots \\ 0 & n\bar{y}(0,1)\bar{x}(0) & n(N-1)\bar{y}(0,2)\bar{x}(0) & \cdots \\ 0 & 0 & 0 & \cdots \\ \vdots & \vdots & \vdots & \end{bmatrix}.$$

Combining $\mathbf{F}^{(0)}$ and $\mathbf{F}^{(1)}$, and setting $\bar{x}(0) = 1$, yields

$$\mathbf{J}_{11} = \begin{bmatrix} 0 & N(\lambda - n\beta\bar{y}(0,1)) & -N(N-1)\bar{y}(0,2) & \cdots & & \\ 0 & -\lambda + n\beta\bar{y}(0,1) & (N-1)(\lambda + n\beta\bar{y}(0,2)) & \cdots & & \\ 0 & 0 & -2\lambda & \lambda(N-2) & 0 & \cdots \\ 0 & 0 & 0 & -3\lambda & \lambda(N-3) & \\ \vdots & \vdots & \ddots & \ddots & \ddots & \end{bmatrix}$$

This is an upper triangular matrix with eigenvalues given by $0, -\lambda + n\beta\bar{y}(0,1), -2\lambda, \ldots, -N\lambda$. All eigenvalues are non-positive if $n\beta < \frac{\lambda}{\bar{y}(0,1)}$, which is equivalent to the critical level β^c we have identified above. Hence, if $\beta < \beta^c$ then the trivial solution is asymptotically stable.

From the threshold β^c we finally find that

$$\frac{\partial \beta^c}{\partial \lambda} = \frac{b\eta\lambda\tau\left(e^{c\eta}+2\right)+nW(\cdot)}{b\eta\lambda n\tau(W(\cdot)+1)} > 0,$$

$$\frac{\partial \beta^c}{\partial c} = \frac{b\eta\lambda\tau e^{c\eta}\left(e^{c\eta}+1\right)+nW(\cdot)\left(2e^{c\eta}+1\right)}{bn\tau(W(\cdot)+1)\left(e^{c\eta}+1\right)} > 0,$$

$$\frac{\partial \beta^c}{\partial b} = -\frac{W(\cdot)\left(b\eta\lambda\tau\left(e^{c\eta}+2\right)+nW(\cdot)\right)}{b^2\eta n\tau(W(\cdot)+1)} < 0,$$

$$\frac{\partial \beta^c}{\partial \tau} = -\frac{W(\cdot)\left(b\eta\lambda\tau\left(e^{c\eta}+2\right)+nW(\cdot)\right)}{b\eta n\tau^2(W(\cdot)+1)} < 0,$$

where $W(\cdot)$ can be explicitly written as $W\left(\frac{1}{n}\left(e^{c\eta}+1\right)(b\eta\lambda\tau)e^{\frac{\eta\left(cn-b\lambda\tau\left(e^{c\eta}+2\right)\right)}{n}}\right)$. This concludes the proof. □

We next provide an explicit expression for $N = 1, 2$ of Equation (49). The difference equations for the eigenvalues μ together with the boundary conditions admit a closed form solution only in the cases of $N = 1$ and $N = 2$. In the case of $N = 1$ we obtain the two eigenvalues $\mu_1 = 0$, $\mu_2 = -(\gamma+\lambda)$. The corresponding left eigenvector is given by $\mathbf{v}_1 = \left[\frac{\lambda}{\gamma}, 1\right]^\top$ and $\mathbf{v}_2 = [-1, 1]^\top$, while the right eigenvector is $\mathbf{w}_1 = [1, 1]^\top$ and $\mathbf{w}_2 = \left[-\frac{\lambda}{\gamma}, 1\right]^\top$. In the case of $N = 2$ we get the three eigenvalues $\mu_1 = 0$, $\mu_{2,3} = -\frac{1}{2}\left(\alpha+3(\gamma+\lambda) \pm \sqrt{\alpha^2-2\alpha(\gamma-3\lambda)+(\gamma+\lambda)^2}\right)$. Observe that $\alpha^2-2\alpha(\gamma-3\lambda)+(\gamma+\lambda)^2 \geq \alpha^2-2\alpha\gamma+\gamma^2 = (\alpha-\gamma)^2 \geq 0$, so that all eigenvalues are real. The corresponding left eigenvectors are given by $\mathbf{v}_1 = \left[\frac{\lambda^2}{\gamma(\alpha+\gamma)}, \frac{\lambda}{\alpha+\gamma}, 1\right]^\top$, and

$$\mathbf{v}_2 = \left[\frac{-\alpha+\gamma-\lambda+\sqrt{\alpha^2-2\alpha(\gamma-3\lambda)+(\gamma+\lambda)^2}}{2(\alpha+\gamma)}, -\frac{\alpha+3\gamma-\lambda+\sqrt{\alpha^2-2\alpha(\gamma-3\lambda)+(\gamma+\lambda)^2}}{4(\alpha+\gamma)}, 1\right]^\top$$

$$\mathbf{v}_3 = \left[-\frac{\alpha-\gamma+\lambda+\sqrt{\alpha^2-2\alpha(\gamma-3\lambda)+(\gamma+\lambda)^2}}{2(\alpha+\gamma)}, \frac{-\alpha-3\gamma+\lambda+\sqrt{\alpha^2-2\alpha(\gamma-3\lambda)+(\gamma+\lambda)^2}}{4(\alpha+\gamma)}, 1\right]^\top,$$

while the right eigenvectors are $\mathbf{w}_1 = [1, 2, 1]^\top$, and

$$\mathbf{w}_2 = \left[\frac{\gamma\left(-\alpha+\gamma-\lambda+\sqrt{\alpha^2-2\alpha(\gamma-3\lambda)+(\gamma+\lambda)^2}\right)}{2\lambda^2}, -\frac{\alpha+3\gamma-\lambda+\sqrt{\alpha^2-2\alpha(\gamma-3\lambda)+(\gamma+\lambda)^2}}{2\lambda}, 1\right]^\top$$

$$\mathbf{w}_3 = \left[\frac{\gamma\left(-\alpha+\gamma-\lambda-\sqrt{\alpha^2-2\alpha(\gamma-3\lambda)+(\gamma+\lambda)^2}\right)}{2\lambda^2}, \frac{-\alpha-3\gamma+\lambda+\sqrt{\alpha^2-2\alpha(\gamma-3\lambda)+(\gamma+\lambda)^2}}{2\lambda}, 1\right]^\top.$$

We find that in all cases considered (for $\theta = 1$, β small and $N = 1, 2$) the Jacobian possesses only real eigenvalues.

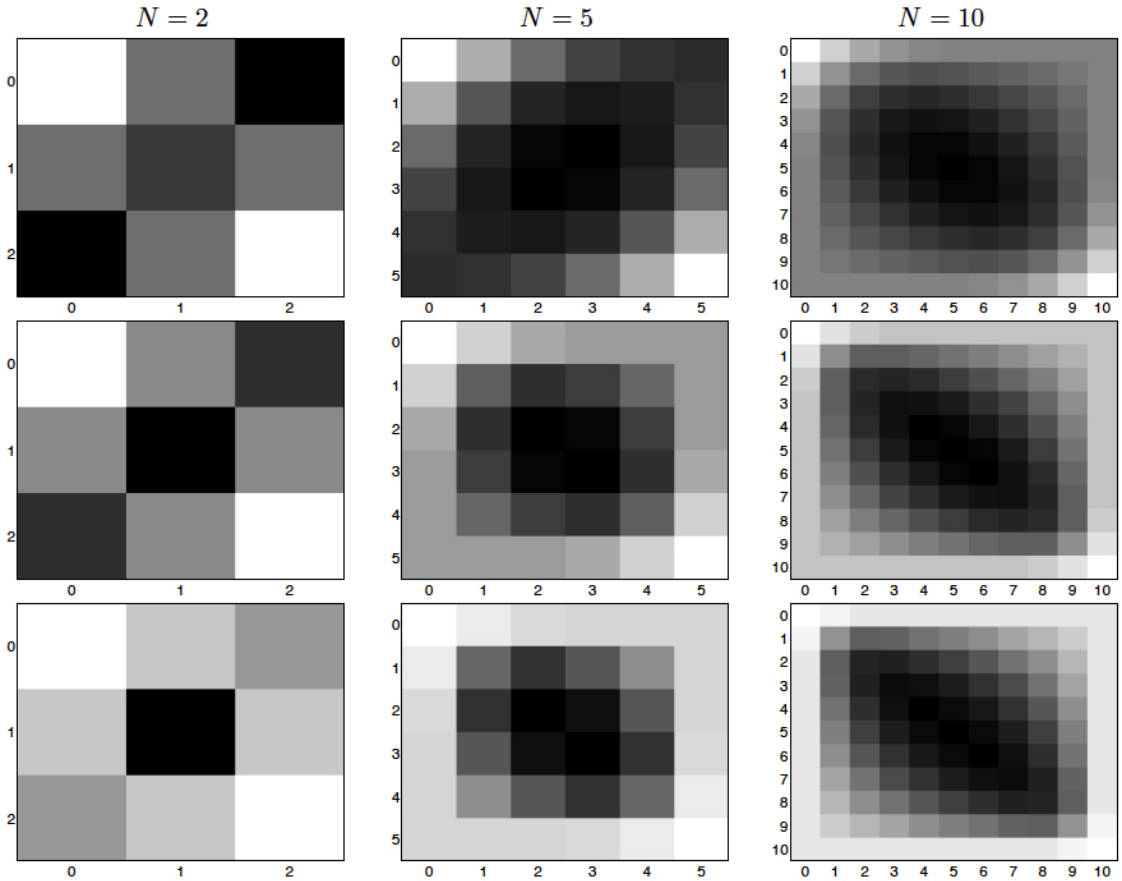


Figure C.2: Examples of the stationary expected fraction of links $\bar{y}(s, s')$, $0 \leq s, s' \leq N$ of Equation (16) with $c = 1$, $\theta = 1$, $\beta\tau b = 1$, for $\eta = 1, 2, 3$ (rows) and $N = 2, 5, 10$ (columns) (where higher values are black and lower values are white). We observe that with increasing values of η (and $c > 0$) the number of links is highest along the diagonal with firms having similar portfolio sizes, except for the upper left and lower right corners. This indicates assortative matching. That is, firms with similar portfolio sizes tend to be connected, however, their portfolios need to be composed of different technologies.

Online Appendix for "Endogenous Technology Cycles in Dynamic R&D Networks"

Michael D. König[a] and Tim Rogers[b]

[a] *Department of Economics, University of Zurich, Schönberggasse 1, CH-8001 Zurich, Switzerland.*
[b] *Department of Mathematical Sciences, University of Bath, BA2 7AY, U.K.*

Contents

D. Table of Symbols and Notations

Technology Space		Section 3.1
$N \in \mathbb{N}$	Number of technologies	
$\mathcal{H}^N = \{0,1\}^N$	Technology space	
$\mathbf{h}_i \in \mathcal{H}^N$	Technology portfolio of firm i	
$\mathbf{h}_i^c = \mathbf{u} - \mathbf{h}_i$	Complement of the technology portfolio of firm i	
$\mathrm{S}(\mathbf{h}_i) = \langle \mathbf{h}_i, \mathbf{u} \rangle$	Technology stock of firm i	
$\bar{h} = \frac{1}{n}\sum_{i=1}^n \mathrm{S}(\mathbf{h}_i)$	Average technology stock	
Market Level		Section 3.1
$n \in \mathbb{N}$	Number of firms	
$\mathcal{N} = \{1, \dots, n\}$	Set of firms	
$\theta \in \{0,1\}$	Competition parameter	
Firm Level		Section 3.1
$A_i = a + b\lvert \mathsf{S}(\mathbf{h}_i)\rvert$	Firm i's productivity	
$a, b \in \mathbb{R}_+$	Productivity parameters	
$\pi_i = \theta A_i + (1-\theta) A_i n \frac{A_i}{\sum_{j=1}^n A_j} - c d_i$	Firm i's profit	
Network		Section 3.1
$\mathcal{G}(n)$	Set of networks with n firms	
$\mathbf{A} = (a_{ij})_{1 \le i,j \le n}$	Adjacency matrix	
$d_i = \sum_{j=1}^n a_{ij}$	Number of collaborations of firm i, i.e. the degree of i	
$\bar{d} = \frac{1}{n}\sum_{i=1}^n d_i$	Average degree	
Innovation		Section 3.2
$\gamma, \alpha \in \mathbb{R}_+$	In-house R&D parameters	
$\beta \in \mathbb{R}_+$	Technology Spillover parameters	
$\nu_{ik} = \gamma + \alpha \sum_{l=1}^N h_{il} + \beta \sum_{j=1}^n a_{ij} h_{jk}$	Innovation arrival rate	
$\lambda \in \mathbb{R}_+$	Obsolescence rate	
Link Formation		Section 3.3
$\tau \in \mathbb{R}_+$	Alliance duration	
$\rho \in \mathbb{R}_+$	Link creation/expiration rate	
$c \in \mathbb{R}_+$	Linking cost	
$V_i^\tau(\mathbf{h}, G + ij) = V_i^\tau(\mathbf{h}, G) + \beta g_{\theta,\tau}(\bar{h})(1 + b\lvert \mathsf{S}(\mathbf{h}_i)\rvert)^{1-\theta} \langle \mathbf{h}_i^c, \mathbf{h}_j \rangle - c$	Present discounted profit of firm i from collaborating with firm j	
$g_{\theta,\tau}(\bar{h}) = \tau b \left(\theta + 2\frac{1-\theta}{1+b\bar{h}}\right)$	Competition effect in marginal profits from collaborations	
$\vartheta \in \mathbb{R}_+$	Uncertainty parameter	
Entry and Exit		Section 3.4
$\xi \in \mathbb{R}_+$	Exit rate	
$\chi \in \mathbb{R}_+$	Entry rate	
$\kappa \in \mathbb{R}_+$	Entry cost	
$f_{\theta,\kappa}^\vartheta(\bar{h}) = \Lambda^\vartheta\left(\theta + \frac{1-\theta}{1+b\bar{h}_t} - \kappa\right)$	Entry probability	
$\Lambda^\vartheta(x) = \frac{e^{\vartheta x}}{1+e^{\vartheta x}}$	Logistic function	

E. Network Definitions and Characterizations

A *network* (*graph*) G is the pair $(\mathcal{N}, \mathcal{E})$ consisting of a set of *nodes* (*vertices*) $\mathcal{N} = \{1, \ldots, n\}$ and a set of edges (*links*) $\mathcal{E} \subset \mathcal{N} \times \mathcal{N}$ between them. We denote by $\mathcal{G}(n)$ the set of graphs with n nodes. A link (i, j) is *incident* with nodes i and j. The *neighborhood* of a node $i \in \mathcal{N}$ is the set $\mathcal{N}_i = \{j \in \mathcal{N} : (i, j) \in \mathcal{E}\}$. The *degree* d_i of a node $i \in \mathcal{N}$ gives the number of links incident to node i. Clearly, $d_i = |\mathcal{N}_i|$. Let $\mathcal{N}_i^{(2)} = \bigcup_{j \in \mathcal{N}_i} \mathcal{N}_j \setminus (\mathcal{N}_i \cup \{i\})$ denote the second-order neighbors of node i. Similarly, the k-th order neighborhood of node i is defined recursively from $\mathcal{N}_i^{(0)} = \{i\}$, $\mathcal{N}_i^{(1)} = \mathcal{N}_i$ and $\mathcal{N}_i^{(k)} = \bigcup_{j \in \mathcal{N}_i^{(k-1)}} \mathcal{N}_j \setminus \left(\bigcup_{l=0}^{k-1} \mathcal{N}_i^{(l)} \right)$. A *walk* in G of length k from i to j is a sequence $\langle i_0, i_1, \ldots, i_k \rangle$ of nodes such that $i_0 = i$, $i_k = j$, $i_p \neq i_{p+1}$, and i_p and i_{p+1} are (directly) linked, that is $i_p i_{p+1} \in \mathcal{E}$, for all $0 \leq p \leq k-1$. Nodes i and j are said to be *indirectly linked* in G if there exists a walk from i to j in G containing nodes other than i and j. A pair of nodes i and j is connected if they are either directly or indirectly linked. A node $i \in \mathcal{N}$ is *isolated* in G if $\mathcal{N}_i = \emptyset$. The network G is said to be *empty* (denoted by $\overline{K}_n$) when all its nodes are isolated.

A *subgraph*, G', of G is the graph of subsets of the nodes, $\mathcal{N}(G') \subseteq \mathcal{N}(G)$, and links, $\mathcal{E}(G') \subseteq \mathcal{E}(G)$. A graph G is *connected*, if there is a path connecting every pair of nodes. Otherwise G is disconnected. The *components* of a graph G are the maximally connected subgraphs. A component is said to be *minimally connected* if the removal of any link makes the component disconnected.

Given a graph G and a set $\mathcal{S} \subseteq \mathcal{N}$, we say that $G_{\mathcal{S}}$ is the *subgraph* G *induced* $\mathcal{S}$ whenever the adjacency matrix of $G_{\mathcal{S}}$ is $\mathbf{A}_{\mathcal{S}}$. We write $G_{-\mathcal{S}}$ to denote the network $G_{\mathcal{N} \setminus \mathcal{S}}$, that is $G_{-\mathcal{S}}$ is the network that results after eliminating all the nodes in $\mathcal{S}$.

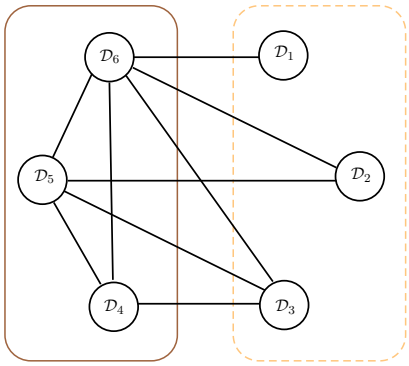
A *dominating set* for a graph $G = (\mathcal{N}, \mathcal{E})$ is a subset $\mathcal{S} \subseteq \mathcal{N}$ such that every node not in $\mathcal{S}$ is connected to at least one member of S by a link. An *independent set* is a set of nodes in a graph in which no two nodes are adjacent. For example the central node in a star $K_{1,n-1}$ forms a dominating set while the peripheral nodes form an independent set.

Let $G = (\mathcal{N}, \mathcal{E})$ be a graph whose distinct positive degrees are $d_{(1)} < d_{(2)} < \ldots < d_{(k)}$, and let $d_0 = 0$ (even if no agent with degree 0 exists in G). Further, define $\mathcal{D}_i = \{v \in \mathcal{N} : d_v = d_{(i)}\}$ for $i = 0, \ldots, k$. Then the set-valued vector $\mathcal{D} = (\mathcal{D}_0, \mathcal{D}_1, \ldots, \mathcal{D}_k)$ is called the *degree partition* of G. A *nested split graph* is a graph with a nested neighborhood structure such that the set of neighbors of each node is contained in the set of neighbors of each higher degree node [Cvetkovic and Rowlinson, 1990; Mahadev and Peled, 1995]. Let $\mathcal{D} = (\mathcal{D}_0, \mathcal{D}_1, \ldots, \mathcal{D}_k)$ be the degree partition of a nested split graph $G = (\mathcal{N}, \mathcal{E})$. Then the nodes $\mathcal{N}$ can be partitioned in independent sets $\mathcal{D}_i$, $i = 1, \ldots, \lfloor \frac{k}{2} \rfloor$ and a dominating set $\bigcup_{i=\lfloor \frac{k}{2} \rfloor + 1}^{k} \mathcal{D}_i$ in the graph $G' = (\mathcal{N} \setminus \mathcal{D}_0, \mathcal{E})$. Moreover, the neighborhoods of the nodes are nested. In particular, for each node $v \in \mathcal{D}_i$, $\mathcal{N}_v = \bigcup_{j=1}^{i} \mathcal{D}_{k+1-j}$ if $i = 1, \ldots, \lfloor \frac{k}{2} \rfloor$ if $i = 1, \ldots, k$, while $\mathcal{N}_v = \bigcup_{j=1}^{i} \mathcal{D}_{k+1-j} \setminus \{v\}$ if $i = \lfloor \frac{k}{2} \rfloor + 1, \ldots, k$. See also the left panel in Figure E.1.

In a *complete* graph K_n, every node is adjacent to every other node. The graph in which no pair of nodes is adjacent is the empty graph $\overline{K}_n$. A *clique* $K_{n'}$, $n' \leq n$, is a complete subgraph of the network G. A graph is *k-regular* if every node i has the same number of links $d_i = k$ for all $i \in \mathcal{N}$. The complete graph K_n is $(n-1)$-regular. The cycle C_n is 2-regular. In a *bipartite graph* there exists a partition of the nodes in two disjoint sets $\mathcal{V}_1$ and $\mathcal{V}_2$ such that each link connects a node in $\mathcal{V}_1$ to a node in $\mathcal{V}_2$. $\mathcal{V}_1$ and $\mathcal{V}_2$ are independent sets with cardinalities n_1 and n_2, respectively. In a complete bipartite graph K_{n_1,n_2} each node in $\mathcal{V}_1$ is connected to each other node in $\mathcal{V}_2$. The *star* $K_{1,n-1}$ is a complete bipartite graph in which $n_1 = 1$ and $n_2 = n-1$.

The *complement* of a graph G is a graph $\overline{G}$ with the same nodes as G such that any two nodes of $\overline{G}$ are adjacent if and only if they are not adjacent in G. For example the complement of the complete graph K_n is the empty graph $\overline{K}_n$.

Let $\mathbf{A}$ be the symmetric $n \times n$ *adjacency matrix* of the network G. The element $a_{ij} \in \{0, 1\}$ indicates if there exists a link between nodes i and j such that $a_{ij} = 1$ if $(i, j) \in \mathcal{E}$ and $a_{ij} = 0$ if $(i, j) \notin \mathcal{E}$. The k-th power of the adjacency matrix is related to walks of length k in the graph. In particular, $\left(\mathbf{A}^k\right)_{ij}$ gives the number of walks of length k from node i to node j. The *eigenvalues* of the adjacency matrix $\mathbf{A}$ are the numbers $\lambda_1, \lambda_2, \ldots, \lambda_n$ such that $\mathbf{A}\mathbf{v}_i = \lambda_i \mathbf{v}_i$ has a nonzero solution



$$\mathbf{A} = \begin{pmatrix} 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 1 & 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Figure E.1: (Left panel) Representation of nested split graphs and their degree partitions $\mathcal{D}$ with corresponding adjacency matrices $\mathbf{A}$. A line between $\mathcal{D}_i$ and $\mathcal{D}_j$ indicates that every node in $\mathcal{D}_i$ is adjacent to every node in $\mathcal{D}_j$. The partitions included in the solid frame ($\mathcal{D}_i$ with $\lfloor \frac{k}{2} \rfloor + 1 \leq i \leq k$) are the dominating subsets while the partitions in the dashed frame ($\mathcal{D}_i$ with $1 \leq i \leq \lfloor \frac{k}{2} \rfloor$) are the independent sets with $k = 6$. The illustration follows Mahadev and Peled [1995, p. 11]. (Right panel) The corresponding stepwise adjacency matrix $\mathbf{A}$ with elements a_{ij} satisfying the following condition: if $i < j$ and $a_{ij} = 1$ then $a_{hk} = 1$ whenever $h < k \leq j$ and $h \leq i$.

vector $\mathbf{v}_i$, which is an *eigenvector* associated with λ_i for $i = 1, \ldots, n$. Since the adjacency matrix $\mathbf{A}$ of an undirected graph G is real and symmetric, the eigenvalues of $\mathbf{A}$ are real, $\lambda_i \in \mathbb{R}$ for all $i = 1, \ldots, n$. Moreover, if $\mathbf{v}_i$ and $\mathbf{v}_j$ are eigenvectors for different eigenvalues, $\lambda_i \neq \lambda_j$, then $\mathbf{v_i}$ and $\mathbf{v}_j$ are orthogonal, i.e. $\mathbf{v}_i^\top \mathbf{v}_j = 0$ if $i \neq j$. In particular, $\mathbb{R}^n$ has an orthonormal basis consisting of eigenvectors of $\mathbf{A}$. Since $\mathbf{A}$ is a real symmetric matrix, there exists an orthogonal matrix $\mathbf{S}$ such that $\mathbf{S}^\top \mathbf{S} = \mathbf{S}\mathbf{S}^\top = \mathbf{I}$ (that is $\mathbf{S}^\top = \mathbf{S}^{-1}$) and $\mathbf{S}^\top \mathbf{A} \mathbf{S} = \mathbf{D}$, where $\mathbf{D}$ is the diagonal matrix of eigenvalues of $\mathbf{A}$ and the columns of $\mathbf{S}$ are the corresponding eigenvectors. The *Perron-Frobenius eigenvalue* $\lambda_{\text{PF}}(G)$ is the *largest real eigenvalue* of $\mathbf{A}$ associated with G, i.e. all eigenvalues λ_i of $\mathbf{A}$ satisfy $|\lambda_i| \leq \lambda_{\text{PF}}(G)$ for $i = 1, \ldots, n$ and there exists an associated nonnegative eigenvector $\mathbf{v}_{\text{PF}} \geq 0$ such that $\mathbf{A}\mathbf{v}_{\text{PF}} = \lambda_{\text{PF}}(G)\mathbf{v}_{\text{PF}}$. For a connected graph G the adjacency matrix $\mathbf{A}$ has a unique largest real eigenvalue $\lambda_{\text{PF}}(G)$ and a positive associated eigenvector $\mathbf{v}_{\text{PF}} > 0$. There exists a relation between the number of walks in a graph and its eigenvalues. The number of closed walks of length k from a node i in G to herself is given by $\left(\mathbf{A}^k\right)_{ii}$ and the total number of closed walks of length k in G is $\text{tr}\left(\mathbf{A}^k\right) = \sum_{i=1}^n \left(\mathbf{A}^k\right)_{ii} = \sum_{i=1}^n \lambda_i^k$. We further have that $\text{tr}\left(\mathbf{A}\right) = 0$, $\text{tr}\left(\mathbf{A}^2\right)$ gives twice the number of links in G and $\text{tr}\left(\mathbf{A}^3\right)$ gives six times the number of triangles in G.

A nested split graph is characterized by a *stepwise adjacency matrix* $\mathbf{A}$, which is a symmetric, binary $(n \times n)$-matrix with elements a_{ij} satisfying the following condition: if $i < j$ and $a_{ij} = 1$ then $a_{hk} = 1$ whenever $h < k \leq j$ and $h \leq i$. See also the right panel in Figure E.1. Both, the complete graph, K_n, as well as the star $K_{1,n-1}$, are particular examples of nested split graphs. Nested split graphs are also the graphs which maximize the largest eigenvalue, $\lambda_{\text{PF}}(G)$, [Brualdi and Solheid, 1986], and they are the ones that maximize the degree variance [Peled et al., 1999]. See e.g. König et al. [2014] for further properties.

F. Additional Results for Section 5.1

F.1. Innovation Gains from R&D Collaborations

Using Equation (19) from the last part of Proposition 2, it follows that the average stock of knowledge is given by

$$\bar{h} = \sum_{s=1}^{N} s \binom{N}{s} \bar{x}(s) = \bar{h}_0 + \frac{n\beta}{\gamma} \frac{\Omega}{\Psi^2} \left(\frac{\sum_{s=1}^{N} s \binom{N}{s} \omega(s)}{\Omega} - \bar{h}_0 \right) + O(\beta^2),$$

where $\bar{h}_0 \equiv \bar{h}\big|_{\beta=0}$ (see Proposition 2). The stationary average stock of knowledge $\bar{h}$ can be seen in Figure F.1 for varying values of γ and λ. The innovation gains from R&D collaborations are then

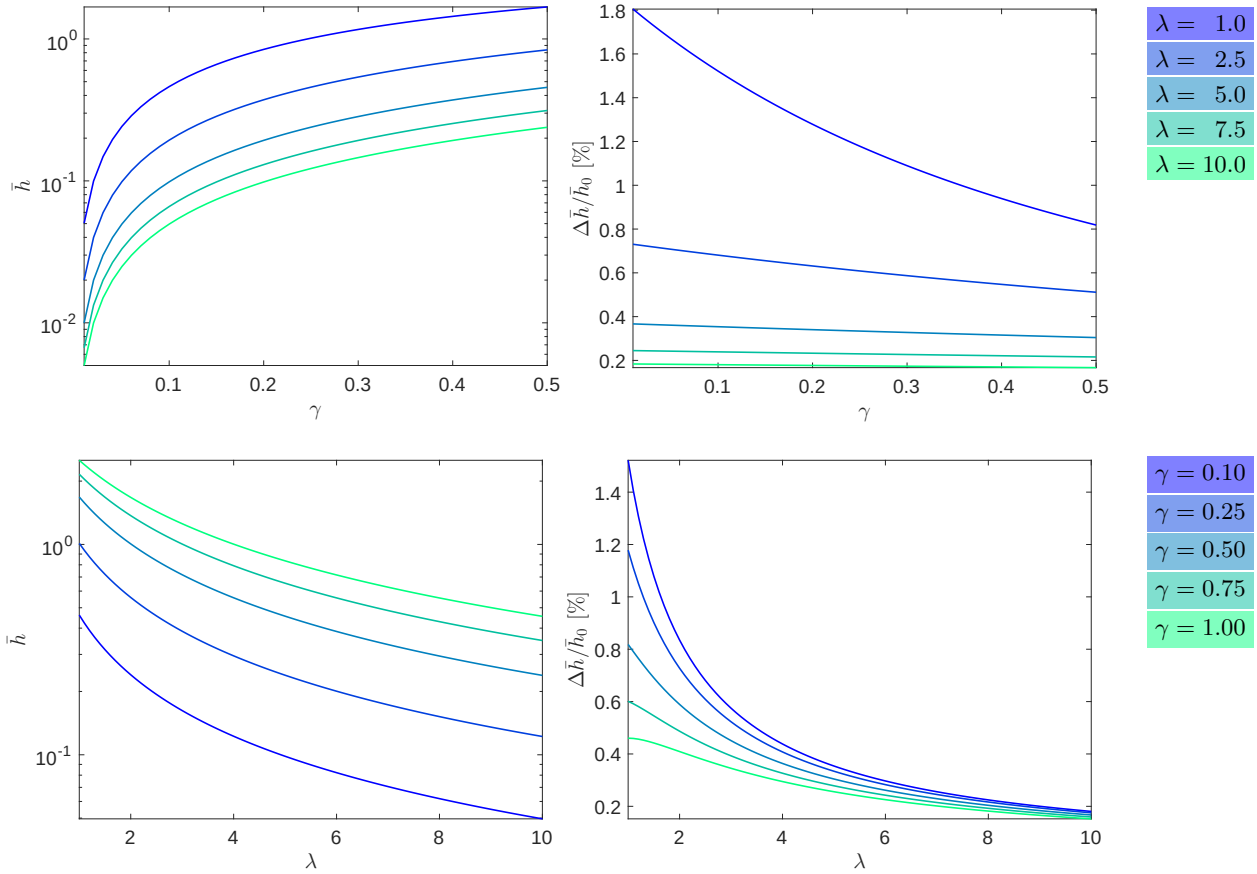


Figure F.1: (Top left panel) The stationary average stock of knowledge $\bar{h}$ over different values of $\gamma \in [0, 0.5]$ for varying values of $\lambda \in \{1, 2.5, 5.0, 7.5, 10\}$. (Top right panel) The relative percentage gain in the average stock of knowledge, $\Delta\bar{h}/\bar{h}_0$, for the same parameters. (Bottom left panel) The stationary average stock of knowledge $\bar{h}$ over different values of $\lambda \in [1, 10]$ for varying values of $\gamma \in \{0.1, 0.25, 0.50, 0.75, 1\}$. (Bottom right panel) The relative percentage gain in the average stock of knowledge, $\Delta\bar{h}/\bar{h}_0$, for the same parameters. The parameters are $\theta = 1$, $N = 5$, $c = 0.1$, $\eta = 1$, $\alpha = 0$, $\tau = 0.01$, $b = 1$ and $\beta = 0.1$.

given by

$$\Delta\bar{h} \equiv \bar{h} - \bar{h}_0 = \frac{n\beta}{\gamma}\frac{\Omega}{\Psi^2}\left(\frac{\sum_{s=1}^{N} s\binom{N}{s}\omega(s)}{\Omega} - \bar{h}_0\right) + O(\beta^2).$$

The relative gains from R&D collaborations, $\Delta\bar{h}/\bar{h}_0$, are illustrated in the right panels in Figure F.1. As we have assumed that β is small the figure tends to underestimate the increase in the average stock of knowledge due to collaborations. We also find that the relative gains decrease with the in-house R&D success rate γ, and the knowledge obsolescence rate λ.

F.2. Power-Law Degree Distributions

It is possible to characterizes the stationary knowledges stocks and degree distributions in the case of vanishing spillovers ($\beta = 0$), and show that they decay as a power-law (see also Figure 4), even in the presence of entry end exit. For this purpose, assume that $\beta = \lambda = \vartheta = 0$, $\alpha = \tilde{\alpha}/N$, $\gamma = \tilde{\gamma}/N$, in Equation (40). Then letting $N \to \infty$ we obtain

$$\frac{d\tilde{x}_t(s)}{dt} = (\tilde{\gamma} + \tilde{\alpha}(s-1))\tilde{x}_t(s-1) - (\tilde{\gamma} + \tilde{\alpha}s + \xi)\tilde{x}_t(s) + \chi/2\left(1 - \sum_{s'=0}^{\infty}\tilde{x}_t(s')\right)\mathbb{1}_{\{s=0\}}.$$

At stationarity we find a recursion relation for $x(s)$ for $s \geq 2$, $\tilde{x}_t(s) = \frac{(\tilde{\gamma}+\tilde{\alpha}(s-1))}{(\tilde{\gamma}+\tilde{\alpha}s+\xi)}\tilde{x}_t(s-1)$. Up to a multiplicative constant, the solution is given by $\tilde{x}(s) \propto \frac{\Gamma(\tilde{\gamma}/\tilde{\alpha}+s)}{\Gamma(1+\tilde{\gamma}/\tilde{\alpha}+\xi/\tilde{\alpha}+s)}$. Taking Stirling's approximation to the Gamma functions,[51] we obtain the power-law tail $\tilde{x}(s) \sim s^{-(1+\xi/\tilde{\alpha})}$. An illustration can be see in the left panel of Figure 4.

For the degree distribution we first have to check what is going on with g. Set $\theta = 1$ for

simplicity. As $N \to \infty$, for $s' < s$ the sum in Footnote 28 is dominated by the last term (check all the combinatorial factors to see this) so for small β we get the relatively simple expression

$$g^{\eta}_{1,\tau}(s,s') = \frac{1}{(1+e^c)^2} + \beta \frac{b\eta\tau e^c}{(1+e^c)^3}(s+s') + O(\beta^2)\,.$$

Note that because of the factors of τ and b, it should not be too hard to take small β (so that spillovers are not so important for technology stock distribution) but still have correlation to linking. The density of edges between firms with technology stocks s, s' is

$$\bar{y}(s,s') = \frac{g^{\eta}_{1,\tau}(s,s')}{1+g^{\eta}_{1,\tau}(s,s')} = \frac{A+B(s+s')}{1+A+B(s+s')}\,, \tag{50}$$

for constants A, B. The knowledge stock distribution is

$$f(s) = \frac{\tilde{x}(s)}{\sum_{s'}\tilde{x}(s')} = \frac{\Gamma(1+\tilde{\gamma}/\tilde{\alpha}+\xi/\tilde{\alpha})\Gamma(\tilde{\gamma}/\tilde{\alpha}+s)}{\Gamma(\tilde{\gamma}/\alpha)_2F_1(1,\tilde{\gamma}/\alpha,1+\tilde{\gamma}/\tilde{\alpha}+\xi/\tilde{\alpha};1)\Gamma(1+\tilde{\gamma}/\tilde{\alpha}+\xi/\tilde{\alpha}+s)}\,,$$

with mean

$$\langle s\rangle \equiv \mathbb{E}(|\,\mathsf{S}(\mathbf{h}_i)|) = \frac{\Gamma(1+\tilde{\gamma}/\tilde{\alpha}+\xi/\tilde{\alpha})\Gamma(\tilde{\gamma}/\tilde{\alpha})_2F_1(2,\tilde{\gamma}/\tilde{\alpha},1+\tilde{\gamma}/\tilde{\alpha}+\xi/\tilde{\alpha};1)}{\Gamma(2+\tilde{\gamma}/\alpha+\tilde{\gamma}/\tilde{\alpha})\Gamma(\tilde{\gamma}/\tilde{\alpha})_2F_1(1,\tilde{\gamma}/\tilde{\alpha},1+\tilde{\gamma}/\tilde{\alpha}+\xi/\tilde{\alpha};1)}\,.$$

Following section 3.4 in Hurd [2015] we find the degree distribution given by the mixed Poisson $P(k) = \sum_s f(s)e^{-\mu(s)}\frac{\mu(s)^k}{k!}$, where $\mu(s) \equiv A + B(s+\langle s\rangle)$. Note that a mixed Poisson distribution with an unbounded distribution of the mixing variable inherits the tail distribution of the mixing variable [Hurd, 2015]. Hence, the tail of the degree distribution decays with the same power-law exponent as the knowledge stocks distribution. An illustration can be see in the right panel of Figure 4.

G. Additional Results for Section 5.2

G.1. A Single Technology ($N = 1$)

The case of a single technology, albeit being very restrictive, is interesting to analyze because it is more tractable and it allows to gain insights which can also be observed for the more general cases. In particular, if $N = 1$ then the only possible technology profiles are $\mathbf{h} \in \{0,1\}$.[52] We then can state the following proposition, characterizing the stationary states and their stability properties.

Proposition 3. *Consider large ρ such that terms of the order $o(\rho)$ can be neglected in Equation (11) and let $N = 1$. Denote by $x(h) \equiv \lim_{t\to\infty} x_t(h)$ and $y(h,h') \equiv \lim_{t\to\infty} y_t(h,h')$ the stationary states of the dynamic system in Equations (10) and (11), respectively, for $0 \le h, h', \le 1$.*

(i) ***Threshold:*** *We have that $x(0) = 1$ is an asymptotically stable fixed point in the limit of $\gamma \to 0$ if $\beta < \beta^c$, where β^c solves the following equation:*

$$\beta = \frac{(\lambda+\xi)\left(\xi+\chi+\xi e^{(\kappa-1)\vartheta}\right)}{n\chi}\frac{e^{(2-\theta)b\beta\eta\tau}}{2e^{b\beta\eta\tau}+e^{\eta((2-\theta)b\beta\tau+c)}+e^{c\eta}+e^{2c\eta}}, \quad \theta\in\{0,1\}.$$

If, in addition, we assume that $\xi = \chi = 0$ (no entry or exit) then we can compute β^c explicitly as

$$\beta^c = \frac{\lambda\left(2+e^{c\eta}\right)}{n} + \frac{1}{(2-\theta)b\eta\tau}W\left(\frac{(2-\theta)b\eta\lambda\tau\left(1+e^{c\eta}\right)}{n}e^{\frac{\eta\left(cn-(2-\theta)b\lambda\tau\left(2+e^{c\eta}\right)\right)}{n}}\right), \quad \theta\in\{0,1\}, \tag{51}$$

[52] Note that $\bar{h} = x(1)$ and $\bar{d}_t = (2y_t(0,1)x_t(0)x_t(1) + y_t(0,0)x_t(0)^2 + y_t(1,1)x_t(1)^2)n$.

which is equivalent to Equation (17) in the case of $\theta = 1$, and where $W(x)$ is the Lambert W function (or product-log) implicitly defined by $W(x)e^{W(x)} = x$.

(ii) ***Single-product monopolies:*** *Let $\theta = 1$ then the stationary state is given by*

$$x(0) = \frac{T_+(\cdot)(1+g)}{2n\beta g(f\chi+\xi)},\ x(1) = \frac{2\gamma f\chi}{T_-(\cdot)},\ y(0,1) = \frac{g_{01}}{1+g_{01}},\ y(0,0) = \frac{g_{00}}{1+g_{00}},\ y(1,1) = \frac{g_{11}}{1+g_{11}},$$

where $g_{01} \equiv g^{\eta}_{1,\tau}(0,1)$, $g_{00} \equiv g^{\eta}_{1,\tau}(0,0)$, $g_{11} \equiv g^{\eta}_{1,\tau}(1,1)$, $f \equiv f^{\vartheta}_1$ and functions $T_\pm(\cdot)$ of the parameters whose expressions can be found in the proof of the proposition. If also $\xi = \chi = 0$ (no entry or exit) then the stationary state is asymptotically stable and given by

$$x(1) = \frac{\beta gn - (1+g)(\gamma+\lambda) + \sqrt{((1+g)(\gamma+\lambda) - \beta gn)^2 + 4\beta\gamma g(1+g)n}}{2\beta gn}, \tag{52}$$

where $x(0) = 1 - x(1)$. Moreover, in the limit of $\gamma \to 0$ the non-trivial solution ($x(1) > 0$) is given by

$$x(1) = 1 - \frac{\lambda\left(2 + e^{c\eta} + (1+e^{c\eta})\, e^{\eta(c-b\beta\tau)}\right)}{\beta n}. \tag{53}$$

In all the above cases the Jacobian has only real eigenvalues, so that the solution trajectories for $x_t(0), x_t(1)$, $y_t(0,0)$, $y_t(1,1)$ and $y_t(0,1)$ do not exhibit oscillatory behavior.

(iii) ***High uncertainty:*** *In the case of $\eta \to 0$, $\vartheta \to 0$ and $\theta \in \{0,1\}$ the asymptotically stable stationary state is given by*

$$x(0) = \frac{5(2\xi+\chi)(\gamma+\lambda+\xi) + \beta n\chi - S(\cdot)}{2\beta n(2\xi+\chi)},\ x(1) = \frac{10\gamma\chi}{5(2\xi+\chi)(\gamma+\lambda+\xi) - \beta n\chi + S(\cdot)},$$

and $y(0,1) = y(0,0) = y(1,1) = \frac{1}{5}$ with a function $S(\cdot)$ of the parameters whose expression can be found in the proof of the proposition. If also $\xi = \chi = 0$ (no entry or exit) then

$$x(1) = \frac{n\beta - 5(\gamma+\lambda) + \sqrt{(5(\gamma+\lambda) - \beta n)^2 + 20\beta\gamma n}}{2\beta n}, \tag{54}$$

and $x(0) = 1 - x(1)$. In all the above cases the Jacobian has only real eigenvalues, so that the solution trajectories for $x_t(0), x_t(1)$, $y_t(0,0)$, $y_t(1,1)$ and $y_t(0,1)$ do not exhibit oscillatory behavior.

(iv) ***Low uncertainty:*** *When $\eta \to \infty$, $\vartheta \to \infty$, $\theta \in \{0,1\}$ and $c > 0$ then the asymptotically stable stationary state is given by*

$$x(0) = \frac{\chi(\lambda+\xi)}{(\xi+\chi)(\gamma+\lambda+\xi)},\ x(1) = \frac{\gamma\chi}{(\xi+\chi)(\gamma+\lambda+\xi)},$$

and $y(0,1) = y(0,0) = y(1,1) = 0$, given that $\kappa < 1$ when $\theta = 1$ and $1 + \frac{b\gamma\chi}{(\xi+\chi)(\gamma+\lambda+\xi)} < \frac{1}{\kappa}$ when $\theta = 0$. If also $\xi = \chi = 0$ (no entry or exit) then $x(1) = \frac{\gamma}{\lambda+\gamma}$ and $x(0) = 1 - x(1) = \frac{\lambda}{\lambda+\gamma}$. Otherwise the stationary state is $x(0) = x(1) = y(0,1) = 0$. In all the above cases the Jacobian has only real eigenvalues (the eigenvalue spectrum is given by $\{-\gamma-\lambda-\xi, -\rho, -\xi-\chi\}$), so that the solution trajectories for $x_t(0), x_t(1)$, $y_t(0,0)$, $y_t(1,1)$ and $y_t(0,1)$ do not exhibit oscillatory behavior.

The left panel in Figure G.1 illustrates the stationary average knowledge stock, $\bar{h}$, as a function of β together with the threshold β^c from Equation (51) in part (i) of Proposition 3. A significant fraction of firms will have knowledge of the technology once the spillover parameter β exceeds the critical value β^c. The same comparative statics as in Equation (20) hold for the critical level β^c. Further, from Equation (51) we find that the threshold β^c is lower in the multi-product competition case ($\theta = 0$) than in the independent markets case ($\theta = 1$). Hence, introducing product competition lowers the threshold above which innovation can take off in the economy. Moreover, from Equation

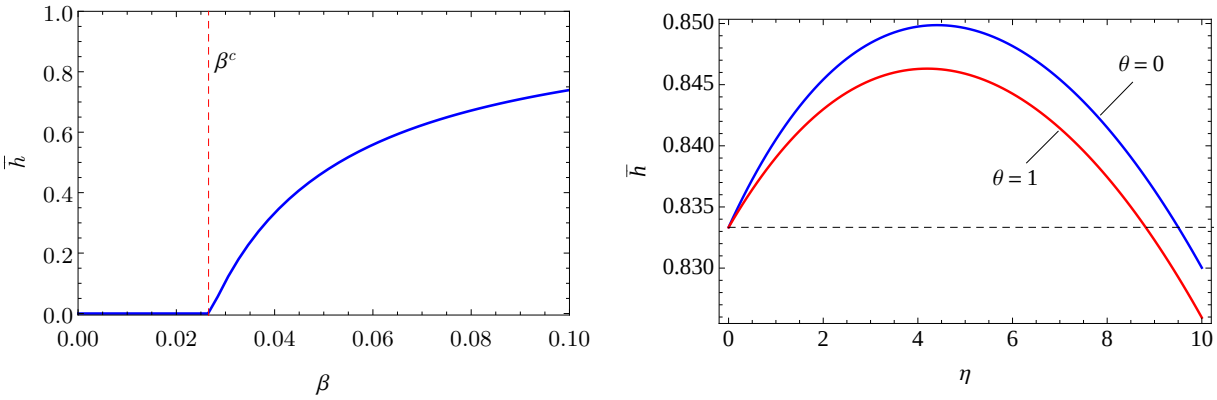


Figure G.1: (Left panel) Stationary average knowledge stock $\bar{h}$ as a function of β for $\gamma = 0$. The dashed line indicates the threshold β^c from Equation (51). (Right panel) The stationary average knowledge stock $\bar{h}$ as a function of η for $\theta \in \{0, 1\}$, $\lambda = 1$, $\beta = 30$, $b = 1$, $\tau = 0.01$, $c = 0.1$, $\xi = \chi = 0$ and $\gamma = 0$. The dashed line indicates the solution from Equation (54). The average knowledge stock in the multi-product competitive case ($\theta = 0$) is always higher than in the single-product monopolies case ($\theta = 1$).

(53) in part (ii) of Proposition 3 we find that the asymptotic average stocks of knowledge, $\bar{h} \equiv \lim_{t\to\infty} \bar{x}_t(1)$, satisfy: $\frac{\partial \bar{h}}{\partial \lambda} < 0$, $\frac{\partial \bar{h}}{\partial c} < 0$, $\frac{\partial \bar{h}}{\partial \beta} > 0$, $\frac{\partial \bar{h}}{\partial b} > 0$, and $\frac{\partial \bar{h}}{\partial \tau} > 0$.[53] The change in $\bar{h}$ with η is non-monotonic, and $\bar{h}$ being a concave function of η, where $\frac{\partial \bar{h}}{\partial \eta} > 0$ if $c\left(1 + e^{b\beta\eta\tau} + 2e^{c\eta}\right) < b\beta\tau\left(1 + e^{c\eta}\right)$ and $\frac{\partial \bar{h}}{\partial \eta} < 0$ otherwise. This is shown in the right panel of Figure G.1. The intuition behind this non-monotonicity is similar to the explanation for the convexity of the threshold β^c that we have given in the previous section (cf. Figure 5). When the noise is large (and η is small) then the average knowledge stock $\bar{h}$ is increasing with increasing values of η, because lower levels of noise (higher η) let firms form collaborations with other firms that possess complementary technology portfolios (which are more profitable). These collaborations typically involve larger firms with larger technology portfolios. When larger firms are preferentially selected for forming collaborations then this leads to more centralized network structures, typically exhibiting better diffusion properties [cf. Jackson and Rogers, 2007]. In contrast, when the level of noise falls even more, and firms form only the most profitable links, then due to the linking cost and the selectivity of the firms in terms of their collaboration partners, this leads to an overall decline in the number of collaborations that are being formed. The resulting network becomes increasingly sparse. This then weakens its diffusion properties and ultimately leads to a decline in the average stock of knowledge, $\bar{h}$. Moreover, as Figure G.1 illustrates, we find that the average knowledge stock, $\bar{h}$, in the multi-product competition case ($\theta = 0$) is higher than in the single-product monopolies case ($\theta = 1$). From Equation (54) in part (iii) of Proposition 3 we also find that $\frac{\partial \bar{h}}{\partial \gamma} > 0$.[53]

The cases (ii)–(iv) considered in Proposition 3 have in common that the solution trajectories do not exhibit any oscillatory behavior. However, such oscillations can be observed for intermediate levels of the R&D uncertainty in the multi-product competitive case. Such an oscillatory solution for $\bar{d}_t$ together with a numerical simulation of the stochastic process can be seen in Figure G.2.[54] A phase diagram showing the imaginary part of the eigenvalues of the Jacobian associated with the system of ODEs of Equations (10) and (11) when $N = 1$ for varying values of χ and β can be seen in in the left panel in Figure 7. Oscillations are indicated by the appearance of complex eigenvalues [cf. Khalil, 2002]. The phase diagram illustrates that only a mix of entry/exit ($\chi, \xi > 0$) and spillovers through link formation ($\beta > 0$) can give rise to oscillations. This can be explained as follows: When a new technology is discovered, there is a large market for this technology (where the average knowl-

[53] See also the proof of Proposition 3 in Appendix C.

[54] A numerical simulation of the stochastic process introduced in Sections 3.2, 3.3 and 3.4 can be implemented using the "next reaction method" for simulating a continuous time Markov chain [cf. Anderson, 2012]. However, such simulations become computationally infeasible for higher values of n and/or N.

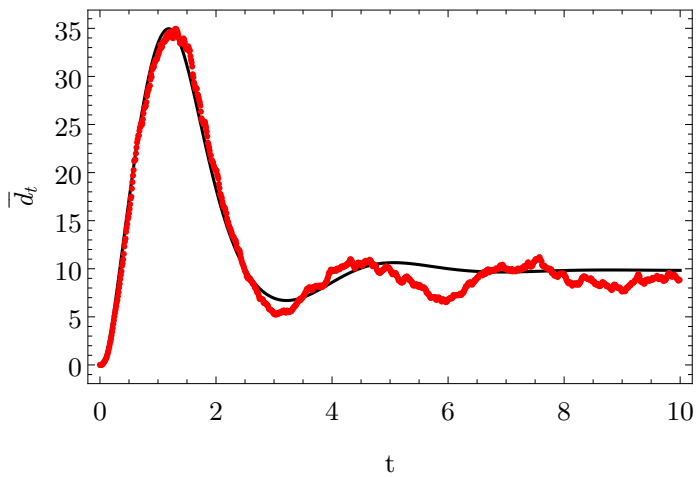


Figure G.2: The solution of the ODEs in Equations (10) and (11) for $N = 1$, respectively, for $\bar{d}_t = (2y_t(0,1)x_t(0)x_t(1) + y_t(0,0)x_t(0)^2 + y_t(1,1)x_t(1)^2)n$ with $n = 5000$, $c = 0$, $\beta = 0.03$, $b = 10$, $\eta = 10$, $\vartheta = 10$, $\gamma = 0.01$, $\lambda = 1$, $\xi = 1$, $\chi = 1$, $\kappa = 1$, $\rho = 10$, and $\tau = 0.01$. The red dots indicate a numerical simulation of the stochastic process averaged across 5000 simulation runs.

edge stock/productivity is low) and firms have strong incentives to enter and form collaborations which allow them to get access to the technology. However, once the technology has sufficiently diffused through the network, the market size shrinks (the average knowledge stock/productivity increases), and so do the incentives for market entry and the formation of collaborations, until a new innovation arrives. The economy then experiences periods of high collaborative activity followed by periods of low collaborative activity. The over- and undershooting of the average degree compared to the stationary state that we observe in these oscillations stems from the delay in the adjustment of the market (through entry and exit) to changes in the average technology stocks of the firms and the impact this has on the marginal profits from collaborations.

Proof of Proposition 3. In the following we consider the special case of a single technology, $N = 1$. From Equation (36) we then get

$$\begin{aligned} \frac{d\bar{x}_t(0)}{dt} &= -\gamma\bar{x}_t(0) + \lambda\bar{x}_t(1) - n\beta\bar{y}_t(0,1)\bar{x}_t(0)\bar{x}_t(1) - \xi\bar{x}_t(0) + \chi f^{\vartheta}_{\theta,\kappa}(\bar{x}_t(1))(1 - \bar{x}_t(0) - \bar{x}_t(1)), \\ \frac{d\bar{x}_t(1)}{dt} &= \gamma\bar{x}_t(0) - \lambda\bar{x}_t(1) + n\beta\bar{y}_t(0,1)\bar{x}_t(0)\bar{x}_t(1) - \xi\bar{x}_t(1), \end{aligned} \tag{55}$$

and from Equation (37) we get

$$\begin{aligned} \frac{d\bar{y}_t(0,1)}{dt} &= \rho g^{\eta}_{\theta,\tau}(0,1;\bar{x}_t(1)) - \rho(1 + g^{\eta}_{\theta,\tau}(0,1;\bar{x}_t(1)))\bar{y}_t(0,1) + o(\rho), \\ \frac{d\bar{y}_t(0,0)}{dt} &= \rho g^{\eta}_{\theta,\tau}(0,0;\bar{x}_t(1)) - \rho(1 + g^{\eta}_{\theta,\tau}(0,0;\bar{x}_t(1)))\bar{y}_t(0,1) + o(\rho), \\ \frac{d\bar{y}_t(1,1)}{dt} &= \rho g^{\eta}_{\theta,\tau}(1,1;\bar{x}_t(1)) - \rho(1 + g^{\eta}_{\theta,\tau}(1,1;\bar{x}_t(1)))\bar{y}_t(0,1) + o(\rho), \end{aligned} \tag{56}$$

where from Equation (8) it follows that

$$g^{\eta}_{\theta,\tau}(0,1;\bar{x}_t(1)) = \begin{cases} \Lambda^{\eta}\left((\beta\tau b - c) \cdot \Lambda^{\eta}(-c)\right. & \text{if } \theta = 1, \\ \Lambda^{\eta}\left(\frac{2\beta\tau b}{1+b\bar{x}_t(1)} - c\right) \cdot \Lambda^{\eta}(-c) & \text{if } \theta = 0, \end{cases}, \tag{57}$$

and $g^{\eta}_{\theta,\tau}(0,0;\bar{x}_t(1)) = g^{\eta}_{\theta,\tau}(1,1;\bar{x}_t(1)) = \Lambda^{\eta}(-c)^2$, for $\theta \in \{0,1\}$, with $\Lambda^{\eta} : x \mapsto e^{\eta x}/(1 + e^{\eta x})$ being the logistic function, and that

$$f^{\vartheta}_{\theta}(\bar{x}_t(1)) = \begin{cases} \Lambda^{\vartheta}(1 - \kappa) & \text{if } \theta = 1, \\ \Lambda^{\vartheta}\left(\frac{1}{1+b\bar{x}_t(1)} - \kappa\right) & \text{if } \theta = 0. \end{cases}$$

We now give a proof of part (i) of the proposition. We are considering a stationary state with $\bar{x}(1) = 0$ and assume that $\gamma = 0$. When $\bar{x}(1) = 0$ then the fixed point for $\bar{x}(0)$ solves $0 = f\chi(1-\bar{x}(0)) - \xi\bar{x}(0)$, from which we get $\bar{x}(0) = \frac{f\chi}{f\chi+\xi}$. Further, in the stationary state we must have that $\bar{y}_{01} = \frac{g_{01}}{1+g_{01}}$. The Jacobian is then given by

$$\mathbf{J} = \begin{bmatrix} -\xi - f\chi & \lambda + f\chi\left(-\frac{g_{01}n\beta}{(1+g_{01})(\xi+f\chi)} - 1\right) & 0 \\ 0 & -\lambda - \xi + \frac{fg_{01}n\beta\chi}{(1+g_{01})(\xi+f\chi)} & 0 \\ 0 & 0 & -(1+g_{01})\rho \end{bmatrix},$$

and the corresponding eigenvalues are $\mu_1 = -\frac{f\chi(g_{01}(\lambda-n\beta+\xi)+\lambda+\xi)+(1+g_{01})\xi(\lambda+\xi)}{(1+g_{01})(f\chi+\xi)}$, $\mu_2 = -(1+g_{01})\rho$ and $\mu_3 = -f\chi-\xi$. The eigenvalues μ_2 and μ_3 are real and negative by definition. The eigenvalue μ_1 is real and negative if $\beta < \beta^c \equiv \frac{1+g_{01}}{g_{01}}\frac{(\lambda+\xi)(f\chi+\xi)}{fn\chi}$. When $\bar{x}(1) = 0$ then $f = \frac{e^{(1-\kappa)\vartheta}}{e^{(1-\kappa)\vartheta}+1}$, so that we can write the threshold as $\beta^c = \frac{1+g_{01}}{g_{01}}\frac{(\lambda+\xi)\left(\xi+\chi+\xi e^{(\kappa-1)\vartheta}\right)}{n\chi}$. In the case of $\theta = 1$ (no competition) we have that $g_{01} = \frac{e^{c(-\eta)}e^{\eta(\beta b\tau-c)}}{\left(e^{c(-\eta)}+1\right)\left(e^{\eta(\beta b\tau-c)}+1\right)}$ so that $\frac{1+g_{01}}{g_{01}} = \frac{e^{b\beta\eta\tau}}{2e^{b\beta\eta\tau}+e^{\eta(b\beta\tau+c)}+e^{c\eta}+e^{2c\eta}}$, while in the case of $\theta = 0$ (competition) we get $g_{01} = \frac{e^{c(-\eta)}e^{\eta(2\beta b\tau-c)}}{\left(1+e^{c(-\eta)}\right)\left(1+e^{\eta(2\beta b\tau-c)}\right)}$ and therefore $\frac{1+g_{01}}{g_{01}} = \frac{e^{2b\beta\eta\tau}}{2e^{2b\beta\eta\tau}+e^{\eta(2b\beta\tau+c)}+e^{c\eta}+e^{2c\eta}}$. With $\theta \in \{0,1\}$ the two cases can be written more compactly as follows $\beta = \frac{(\lambda+\xi)\left(\xi+\chi+\xi e^{(\kappa-1)\vartheta}\right)}{n\chi}$ $\frac{e^{(2-\theta)b\beta\eta\tau}}{2e^{b\beta\eta\tau}+e^{\eta((2-\theta)b\beta\tau+c)}+e^{c\eta}+e^{2c\eta}}$. Next, assume also that there is no entry and exit, that is, we set $\xi = \chi = 0$. Then we must have that $x(0) = 1$ if $x(1) = 0$, and the Jacobian is given by

$$\mathbf{J} = \begin{bmatrix} 0 & \lambda - \frac{g_{01}n\beta}{g_{01}+1} & 0 \\ 0 & \frac{g_{01}n\beta}{g_{01}+1} - \lambda & 0 \\ 0 & 0 & -(1+g_{01})\rho \end{bmatrix}.$$

The corresponding eigenvalues are given by $\mu_1 = \frac{\beta g_{01}n}{g_{01}+1} - \lambda$, $\mu_2 = 0$ and $\mu_3 = -(1+g_{01})\rho$. The critical value for β such that the eigenvalue μ_1 is negative is then given by $\beta = \frac{(1+g_{01})\lambda}{g_{01}n}$. In the case of $\theta = 1$ this is $\beta = \frac{1}{n}\lambda e^{-b\beta\eta\tau}\left(2e^{b\beta\eta\tau} + e^{\eta(b\beta\tau+c)} + e^{c\eta} + e^{2c\eta}\right)$. Solving for β yields $\beta = \frac{1}{b\eta\tau}\, W\left(\frac{1}{n}b\eta\lambda\tau\left(e^{c\eta}+1\right)e^{\frac{\eta\left(cn-b\lambda\tau\left(e^{c\eta}+2\right)\right)}{n}}\right) + \frac{\lambda\left(e^{c\eta}+2\right)}{n}$, where $W(x)$ is the Lambert W function (or product-log), which is implicitly defined by $W(x)e^{W(x)} = x$. Similarly, in the case of $\theta = 0$ we get $\beta = \frac{1}{n}\lambda\left(e^{\eta(c-2b\beta\tau)}\left(e^{2b\beta\eta\tau} + e^{c\eta} + 1\right) + 2\right)$. Solving for β yields $\beta = \frac{\lambda(2+e^{c\eta})}{n} + \frac{1}{2b\eta\tau}$ $W\left(\frac{1}{n}2b\eta\lambda\tau\left(e^{c\eta}+1\right)e^{\frac{\eta\left(cn-2b\lambda\tau\left(e^{c\eta}+2\right)\right)}{n}}\right)$. With $\theta \in \{0,1\}$ the two cases can be written more compactly as follows $\beta = \frac{\lambda(2+e^{c\eta})}{n} + \frac{1}{b\eta(2-\theta)\tau}W\left(\frac{1}{n}b\eta(2-\theta)\lambda\tau\left(e^{c\eta}+1\right)\exp\left(\frac{\eta\left(cn-b(2-\theta)\lambda\tau\left(e^{c\eta}+2\right)\right)}{n}\right)\right)$.

We next give a proof of part (ii) of the proposition. The fixed point for the distribution of knowledge stocks is the solution of $0 = f\chi(1-\bar{x}(0)-\bar{x}(1)) - \beta n\bar{x}(0)\bar{x}(1)\bar{y}(0,1) - \gamma\bar{x}(0) - \xi\bar{x}(0) + \lambda\bar{x}(1)$ and $0 = \beta n\bar{x}(0)\bar{x}(1)\bar{y}(0,1) + \gamma\bar{x}(0) - \lambda\bar{x}(1) - \xi\bar{x}(1)$. Solving for $\bar{x}(0)$ and $\bar{x}(1)$ yields the admissible (non-negative) solution

$$\bar{x}(0) = \frac{\gamma\xi + \gamma f\chi + f\lambda\chi + f\xi\chi - S + \beta fn\chi y(0,1) + \lambda\xi + \xi^2}{2\beta ny(0,1)(f\chi+\xi)},$$

$$\bar{x}(1) = \frac{2\gamma f\chi}{\gamma\xi + \gamma f\chi + f\lambda\chi + f\xi\chi + S - \beta fn\chi y(0,1) + \lambda\xi + \xi^2},$$

where we have denoted by $S \equiv \sqrt{((\gamma+\lambda+\xi)(f\chi+\xi) + \beta fn\chi y(0,1))^2 - 4\beta fn\chi y(0,1)(\lambda+\xi)(f\chi+\xi)}$. Next, in the stationary state we have that

$$\bar{y}(0,1) = \frac{g_{01}}{1+g_{01}}, \tag{58}$$

and similarly, $\bar{y}(0,0)$ and $\bar{y}(1,1)$ can be computed. Inserting into $\bar{x}(0)$ and $\bar{x}(1)$ gives $\bar{x}(0) = \frac{T_+(1+g)}{2n\beta g(f\chi+\xi)}$,$\bar{x}(1) = \frac{2\gamma f\chi}{T_-}$, where we have denoted by

$$S \equiv \sqrt{\left((\gamma+\lambda+\xi)(f\chi+\xi) + \frac{\beta f g_{01} n\chi}{1+g_{01}}\right)^2 - \frac{4\beta f g_{01} n\chi(\lambda+\xi)(f\chi+\xi)}{1+g_{01}}},$$

and $T_\pm \equiv \xi(\gamma+\lambda+\xi) + \frac{f\chi(\gamma+g_{01}(\gamma+\lambda\pm\beta n+\xi)+\lambda+\xi)}{1+g_{01}} \mp S$.

Next, if also $\xi = \chi = 0$ then we have that $\bar{x}(0) = 1 - \bar{x}(1)$ and $\bar{x}(1)$ is the solution of $0 = \beta n(1-x(1))\bar{x}(1)\bar{y}(0,1) + \gamma(1-\bar{x}(1)) - \lambda\bar{x}(1)$. Solving for $\bar{x}(1)$ gives

$$\bar{x}(1) = \frac{-\gamma-\lambda+\sqrt{(\gamma+\lambda+\beta(-n)\bar{y}(0,1))^2 + 4\beta\gamma n\bar{y}(0,1)} + \beta n\bar{y}(0,1)}{2\beta n\bar{y}(0,1)},$$

where $\bar{y}(0,1)$ is given by Equation (58). Inserting yields

$$\bar{x}(1) = \frac{\beta gn - (g+1)(\gamma+\lambda) + \sqrt{((g+1)(\gamma+\lambda)-\beta gn)^2 + 4\beta\gamma g(g+1)n}}{2\beta gn}.$$

Next we analyze the stability of the fixed point. The Jacobian in the case of $\xi = \chi = 0$ is

$$\mathbf{J} = \begin{bmatrix} -nx(1)y(0,1)\beta - \gamma & -nx(0)y(0,1)\beta + \lambda & -nx(0)x(1)\beta \\ nx(1)y(0,1)\beta + \gamma & nx(0)y(0,1)\beta - \lambda & nx(0)x(1)\beta \\ 0 & 0 & -(1+g_{01})\rho \end{bmatrix},$$

with the eigenvalues $\mu_1 = -\gamma - \lambda + \beta n\bar{y}(0,1)(\bar{x}(0) - \bar{x}(1))$, $\mu_2 = -(1+g_{01})\rho$ and $\mu_3 = 0$. These eigenvalues are all real. The eigenvalue μ_1 is negative if $0 > -\gamma - \lambda + \beta n\bar{y}(0,1)(\bar{x}(0) - \bar{x}(1)) = -\frac{1}{1+g_{01}}\sqrt{(\gamma + g_{01}(\gamma+\lambda+\beta(-n)) + \lambda)^2 + 4\beta\gamma g_{01}(1+g_{01})n}$, which holds by definition. Moreover, in the limit of $\gamma \to 0$ we obtain the fixed point $\bar{x}(1) = 1 - \frac{1}{\beta n}\lambda(2 + e^{c\eta} + (1+e^{c\eta})\, e^{\eta(c-b\beta\tau)})$. Note that

$$\begin{aligned}
\frac{\partial\bar{x}(1)}{\partial\lambda} &= -\frac{e^{\eta(c-b\beta\tau)}\left(e^{b\beta\eta\tau} + e^{c\eta} + 1\right) + 2}{\beta n} < 0, \\
\frac{\partial\bar{x}(1)}{\partial c} &= -\frac{\eta\lambda e^{\eta(c-b\beta\tau)}\left(e^{b\beta\eta\tau} + 2e^{c\eta} + 1\right)}{\beta n} < 0, \\
\frac{\partial\bar{x}(1)}{\partial\beta} &= \frac{\lambda e^{-b\beta\eta\tau}\left(2e^{b\beta\eta\tau} + e^{c\eta}(b\beta\eta\tau+1) + e^{2c\eta}(b\beta\eta\tau+1) + e^{\eta(b\beta\tau+c)}\right)}{n\beta^2} > 0, \\
\frac{\partial\bar{x}(1)}{\partial b} &= \frac{\eta\lambda\tau\left(e^{c\eta}+1\right)e^{\eta(c-b\beta\tau)}}{n} > 0, \\
\frac{\partial\bar{x}(1)}{\partial\tau} &= \frac{b\eta\lambda\left(e^{c\eta}+1\right)e^{\eta(c-b\beta\tau)}}{n} > 0.
\end{aligned}$$

Moreover, we find that $\frac{\partial\bar{x}(1)}{\partial\eta} = \frac{1}{n\beta}\lambda e^{\eta(c-b\beta\tau)}\left(b\beta\tau\left(e^{c\eta}+1\right) - c\left(e^{b\beta\eta\tau} + 2e^{c\eta} + 1\right)\right)$, which is positive if $c(1 + e^{b\beta\eta\tau} + 2e^{c\eta}) < b\beta\tau\left(1 + e^{c\eta}\right)$.

We next provide a proof of part (iii) of the proposition. In the case of large shocks, where $\eta \to 0$ and $\vartheta \to 0$ we have that $g_{01} = \frac{1}{4}$, $y(0,1) = \frac{1}{5}$ and $f = \frac{1}{2}$ for both $\theta \in \{0,1\}$. Similarly, we find that $\bar{y}(0,0) = \bar{y}(1,1) = \frac{1}{5}$. The stationary state solves the equations $0 = \frac{1}{2}\chi(1 - \bar{x}(0) - \bar{x}(1)) - \beta n\bar{x}(0)\bar{x}(1)\frac{1}{5} + -\gamma\bar{x}(0) - \xi\bar{x}(0) + \lambda\bar{x}(1)$, and $0 = \beta n\bar{x}(0)\bar{x}(1)\frac{1}{5} + \gamma\bar{x}(0) - \lambda\bar{x}(1) - \xi\bar{x}(1)$. Solving for $\bar{x}(0)$ gives $\bar{x}(0) = \frac{5(2\xi+\chi)(\gamma+\lambda+\xi) - S + \beta n\chi}{2\beta n(2\xi+\chi)}$, and for $\bar{x}(1)$ we get $\bar{x}(1) = \frac{10\gamma\chi}{5(2\xi+\chi)(\gamma+\lambda+\xi) + S + \beta - n\chi}$, where we have denoted by $S \equiv \sqrt{25\gamma^2(2\xi+\chi)^2 + 10\gamma(2\xi+\chi)(5(\lambda+\xi)(2\xi+\chi) + \beta n\chi) + (\beta n\chi - 5(\lambda+\xi)(2\xi+\chi))^2}$.

The Jacobian is given by

$$\mathbf{J} = \begin{bmatrix} -\frac{1}{5}nx(1)\beta - \gamma - \xi - \frac{\chi}{2} & -\frac{1}{5}nx(0)\beta + \lambda - \frac{\chi}{2} & -nx(0)x(1)\beta \\ \frac{1}{5}nx(1)\beta + \gamma & \frac{1}{5}nx(0)\beta - \lambda - \xi & nx(0)x(1)\beta \\ 0 & 0 & -(1+g_{01})\rho \end{bmatrix},$$

with the eigenvalues $\mu_1 = \frac{1}{5}(\beta n(\bar{x}(0) - \bar{x}(1)) - 5(\gamma + \lambda + \xi))$, $\mu_2 = -(1+g_{01})\rho$ and $\mu_3 = -\xi - \frac{\chi}{2}$. The eigenvalues μ_2 and μ_3 are negative by definition. Inserting $\bar{x}(0)$ and $\bar{x}(1)$ from above shows that the eigenvalue μ_1 can be written as follows

$$\mu_1 = -\frac{\sqrt{25\gamma^2(2\xi+\chi)^2 + 10\gamma(2\xi+\chi)(5(\lambda+\xi)(2\xi+\chi) + \beta n\chi) + (\beta n\chi - 5(\lambda+\xi)(2\xi+\chi))^2}}{5(2\xi+\chi)},$$

which shows that also μ_1 is negative.

Next, if also $\xi = \chi = 0$ then the fixed point simplifies to $\bar{x}(1) = \frac{\beta n - 5(\gamma+\lambda) + \sqrt{(5(\gamma+\lambda)-\beta n)^2 + 20\beta\gamma n}}{2\beta n}$, while $\bar{x}(0) = 1 - \bar{x}(1)$. From the above equation we find that $\frac{\partial \bar{x}(1)}{\partial \gamma} = \frac{1}{2\beta n}$ $(5(\beta n + 5(\gamma+\lambda))$ $/((\beta n - 5(\gamma+\lambda))^2 + 20\beta n\gamma)^{1/2} - 5)$. This derivative is non-negative for all $\beta > 0$ and attains its unique maximum at $\beta = \frac{5(\gamma+\lambda)}{n}$.

We finally give a proof of part (iv) of the proposition. In the case of small shocks, when $\eta \to 0$ and $\vartheta \to 0$, we find that $g_{01} = 0$ and consequently $\bar{y}(0,1) = 0$ provided that $c > 0$. Similarly, we find that $\bar{y}(0,0) = \bar{y}(1,1) = 0$. Consider first the case of $\theta = 1$. Then $f = 1$ if $\kappa < 1$ and zero otherwise. If $f = 0$ then $\bar{x}(0) = \bar{x}(1) = 0$. If $f = 1$ then $\bar{x}(0)$ and $\bar{x}(1)$ are the solutions of $0 = -\gamma\bar{x}(0) - \xi\bar{x}(0) + \chi(1 - \bar{x}(0) - \bar{x}(1)) + \lambda\bar{x}(1)$, and $0 = \gamma\bar{x}(0) - \lambda\bar{x}(1) - \xi\bar{x}(1)$. Solving for $\bar{x}(0)$ and $\bar{x}(1)$ yields $\bar{x}(0) = \frac{\chi(\lambda+\xi)}{(\xi+\chi)(\gamma+\lambda+\xi)}$, $\bar{x}(1) = \frac{\gamma\chi}{(\xi+\chi)(\gamma+\lambda+\xi)}$. The Jacobian, evaluated at these fixed point, is given by

$$\mathbf{J} = \begin{bmatrix} -\gamma - \xi - \chi & \lambda - \chi & -\frac{n\beta\gamma(\lambda+\xi)\chi^2}{(\gamma+\lambda+\xi)^2(\xi+\chi)^2} \\ \gamma & -\lambda - \xi & \frac{n\beta\gamma(\lambda+\xi)\chi^2}{(\gamma+\lambda+\xi)^2(\xi+\chi)^2} \\ 0 & 0 & -\rho \end{bmatrix}.$$

The corresponding eigenvalues are all real, negative, and given by $-\gamma - \lambda - \xi$, $-\rho$, $-\xi - \chi$. If also $\xi = \chi = 0$ then we find from the above that $\bar{x}(1) = \frac{\gamma}{\gamma+\lambda}$ and $\bar{x}(0) = \frac{\lambda}{\gamma+\lambda}$. Next, we consider the case of $\theta = 0$. When $\kappa > \frac{1}{1+b\bar{x}(1)}$ then $f = 0$ and $f = 1$ otherwise. If $f = 0$ then $\bar{x}(0) = \bar{x}(1) = 0$. If $f = 1$ then we get the same solution for $\bar{x}(0)$ and $\bar{x}(1)$ as for $\theta = 1$. Moreover, the condition $\kappa < \frac{1}{1+b\bar{x}(1)}$ can then be written as follows $\frac{1}{\kappa} > 1 + \frac{b\gamma\chi}{(\xi+\chi)(\gamma+\lambda+\xi)}$. This completes the proof. □

G.2. Two Technologies ($N = 2$)

We next identify the stationary states of the stochastic process and their stability properties in the case of $N = 2$ technologies. We will see that similar results as in the previous section can be obtained, however, differently to the case of a single technology, here firms have incentives to form collaborations even in the absence of noise.

Proposition 4. *Consider large ρ such that terms of the order $o(\rho)$ can be neglected in Equation (11), $N = 2$ and denote by $\bar{x}(s) \equiv \lim_{t\to\infty} \bar{x}_t(s)$ and $\bar{y}(s,s') \equiv \lim_{t\to\infty} \bar{y}_t(s,s')$ the stationary state of the dynamic system in Equations (10) and (11) with $\bar{x}_t(s)$ and $\bar{y}_t(s,s')$ defined in Equations (13) and (14), respectively, for $0 \le s, s', \le 2$.*

(i) ***Threshold:*** *There exists a threshold such that $x_1 = x_2 = 0$ is a stable fixed point in the limit of $\gamma, \alpha \to 0$ if $\beta < \beta^c$, where the threshold value β^c is given by the solution to the following equation:*

$$\beta = \frac{\left(2 + e^{\eta(c-(2-\theta)b\beta\tau)}\left(1 + e^{(2-\theta)b\beta\eta\tau} + e^{c\eta}\right)\right)(\lambda+\xi)\left((2\gamma+\xi)e^{(\kappa-1)\vartheta} + 2\gamma + \xi + \chi\right)}{n\chi}, \tag{59}$$

for $\theta \in \{0,1\}$. If, in addition, we assume that $\xi = \chi = 0$ (no entry or exit) then we can compute β^c explicitly as given in Equation (51).

(ii) ***High uncertainty:*** *Assume that $\gamma = 0$. Then, in the limit of $\eta \to 0$, $\vartheta \to 0$ the stationary state of the dynamic system in Equation (10) with $\bar{x}(s)$ defined in Equation (13) is given by $\bar{x}(0) = \frac{\chi}{2\xi+\chi}$, $\bar{x}(1) = \bar{x}(2) = 0$ or*

$$\bar{x}(0) = \frac{R(\cdot) - 10\alpha(2\lambda - \xi)(2\xi + \chi)}{8\alpha\beta n(2\xi + \chi)}, \ \bar{x}(2) = \frac{(5(2\xi + \chi)(2\alpha - 4\lambda - 3\xi) + 2\beta n\chi + S(\cdot))T(\cdot)}{80\alpha\beta n(2\lambda + \xi)(2\xi + \chi)^2},$$
$$\bar{x}(1) = \frac{20\alpha^2(2\xi + \chi) + \alpha(20(3\lambda + \xi)(2\xi + \chi) + 4\beta n\chi - 2S(\cdot)) - R(\cdot)}{8\alpha\beta n(2\xi + \chi)},$$

with $R \equiv (2\lambda + \xi)\ (5\xi(2\xi + \chi) - 2\beta n\chi + S(\cdot))$ and functions $S(\cdot)$ and $T(\cdot)$ of the parameters whose expressions can be found in the proof of the proposition. When also $\xi = \chi = 0$ (no entry or exit) then the stationary state is given by $\bar{x}(1) = \bar{x}(2) = 0$, or

$$\bar{x}(1) = \frac{5\alpha^2 + 15\alpha\lambda + \alpha\beta n + \beta\lambda n - S(\cdot)(\alpha + \lambda)}{2\alpha\beta n}, \ \bar{x}(2) = \frac{(\lambda + 2\alpha)S(\cdot) - 10\alpha^2 - 25\alpha\lambda - \beta\lambda n}{2\alpha\beta n},$$

where we have denoted by $S \equiv \sqrt{100\alpha\lambda + (5\alpha + \beta n)^2}$, and $\bar{x}(0) = 1 - 2\bar{x}(1) - \bar{x}(2)$, while the stationary state of Equation (11) with $\bar{y}(s, s')$ defined in Equation (14) is given by $\bar{y}(0, 1) = \bar{y}(0, 2) = \bar{y}(1, 1) = \bar{y}(1, 2) = \frac{1}{5}$. Moreover, when $\lambda \geq \alpha$ then the Jacobian has only real eigenvalues and the solution trajectories for $\bar{x}_t(s)$ and $\bar{y}_t(s, s')$, $s, s' \in \{0, 1, 2\}$, do not exhibit oscillatory behavior.

(iii) ***Low uncertainty:*** *In the limit of $\eta \to \infty$, $\vartheta \to \infty$ starting from an empty graph, $\overline{K}_n$, the stationary state of Equation (11) with $\bar{y}(s, s')$ defined in Equation (14) is given by $\bar{y}(0, 1) = \bar{y}(0, 2) = \bar{y}(1, 2) = 0$ and $\bar{y}(1, 1) = \frac{1}{3}$, while the stationary state of the dynamic system in Equation (10) with $\bar{x}(s)$ defined in Equation (13) is given by*

$$\bar{x}(0) = \frac{\beta n\xi\chi(2\gamma + \xi) + \sqrt{3}\lambda S(\cdot)}{\beta n(2\gamma + \xi)^2(\xi + \chi)} - \frac{3\lambda(\alpha(2\gamma + \xi) + (\gamma + \lambda + \xi)(2(\gamma + \lambda) + \xi))}{\beta n(2\gamma + \xi)^2},$$
$$\bar{x}(1) = \frac{\sqrt{3}S(\cdot)}{2\beta n(2\gamma + \xi)(\xi + \chi)} - \frac{3(\alpha(2\gamma + \xi) + (\gamma + \lambda + \xi)(2(\gamma + \lambda) + \xi))}{2\beta n(2\gamma + \xi)}, \ \bar{x}(2) = \frac{T(\cdot)}{\beta n(2\gamma + \xi)^2(\xi + \chi)},$$

with functions $S(\cdot)$ and $T(\cdot)$ of the parameters whose expressions can be found in the proof of the proposition, and we assume that $\beta b\tau > c$, $\kappa < 1$ in the case of $\theta = 1$ and $\frac{2\beta b(b+1)\tau}{2b(\bar{x}(1)+\bar{x}(2))+1} > c$, $\kappa < \frac{1}{2b(\bar{x}(1)+\bar{x}(2))+1}$ in the case of $\theta = 0$. If we further assume that $\xi = \chi = 0$ (no entry or exit) then

$$\bar{x}(1) = \frac{S_1(\cdot) - 3\alpha\gamma - 3(\gamma + \lambda)^2}{2\beta\gamma n}, \ \bar{x}(2) = \frac{3\alpha\gamma(2\gamma + \lambda) + 2\beta\gamma^2 n + (2\gamma + \lambda)\left(3(\gamma + \lambda)^2 - S_2(\cdot)\right)}{2\beta\gamma^2 n},$$

with functions $S_{1,2}(\cdot)$ of the parameters whose expressions can be found in the proof of the proposition. Moreover, in the case of $\theta = 1$ (no competition) the Jacobian has only real eigenvalues, so that the solution trajectories for $\bar{x}_t(s)$ and $\bar{y}_t(s, s')$, $s, s' \in \{0, 1, 2\}$, do not exhibit oscillatory behavior.

A significant fraction of firms has on average a positive technology portfolio size once the spillover parameter β exceeds the critical value β^c in part (i) of Proposition 4. In contrast to the case of $N = 1$ where the parameter α did not play any role, here we find in parts (ii) and (iii) of Proposition 4 that in both cases, high and low uncertainty, we have that $\frac{\partial \bar{h}}{\partial \alpha} > 0$. Note also that, differently to the case of $N = 1$, the links between firms do not cease to exist in the limit of vanishing uncertainty in part (iii) of the proposition as $\eta \to \infty$. Hence, increased in-house R&D capabilities lead to higher average stocks of knowledge, irrespective of the uncertainty involved in R&D collaborations.

Proof of Proposition 4. Note that when $N = 2$ with $s \in \{0, 1, 2\}$ we obtain from Equation (36)

$$\begin{aligned}
\frac{d\bar{x}_t(0)}{dt} &= 2\left(\lambda\bar{x}_t(1) - \gamma\bar{x}_t(0) - \beta(\bar{y}_t(0,1)\bar{x}_t(0)\bar{x}_t(1) + \bar{y}_t(0,2)\bar{x}_t(0)\bar{x}_t(2))\right) - \xi\bar{x}_t(0) \\
&\quad + \chi f^{\vartheta}_{\theta,\kappa}(\bar{h}_t)\left(1 - \bar{x}_t(0) - 2\bar{x}_t(1) - \bar{x}_t(2)\right) \\
\frac{d\bar{x}_t(1)}{dt} &= \gamma\bar{x}_t(0) + \lambda\bar{x}_t(2) - (\lambda + \gamma + \alpha)\bar{x}_t(1) + \beta\left(\bar{y}_t(0,1)\bar{x}_t(0)\bar{x}_t(1) - \bar{y}_t(1,1)\bar{x}_t(1)^2\right. \\
&\quad \left. + \bar{y}_t(0,2)\bar{x}_t(0)\bar{x}_t(2) - \bar{y}_t(1,2)\bar{x}_t(1)\bar{x}_t(2)\right) - \xi\bar{x}_t(1) \\
\frac{d\bar{x}_t(2)}{dt} &= 2\left((\gamma + \alpha)\bar{x}_t(1) - \lambda\bar{x}_t(2) + \beta\left(\bar{y}_t(1,1)\bar{x}_t(1)^2 + \bar{y}_t(1,2)\bar{x}_t(1)\bar{x}_t(2)\right)\right) - \xi\bar{x}_t(0),
\end{aligned} \tag{60}$$

and from Equation (37) we obtain

$$\begin{aligned}
\frac{d\bar{y}_t(0,1)}{dt} &= \frac{1}{2}\rho\tilde{g}(0,1) - \rho\left(1 + \frac{1}{2}\tilde{g}(0,1)\right)\bar{y}_t(0,1) \\
\frac{d\bar{y}_t(0,2)}{dt} &= \rho\tilde{g}(0,2) - \rho\left(1 + \tilde{g}(0,2)\right)\bar{y}_t(0,2) \\
\frac{d\bar{y}_t(1,1)}{dt} &= \frac{1}{4}\rho\tilde{g}(1,1) - \rho\left(1 + \frac{1}{4}\tilde{g}(1,1)\right)\bar{y}_t(1,1) \\
\frac{d\bar{y}_t(1,2)}{dt} &= \frac{1}{2}\rho\tilde{g}(1,2) - \rho\left(1 + \frac{1}{2}\tilde{g}(1,2)\right)\bar{y}_t(1,2).
\end{aligned} \tag{61}$$

From the definition in Equation (15) we find that

$$\tilde{g}(0,1) = \sum_{\substack{\mathbf{h}=(0,0)^\top \\ \mathbf{h}'\in\{(0,1)^\top,(1,0)^\top\}}} \frac{e^{\eta(\beta g_{\theta,\tau}(\mathbf{x})(1+b|\mathsf{S}(\mathbf{h})|)^{1-\theta}\langle \mathbf{h}^c,\mathbf{h}'\rangle - c)}}{1+e^{\eta(\beta g_{\theta,\tau}(\mathbf{x})(1+b|\mathsf{S}(\mathbf{h})|)^{1-\theta}\langle \mathbf{h}^c,\mathbf{h}'\rangle - c)}} \frac{e^{\eta(\beta g_{\theta,\tau}(\mathbf{x})(1+b|\mathsf{S}(\mathbf{h}')|)^{1-\theta}\langle \mathbf{h}'^c,\mathbf{h}\rangle - c)}}{1+e^{\eta(\beta g_{\theta,\tau}(\mathbf{x})(1+b|\mathsf{S}(\mathbf{h}')|)^{1-\theta}\langle \mathbf{h}'^c,\mathbf{h}\rangle - c)}}$$

$$= \begin{cases} 2\frac{e^{-\eta c}}{1+e^{-\eta c}}\frac{e^{\eta(\beta g_{1,\tau}-c)}}{1+e^{\eta(\beta g_{1,\tau}-c)}} = 2\frac{e^{-\eta c}}{1+e^{-\eta c}}\frac{e^{\eta(\beta b\tau - c)}}{1+e^{\eta(\beta b\tau - c)}} & \text{if } \theta = 1, \\ 2\frac{e^{-\eta c}}{1+e^{-\eta c}}\frac{e^{\eta(\beta g_{0,\tau}(\mathbf{x})-c)}}{1+e^{\eta(\beta g_{0,\tau}(\mathbf{x})-c)}} = 2\frac{e^{-\eta c}}{1+e^{-\eta c}}\frac{e^{\eta(\beta\frac{2\tau b}{1+bh_t(\mathbf{x})}-c)}}{1+e^{\eta(\beta\frac{2\tau b}{1+bh_t(\mathbf{x})}-c)}} & \text{if } \theta = 0, \end{cases}$$

$$\tilde{g}(0,2) = \sum_{\substack{\mathbf{h}=(0,0)^\top \\ \mathbf{h}'=(1,1)^\top}} \frac{e^{\eta(\beta g_{\theta,\tau}(\mathbf{x})(1+b|\mathsf{S}(\mathbf{h})|)^{1-\theta}\langle \mathbf{h}^c,\mathbf{h}'\rangle - c)}}{1+e^{\eta(\beta g_{\theta,\tau}(\mathbf{x})(1+b|\mathsf{S}(\mathbf{h})|)^{1-\theta}\langle \mathbf{h}^c,\mathbf{h}'\rangle - c)}} \frac{e^{\eta(\beta g_{\theta,\tau}(\mathbf{x})(1+b|\mathsf{S}(\mathbf{h}')|)^{1-\theta}\langle \mathbf{h}'^c,\mathbf{h}\rangle - c)}}{1+e^{\eta(\beta g_{\theta,\tau}(\mathbf{x})(1+b|\mathsf{S}(\mathbf{h}')|)^{1-\theta}\langle \mathbf{h}'^c,\mathbf{h}\rangle - c)}}$$

$$= \begin{cases} \frac{e^{-\eta c}}{1+e^{-\eta c}}\frac{e^{\eta(2\beta g^n_{1,\tau}-c)}}{1+e^{\eta(2\beta g^n_{1,\tau}-c)}} = \frac{e^{-\eta c}}{1+e^{-\eta c}}\frac{e^{\eta(2\beta b\tau - c)}}{1+e^{\eta(2\beta b\tau - c)}} & \text{if } \theta = 1, \\ \frac{e^{-\eta c}}{1+e^{-\eta c}}\frac{e^{\eta(2\beta g_{0,\tau}(\mathbf{x})-c)}}{1+e^{\eta(2\beta g_{0,\tau}(\mathbf{x})-c)}} = \frac{e^{-\eta c}}{1+e^{-\eta c}}\frac{e^{\eta(2\beta\frac{2\tau b}{1+bh_t(\mathbf{x})}-c)}}{1+e^{\eta(2\beta\frac{2\tau b}{1+bh_t(\mathbf{x})}-c)}} & \text{if } \theta = 0, \end{cases}$$

$$\tilde{g}(1,1) = \sum_{\substack{\mathbf{h}\in\{(0,1)^\top,(1,0)^\top\} \\ \mathbf{h}'\in\{(0,1)^\top,(1,0)^\top\}}} \frac{e^{\eta(\beta g_{\theta,\tau}(\mathbf{x})(1+b|\mathsf{S}(\mathbf{h})|)^{1-\theta}\langle \mathbf{h}^c,\mathbf{h}'\rangle - c)}}{1+e^{\eta(\beta g_{\theta,\tau}(\mathbf{x})(1+b|\mathsf{S}(\mathbf{h})|)^{1-\theta}\langle \mathbf{h}^c,\mathbf{h}'\rangle - c)}} \frac{e^{\eta(\beta g_{\theta,\tau}(\mathbf{x})(1+b|\mathsf{S}(\mathbf{h}')|)^{1-\theta}\langle \mathbf{h}'^c,\mathbf{h}\rangle - c)}}{1+e^{\eta(\beta g_{\theta,\tau}(\mathbf{x})(1+b|\mathsf{S}(\mathbf{h}')|)^{1-\theta}\langle \mathbf{h}'^c,\mathbf{h}\rangle - c)}}$$

$$= \begin{cases} 2\left(\frac{e^{-2\eta c}}{(1+e^{-\eta c})^2} + \frac{e^{2\eta(\beta g_{1,\tau}-c)}}{(1+e^{\eta(\beta g_{1,\tau}-c)})^2}\right) = 2\left(\frac{e^{-2\eta c}}{(1+e^{-\eta c})^2} + \frac{e^{2\eta(\beta b\tau - c)}}{(1+e^{\eta(\beta b\tau - c)})^2}\right) & \text{if } \theta = 1, \\ 2\left(\frac{e^{-2\eta c}}{(1+e^{-\eta c})^2} + \frac{e^{2\eta((1+b)\beta g_{0,\tau}(\mathbf{x})-c)}}{(1+e^{\eta((1+b)\beta g_{0,\tau}(\mathbf{x})-c)})^2}\right) = 2\left(\frac{e^{-2\eta c}}{(1+e^{-\eta c})^2} + \frac{e^{2\eta((1+b)\beta\frac{2\tau b}{1+bh_t(\mathbf{x})}-c)}}{(1+e^{\eta((1+b)\beta\frac{2\tau b}{1+bh_t(\mathbf{x})}-c)})^2}\right) & \text{if } \theta = 0, \end{cases}$$

$$\tilde{g}(1,2) = \sum_{\substack{\mathbf{h}\in\{(0,1)^\top,(1,0)^\top\} \\ \mathbf{h}'=(1,1)^\top}} \frac{e^{\eta(\beta g_{\theta,\tau}(\mathbf{x})(1+b|\mathsf{S}(\mathbf{h})|)^{1-\theta}\langle \mathbf{h}^c,\mathbf{h}'\rangle - c)}}{1+e^{\eta(\beta g_{\theta,\tau}(\mathbf{x})(1+b|\mathsf{S}(\mathbf{h})|)^{1-\theta}\langle \mathbf{h}^c,\mathbf{h}'\rangle - c)}} \frac{e^{\eta(\beta g_{\theta,\tau}(\mathbf{x})(1+b|\mathsf{S}(\mathbf{h}')|)^{1-\theta}\langle \mathbf{h}'^c,\mathbf{h}\rangle - c)}}{1+e^{\eta(\beta g_{\theta,\tau}(\mathbf{x})(1+b|\mathsf{S}(\mathbf{h}')|)^{1-\theta}\langle \mathbf{h}'^c,\mathbf{h}\rangle - c)}}$$

$$= \begin{cases} 2\frac{e^{-\eta c}}{1+e^{-\eta c}}\frac{e^{\eta(\beta g_{0,\tau}-c)}}{1+e^{\eta(\beta g_{0,\tau}-c)}} = 2\frac{e^{-\eta c}}{1+e^{-\eta c}}\frac{e^{\eta(\beta b\tau - c)}}{1+e^{\eta(\beta b\tau - c)}} & \text{if } \theta = 1, \\ 2\frac{e^{-\eta c}}{1+e^{-\eta c}}\frac{e^{\eta((1+b)\beta g_{0,\tau}(\mathbf{x})-c)}}{1+e^{\eta((1+b)\beta g_{0,\tau}(\mathbf{x})-c)}} = 2\frac{e^{-\eta c}}{1+e^{-\eta c}}\frac{e^{\eta((1+b)\beta\frac{2\tau b}{1+bh_t(\mathbf{x})}-c)}}{1+e^{\eta((1+b)\beta\frac{2\tau b}{1+bh_t(\mathbf{x})}-c)}} & \text{if } \theta = 0, \end{cases}$$

where the average stock of knowledge is given by $\bar{h}_t(\mathbf{x}) = 2(\bar{x}_t(1) + \bar{x}_t(2))$. Note that in the limit of $\eta \to 0$ we obtain $\lim_{\eta\to 0}\tilde{g}(0,1) = \frac{1}{2}$, $\lim_{\eta\to 0}\tilde{g}(0,2) = \frac{1}{4}$, $\lim_{\eta\to 0}\tilde{g}(1,1) = 1$, $\lim_{\eta\to 0}\tilde{g}(1,2) = \frac{1}{2}$, while in the limit of $\eta \to \infty$ we obtain $\lim_{\eta\to\infty}\tilde{g}(0,1) = \lim_{\eta\to\infty}\tilde{g}(0,2) = \lim_{\eta\to\infty}\tilde{g}(1,2) = 0$, and

$$\lim_{\eta\to\infty}\tilde{g}(1,1) = 2 \times \begin{cases} \mathbb{1}_{\{\beta b\tau > c\}} & \text{if } \theta = 1, \\ \mathbb{1}_{\{\frac{2b(b+1)\beta\tau}{4b+1} > c\}} & \text{if } \theta = 0. \end{cases}$$

Further, we from Equation (9) have that

$$f^{\vartheta}_{\theta,\kappa}(\bar{h}_t) \equiv \frac{e^{\vartheta\left(\theta + \frac{1-\theta}{1+bh_t} - \kappa\right)}}{1+e^{\vartheta\left(\theta + \frac{1-\theta}{1+bh_t} - \kappa\right)}} = \begin{cases} \frac{e^{\vartheta(1-\kappa)}}{1+e^{\vartheta(1-\kappa)}} & \text{if } \theta = 1, \\ \frac{e^{\vartheta\left(\frac{1}{1+2b(\bar{x}_t(1)+\bar{x}_t(2))} - \kappa\right)}}{1+e^{\vartheta\left(\frac{1}{1+2b(\bar{x}_t(1)+\bar{x}_t(2))} - \kappa\right)}} & \text{if } \theta = 0. \end{cases}$$

We now proof part (i) of the proposition. Denote by $x_0 \equiv \lim_{t\to\infty}\bar{x}_t(0), x_1 \equiv \lim_{t\to\infty}\bar{x}_t(1), x_2 \equiv \lim_{t\to\infty}\bar{x}_t(2)$ and $y_{01} \equiv \lim_{t\to\infty}\bar{y}_t(0,1), y_{02} \equiv \lim_{t\to\infty}\bar{y}_t(0,2), y_{11} \equiv \lim_{t\to\infty}\bar{y}_t(1,1), y_{12} \equiv \lim_{t\to\infty}\bar{y}_t(1,2)$. The fixed point $x_1 = x_2 = 0$ is stable if the Jacobian has only negative eigenvalues. The Jacobian

is given by $\mathbf{J} = \begin{bmatrix} \mathbf{J}_{11} & \mathbf{J}_{12} \\ \mathbf{J}_{21} & \mathbf{J}_{22} \end{bmatrix}$, with

$$\mathbf{J}_{12} = \begin{bmatrix} -2nx_0x_1\beta & -2nx_0x_2\beta & 0 & 0 \\ nx_0x_1\beta & nx_0x_2\beta & -nx_1^2\beta & -nx_1x_2\beta \\ 0 & 0 & 2nx_1^2\beta & 2nx_1x_2\beta \end{bmatrix}$$

$$\mathbf{J}_{22} = \begin{bmatrix} -\rho\left(1+\frac{1}{2}\tilde{g}_{01}\right) & 0 & 0 & 0 \\ 0 & -\rho(1+\tilde{g}_{02}) & 0 & 0 \\ 0 & 0 & -\rho\left(1+\frac{1}{4}\tilde{g}_{11}\right) & 0 \\ 0 & 0 & 0 & -\rho\left(1+\frac{1}{2}\tilde{g}_{12}\right) \end{bmatrix},$$

and $\mathbf{J}_{11} = \mathbf{J}_{11}^{\mathrm{U}} + \mathbf{J}_{11}^{\mathrm{L}} + \mathbf{J}_{11}^{\mathrm{D}}$, where

$$\mathbf{J}_{11}^{\mathrm{U}} = \begin{bmatrix} 0 & -2\beta nx_0y_{01} + 2\lambda - 2f\chi - x_0 + 2x_1 + x_2 - \chi\frac{\partial f}{\partial x_1} & -2\beta nx_0y_{02} - f\chi - x_0 + 2x_1 + x_2 - \chi\frac{\partial f}{\partial x_2} \\ 0 & 0 & n(x_0y_{02} - x_1y_{12})\beta + \lambda \\ 0 & 0 & 0 \end{bmatrix},$$

$$\mathbf{J}_{11}^{\mathrm{L}} = \begin{bmatrix} 0 & 0 & 0 \\ n(x_1y_{01} + x_2y_{02})\beta + \gamma & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

$$\mathbf{J}_{11}^{\mathrm{D}} = \begin{bmatrix} -2n(x_1y_{01} + x_2y_{02})\beta - 2\gamma - \xi - f\chi & 0 & 0 \\ 0 & n(x_0y_{01} - 2x_1y_{11} - x_2y_{12})\beta - \gamma - \lambda - \xi - \alpha & 0 \\ 0 & 0 & 2nx_1y_{12}\beta - 2\lambda - \xi \end{bmatrix}.$$

In the case of $\theta = 0$ we further have that

$$\mathbf{J}_{21} = \begin{bmatrix} 0 & \frac{4b^2e^{\eta\left(\frac{2b\beta\tau}{2b(x_1+x_2)+1}-c\right)}(z_1-1)\beta\eta\rho\tau}{(1+e^{c\eta})\left(1+e^{\eta\left(\frac{2b\beta\tau}{2b(x_1+x_2)+1}-c\right)}\right)^2(2b(x_1+x_2)+1)^2} & \frac{4b^2e^{\eta\left(\frac{2b\beta\tau}{2b(x_1+x_2)+1}-c\right)}(z_1-1)\beta\eta\rho\tau}{(1+e^{c\eta})\left(1+e^{\eta\left(\frac{2b\beta\tau}{2b(x_1+x_2)+1}-c\right)}\right)^2(2b(x_1+x_2)+1)^2} \\ 0 & \frac{4b^2e^{\eta\left(\frac{4b\beta\tau}{2b(x_1+x_2)+1}-c\right)}(z_2-1)\beta\eta\rho\tau}{(1+e^{c\eta})\left(1+e^{\frac{4b\beta\eta\tau}{2b(x_1+x_2)+1}}\right)^2(2b(x_1+x_2)+1)^2} & \frac{4b^2e^{\eta\left(\frac{4b\beta\tau}{2b(x_1+x_2)+1}-c\right)}(z_2-1)\beta\eta\rho\tau}{(1+e^{c\eta})\left(1+e^{\frac{4b\beta\eta\tau}{2b(x_1+x_2)+1}}\right)^2(2b(x_1+x_2)+1)^2} \\ 0 & \frac{8b^2(b+1)\exp\left(2\eta\left(\frac{2b(b+1)\beta\tau}{2b(x_1+x_2)+1}-c\right)\right)(z_3-1)\beta\eta\rho\tau}{\left(1+\exp\left(\eta\left(\frac{2b(b+1)\beta\tau}{2b(x_1+x_2)+1}-c\right)\right)\right)^3(2b(x_1+x_2)+1)^2} & \frac{8b^2(b+1)\exp\left(2\eta\left(\frac{2b(b+1)\beta\tau}{2b(x_1+x_2)+1}-c\right)\right)(z_3-1)\beta\eta\rho\tau}{\left(1+\exp\left(\eta\left(\frac{2b(b+1)\beta\tau}{2b(x_1+x_2)+1}-c\right)\right)\right)^3(2b(x_1+x_2)+1)^2} \\ 0 & \frac{4b^2(b+1)\exp\left(\eta\left(\frac{2b(b+1)\beta\tau}{2b(x_1+x_2)+1}-c\right)\right)(z_4-1)\beta\eta\rho\tau}{(1+e^{c\eta})\left(1+\exp\left(\eta\left(\frac{2b(b+1)\beta\tau}{2b(x_1+x_2)+1}-c\right)\right)\right)^2(2b(x_1+x_2)+1)^2} & \frac{4b^2(b+1)\exp\left(\eta\left(\frac{2b(b+1)\beta\tau}{2b(x_1+x_2)+1}-c\right)\right)(z_4-1)\beta\eta\rho\tau}{(1+e^{c\eta})\left(1+\exp\left(\eta\left(\frac{2b(b+1)\beta\tau}{2b(x_1+x_2)+1}-c\right)\right)\right)^2(2b(x_1+x_2)+1)^2} \end{bmatrix},$$

while in the case of $\theta = 1$ we have that

$$\mathbf{J}_{21} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Note that we obtain the same result for $\theta = 0$ in the special case of $\eta \to \infty$ and $c > \frac{2b(b+1)\beta\tau}{4b+1}$ as well as $\eta \to 0$ or $\tau \to 0$. Moreover, in both cases, evaluated at $x_1 = x_2 = 0$ we obtain

$$\mathbf{J}_{12} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Hence, in both cases when $x_1 = x_2 = 0$ we have that $\det(\mathbf{J} - \mu\mathbf{I}_7) = \det(\mathbf{J}_{11} - \mu\mathbf{I}_3)\det(\mathbf{J}_{22} - \mu\mathbf{I}_4)$.[55] We have that $\det(\mathbf{J}_{22} - \mu\mathbf{I}_4) = \frac{1}{16}(2\mu + \tilde{g}_{01}\rho)(\mu + \tilde{g}_{02}\rho)(4\mu + \tilde{g}_{11}\rho)(2\mu + \tilde{g}_{12}\rho)$, The roots of the characteristic polynomial give us the eigenvalues $\mu_3 = -\rho\left(1 + \frac{1}{2}\tilde{g}_{01}\right)$, $\mu_4 = -\rho(1 + \tilde{g}_{02})$, $\mu_5 = -\rho\left(1 + \frac{1}{4}\tilde{g}_{11}\right)$, $\mu_6 = -\rho\left(1 + \frac{1}{2}\tilde{g}_{12}\right)$. These are all negative an real. Further we have that at $x_1 = x_2 = 0$

$$\begin{aligned}\det(\mathbf{J}_{11} - \mu\mathbf{I}_3) = &(-2\lambda - \mu - \xi)(\gamma(\chi(2f + \frac{\partial f}{\partial x_1}(x_0 - 1)) - 2\lambda + 2\beta n x_0 y_{01}) \\ &+ (2\gamma + f\chi + \mu + \xi)(\alpha + \gamma + \lambda + \mu - \beta n x_0 y_{01} + \xi)) \\ &+ 2(\alpha + \gamma)(2\gamma\lambda - \gamma\chi(f + \frac{\partial f}{\partial x_2}(x_0 - 1)) + (f\chi + \mu + \xi)(\lambda + \beta n x_0 y_{02})).\end{aligned}$$

Setting $\alpha = \gamma = 0$ this simplifies to $\det(\mathbf{J}_{11} - \mu\mathbf{I}_3) = (2\lambda + \mu + \xi)(f\chi + \mu + \xi)(\beta n x_0 y_{01} - \lambda - \mu - \xi)$. The roots are given by $\mu_1 = -\lambda + \beta n x_0 y_{01} - \xi$, $\mu_2 = -2\lambda - \xi$ and $\mu_3 = -f\chi - \xi$. Note that both μ_2 and μ_3 are real and negative by definition. We next show that also μ_1 is negative if β is below a threshold value. When $x(1) = x(2) = 0$ then from Equation (60) we see that the fixed point for $x(0)$ solves $0 = f\chi(1 - x_0) - 2\gamma x_0 - \xi x_0$, from which we get $x(0) = \frac{f\chi}{2\gamma + f\chi + \xi}$. Together with the stationary state from Equation (61), $y_{01} = \frac{\tilde{g}_{01}}{2 + \tilde{g}_{01}}$, we then get that $\mu_1 < 0$ if $\beta < \beta^c \equiv \frac{2 + \tilde{g}(0,1)}{\tilde{g}(0,1)} \frac{(\lambda + \xi)(2\gamma + f\chi + \xi)}{f n \chi}$. Inserting f and $\tilde{g}_{01}$ then yields in the case of $\theta = 0$ $\beta = \left(2 + e^{\eta(c - 2b\beta\tau)}\left(1 + e^{2b\beta\eta\tau} + e^{c\eta}\right)\right) \frac{(\lambda + \xi)\left((2\gamma + \xi)e^{(\kappa - 1)\vartheta} + 2\gamma + \xi + \chi\right)}{n\chi}$, while in the case of $\theta = 1$ we get $\beta = \left(2 + e^{\eta(c - b\beta\tau)}\left(1 + e^{b\beta\eta\tau} + e^{c\eta}\right)\right) \frac{(\lambda + \xi)\left((2\gamma + \xi)e^{(\kappa - 1)\vartheta} + 2\gamma + \xi + \chi\right)}{n\chi}$. With $\theta \in \{0, 1\}$ the two cases can be written more compactly as follows $\beta = \left(2 + e^{\eta(c - (2-\theta)b\beta\tau)}\left(1 + e^{(2-\theta)b\beta\eta\tau} + e^{c\eta}\right)\right) \frac{(\lambda + \xi)\left((2\gamma + \xi)e^{(\kappa - 1)\vartheta} + 2\gamma + \xi + \chi\right)}{n\chi}$.

We next consider the case of $\xi = \chi = 0$ (no entry/exit). Then we have that $x_0 = 1 - 2x_1 - x_2$, and we only need to keep track of x_1 and x_2. The Jacobian is given by $\mathbf{J} = \begin{bmatrix}\mathbf{J}_{11} & \mathbf{J}_{12} \\ \mathbf{J}_{21} & \mathbf{J}_{22}\end{bmatrix}$, with

$$\mathbf{J}_{11} = \begin{bmatrix}(1 - 4x_1 - x_2)y_{01}n\beta - 2n\beta(x_2 y_{02} + x_1 y_{11}) - x_2 y_{12}\beta - 3\gamma - \lambda & y_{02}\beta(1 - 2x_2) - x_1(y_{01} + 2y_{02} + y_{12})n\beta - \gamma + \lambda \\ 2(2x_1 y_{11} n\beta + x_2 y_{12} n\beta + \gamma) & 2x_1 y_{12} n\beta - 2\lambda\end{bmatrix}$$

$$\mathbf{J}_{12} = \begin{bmatrix}x_1(1 - 2x_1 - x_2)n\beta & x_2(1 - 2x_1 - x_2)n\beta & -x_1^2 n\beta & -x_1 x_2 n\beta \\ 0 & 0 & 2x_1^2\beta & 2x_1 x_2 n\beta\end{bmatrix}$$

$$\mathbf{J}_{22} = \begin{bmatrix}-\rho\left(1 + \frac{1}{2}\tilde{g}(0,1)\right) & 0 & 0 & 0 \\ 0 & -\rho(1 + \tilde{g}(0,2)) & 0 & 0 \\ 0 & 0 & -\rho\left(1 + \frac{1}{4}\tilde{g}(1,1)\right) & 0 \\ 0 & 0 & 0 & -\rho\left(1 + \frac{1}{2}\tilde{g}(1,2)\right)\end{bmatrix}.$$

[55] We have used the fact that for any block matrix, $\det\begin{bmatrix}\mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D}\end{bmatrix} = \det(\mathbf{D})\det(\mathbf{A} - \mathbf{B}\mathbf{D}^{-1}\mathbf{C})$ when $\mathbf{D}$ is invertible.

In the case of $\theta = 0$ we further have that

$$\mathbf{J}_{21} = \begin{bmatrix} \frac{4b^2 e^{\eta\left(\frac{2b\beta\tau}{2b(x_1+x_2)+1}-c\right)}(y_{01}-1)\beta\eta\rho\tau}{(1+e^{c\eta})\left(1+e^{\eta\left(\frac{2b\beta\tau}{2b(x_1+x_2)+1}-c\right)}\right)^2(2b(x_1+x_2)+1)^2} & \frac{4b^2 e^{\eta\left(\frac{2b\beta\tau}{2b(x_1+x_2)+1}-c\right)}(y_{01}-1)\beta\eta\rho\tau}{(1+e^{c\eta})\left(1+e^{\eta\left(\frac{2b\beta\tau}{2b(x_1+x_2)+1}-c\right)}\right)^2(2b(x_1+x_2)+1)^2} \\ \frac{4b^2 e^{\eta\left(\frac{4b\beta\tau}{2b(x_1+x_2)+1}-c\right)}(y_{02}-1)\beta\eta\rho\tau}{(1+e^{c\eta})\left(1+e^{\frac{4b\beta\eta\tau}{2b(x_1+x_2)+1}}\right)^2(2b(x_1+x_2)+1)^2} & \frac{4b^2 e^{\eta\left(\frac{4b\beta\tau}{2b(x_1+x_2)+1}-c\right)}(y_{02}-1)\beta\eta\rho\tau}{(1+e^{c\eta})\left(1+e^{\frac{4b\beta\eta\tau}{2b(x_1+x_2)+1}}\right)^2(2b(x_1+x_2)+1)^2} \\ \frac{8b^2(b+1)\exp\left(2\eta\left(\frac{2b(b+1)\beta\tau}{2b(x_1+x_2)+1}-c\right)\right)(y_{11}-1)\beta\eta\rho\tau}{\left(1+\exp\left(\eta\left(\frac{2b(b+1)\beta\tau}{2b(x_1+x_2)+1}-c\right)\right)\right)^3(2b(x_1+x_2)+1)^2} & \frac{8b^2(b+1)\exp\left(2\eta\left(\frac{2b(b+1)\beta\tau}{2b(x_1+x_2)+1}-c\right)\right)(y_{11}-1)\beta\eta\rho\tau}{\left(1+\exp\left(\eta\left(\frac{2b(b+1)\beta\tau}{2b(x_1+x_2)+1}-c\right)\right)\right)^3(2b(x_1+x_2)+1)^2} \\ \frac{4b^2(b+1)\exp\left(\eta\left(\frac{2b(b+1)\beta\tau}{2b(x_1+x_2)+1}-c\right)\right)(y_{12}-1)\beta\eta\rho\tau}{(1+e^{c\eta})\left(1+\exp\left(\eta\left(\frac{2b(b+1)\beta\tau}{2b(x_1+x_2)+1}-c\right)\right)\right)^2(2b(x_1+x_2)+1)^2} & \frac{4b^2(b+1)\exp\left(\eta\left(\frac{2b(b+1)\beta\tau}{2b(x_1+x_2)+1}-c\right)\right)(y_{12}-1)\beta\eta\rho\tau}{(1+e^{c\eta})\left(1+\exp\left(\eta\left(\frac{2b(b+1)\beta\tau}{2b(x_1+x_2)+1}-c\right)\right)\right)^2(2b(x_1+x_2)+1)^2} \end{bmatrix}$$

while in the case of $\theta = 1$ we have that

$$\mathbf{J}_{21} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

Moreover, evaluated at $x_1 = x_2 = 0$ we obtain

$$\mathbf{J}_{12} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Hence, in both cases when $x_1 = x_2 = 0$ we have that$\det(\mathbf{J} - \mu\mathbf{I}_6) = \det(\mathbf{J}_{11} - \mu\mathbf{I}_2)\det(\mathbf{J}_{22} - \mu\mathbf{I}_4)$. We have that $\det(\mathbf{J}_{22} - \mu\mathbf{I}_4) = \frac{1}{16}(2\mu + \tilde{g}(0,1)\rho)(\mu + \tilde{g}(0,2)\rho)(4\mu + \tilde{g}(1,1)\rho)(2\mu + \tilde{g}(1,2)\rho)$. The roots of the characteristic polynomial give us the eigenvalues $\mu_3 = -\rho\left(1 + \frac{1}{2}\tilde{g}(0,1)\right)$, $\mu_4 = -\rho(1 + \tilde{g}(0,2))$, $\mu_5 = -\rho\left(1 + \frac{1}{4}\tilde{g}(1,1)\right)$, $\mu_6 = -\rho\left(1 + \frac{1}{2}\tilde{g}(1,2)\right)$. These are all negative an real. Further we have that at $x_1 = x_2 = 0$ that $\det(\mathbf{J}_{11} - \mu\mathbf{I}_2) = \alpha(2\gamma+\mu) + (\gamma+\lambda+\mu)(2(\gamma+\lambda)+\mu) - n\beta(y_{01}(2\lambda+\mu) + 2y_{02}(\alpha+\gamma))$. Setting $\alpha = \gamma = 0$ this simplifies to $\det(\mathbf{J}_{11} - \mu\mathbf{I}_2) = (2\lambda + \mu)(\lambda + \mu - n\beta y_{01})$. The roots are given by $\mu_1 = n\beta y_{01} - \lambda$ and $\mu_2 = -2\lambda$. Inserting the stationary state $y_{01} = \frac{\tilde{g}_{01}}{2+\tilde{g}_{01}}$ delivers $\beta^c = \frac{\lambda(e^{c\eta}+2)}{n} + \frac{1}{b\eta(2-\theta)\tau} W\left(\frac{b\eta(2-\theta)\lambda\tau(e^{c\eta}+1)}{n} e^{\eta\left(c - \frac{b(2-\theta)\lambda\tau(e^{c\eta}+2)}{n}\right)}\right)$, for $\theta \in \{0,1\}$, where $W(x)$ is the Lambert W function (or product-log), which is implicitly defined by $W(x)e^{W(x)} = x$. This expression for β^c is identical to Equation (51) for the case of $N = 1$. This completes the proof of part (i) of the proposition.

Next, we consider part (ii) of the proposition. In the case of $\eta \to 0$ and $\vartheta \to 0$ we have that $f = \frac{1}{2}$ and $y_{01} = y_{02} = y_{11} = y_{12} = \frac{1}{5}$. Setting also $\gamma = 0$ we obtain for the fixed points $0 = 2\lambda x_1 - \frac{2}{5}\beta n x_0(x_1+x_2) - \xi x_0 - \frac{1}{2}\chi(x_0 + 2x_1 + x_2 - 1)$, $0 = \lambda x_2 + \frac{1}{5}\beta n(x_0 - x_1)(x_1+x_2) - x_1(\alpha+\lambda) - \xi x_1$ and $0 = \frac{2}{5}\beta n x_1^2 + \frac{2}{5}x_1(5\alpha + \beta n x_2) - x_2(2\lambda + \xi)$. Solving for x_0, x_1 and x_2 gives either $x_0 = \frac{\chi}{2\xi+\chi}$, $x_1 = x_2 = 0$ or

$$x_0 = \frac{(2\lambda+\xi)(-2\beta n\chi + 5\xi(2\xi+\chi) + S) - 10\alpha(2\lambda-\xi)(2\xi+\chi)}{8\alpha\beta n(2\xi+\chi)},$$

$$x_1 = \frac{20\alpha^2(2\xi+\chi) + \alpha(20(3\lambda+\xi)(2\xi+\chi) + 4\beta n\chi - 2S) - (2\lambda+\xi)(-2\beta n\chi + 5\xi(2\xi+\chi) + S)}{8\alpha\beta n(2\xi+\chi)},$$

$$x_2 = \frac{1}{80\alpha\beta n(2\lambda+\xi)(2\xi+\chi)^2}(5(2\xi+\chi)(2\alpha - 4\lambda - 3\xi) + 2\beta n\chi + S)T,$$

where we have denoted by $S \equiv (100\alpha^2(2\xi+\chi)^2 + 20\alpha(2\xi+\chi)(5(4\lambda+\xi)(2\xi+\chi) + 2\beta n\chi) + (2\beta n\chi - 5\xi(2\xi+\chi))^2)^{1/2}$ and $T \equiv 20\alpha^2(2\xi+\chi) + \alpha(20(3\lambda+\xi)(2\xi+\chi) + 4\beta n\chi - 2S) - (2\lambda+\xi)(-2\beta n\chi + 5\xi(2\xi+\chi) + S)$ Next, assume also that $\xi = \chi = 0$. Then $x_0 = 1 - 2x_1 - x_2$ and x_1 and x_2 are the solutions to

$0 = \frac{2}{5}\beta x_1(x_1+x_2)+2\alpha x_1-2\lambda x_2$, $0 = \frac{1}{5}\left(-3\beta x_1^2 + x_1(-5\alpha - 5\lambda + \beta - 4\beta x_2) + x_2(5\lambda + (\beta - \beta x_2))\right)$. The solutions are given by then the stationary state is given by $x_1 = x_2 = 0$, or $x_2 = \frac{\lambda S - 10\alpha^2 - 25\alpha\lambda - \beta\lambda n + 2\alpha S}{2\alpha\beta n}$, and $x_1 = \frac{5\alpha^2+15\alpha\lambda+\alpha\beta n+\beta\lambda n-S(\alpha+\lambda)}{2\alpha\beta n}$, where we have denoted by $S \equiv \sqrt{100\alpha\lambda + (5\alpha+\beta n)^2}$. Since the block element $\mathbf{J}_{21}$ of the Jacobian $\mathbf{J}$ is an all zero matrix when $\eta \to 0$, we have that $\det(\mathbf{J}-\mu\mathbf{I}_7) = \det(\mathbf{J}_{11} - \mu\mathbf{I}_3)\det(\mathbf{J}_{22} - \mu\mathbf{I}_4)$. The roots of the characteristic polynomial $\det(\mathbf{J}_{22} - \mu\mathbf{I}_4) = 0$ are all real and negative. Further, the roots of $\det(\mathbf{J}_{11} - \mu\mathbf{I}_3) = 0$ are given by $\mu_{1,2} = \frac{1}{10}(-5\alpha - 5(3\lambda + 2\xi) \pm S + \beta n(x_0 - 2x_1 - 3x_2))$ and $\mu_3 = -\xi - \frac{\chi}{2}$, where we have denoted by $S \equiv (25\alpha^2 + 10\alpha(15\lambda + \beta n(3x_0 - 2x_1 - x_2)) + (5\lambda + \beta n(x_0 + 2x_1 + x_2))^2)^{1/2}$. The eigenvalue μ_3 is real and negative by definition. The eigenvalues μ_1, μ_2 are real, if S is real, and a sufficient condition for this is that $\lambda \geq \alpha$.

Finally, we give a proof of part (iii) of the proposition. When $\beta b\tau > c$ in the case of $\theta = 1$ and $\frac{2\beta b(b+1)\tau}{2b(x_1+x_2)+1} > c$ in the case of $\theta = 0$ then in the limit of $\eta \to \infty$ starting from an empty graph $\overline{K}_n$ we have that $\tilde{g}_{01} = \tilde{g}_{02} = \tilde{g}_{22} = 0$, $\tilde{g}_{11} = 2$, and consequently $\bar{y}(0,1) = \bar{y}(0,2) = \bar{y}(1,2) = 0$ and $\bar{y}(1,1) = \frac{1}{3}$, while $f = \frac{1}{2}$ when $\kappa < 1$ in the case of $\theta = 1$ and $\kappa < \frac{1}{2b(x_1+x_2)+1}$ in the case of $\theta = 0$. The fixed points of Equation (60) then satisfy $0 = \chi - x_0(2\gamma + \xi + \chi) + 2\lambda x_1 - \chi(2x_1 + x_2)$, $0 = -\frac{1}{3}\beta n x_1^2 + \gamma x_0 - x_1(\alpha+\gamma+\lambda+\xi) + \lambda x_2$ and $0 = \frac{2}{3}\beta n x_1^2 + 2x_1(\alpha+\gamma) - x_2(2\lambda+\xi)$. The solution is given by

$$x_0 = \frac{\xi\chi}{(2\gamma+\xi)(\xi+\chi)} - \frac{3\lambda(\alpha(2\gamma+\xi) + (\gamma+\lambda+\xi)(2(\gamma+\lambda)+\xi))}{\beta n(2\gamma+\xi)^2} + \frac{\sqrt{3}\lambda S}{\beta n(2\gamma+\xi)^2(\xi+\chi)},$$
$$x_1 = \frac{1}{2\beta n(2\gamma+\xi)}\left(\frac{\sqrt{3}S}{\xi+\chi} - 3(\alpha(2\gamma+\xi) + (\gamma+\lambda+\xi)(2(\gamma+\lambda)+\xi))\right), \quad x_2 = \frac{T}{\beta n(2\gamma+\xi)^2(\xi+\chi)},$$

where we have denoted by $S \equiv ((\xi+\chi)(3(\xi+\chi)(\alpha(2\gamma+\xi) + (\gamma+\lambda+\xi)(2(\gamma+\lambda)+\xi))^2 + 4\beta\gamma n\chi(2\gamma+\xi)(2\lambda+\xi)))^{1/2}$ and $T \equiv 3\alpha(2\gamma+\xi)(\xi+\chi)(2\gamma+\lambda+\xi) + 3\xi(\gamma+\lambda+\xi)(2\gamma+\lambda+\xi)(2(\gamma+\lambda)+\xi) + \chi(3(\gamma+\lambda+\xi)(2\gamma+\lambda+\xi)(2(\gamma+\lambda)+\xi) + 2\beta\gamma n(2\gamma+\xi)) - \sqrt{3}S(2\gamma+\lambda+\xi)$. When we also set $\xi = \chi = 0$ (no entry/exit) then we need to solve the following system of equations $0 = \gamma - \frac{1}{3}\beta n x_1^2 - x_1(\alpha+3\gamma+\lambda) - \gamma x_2 + \lambda x_2$ and $0 = 2\left(\frac{1}{3}\beta n x_1^2 + x_1(\alpha+\gamma) - \lambda x_2\right)$, while $x_0 = 1 - 2x_1 - x_2$. The fixed point solution is given by $x_1 = \frac{-3\alpha\gamma - 3(\gamma+\lambda)^2 + S_1}{2\beta\gamma n}$, $x_2 = \frac{3\alpha\gamma(2\gamma+\lambda) + 2\beta\gamma^2 n + (2\gamma+\lambda)\left(3(\gamma+\lambda)^2 - S_2\right)}{2\beta\gamma^2 n}$, where we have denoted by $S_1 \equiv (9\left(\alpha\gamma + (\gamma+\lambda)^2\right)^2 + 12\beta\gamma^2\lambda n)^{1/2}$ and $S_2 \equiv (9\gamma^2(\alpha+\gamma)^2 + 18\gamma\lambda^2(\alpha+3\gamma) + 36\gamma\lambda^3 + 9\lambda^4 + 12\gamma^2\lambda(3\alpha+3\gamma+\beta n))^{1/2}$. For $\gamma = 0$ the unique solution is $x_1 = x_2 = 0$. In the case of $\theta = 1$ we have that the block element $\mathbf{J}_{21}$ of the Jacobian $\mathbf{J}$ in Equation (48) is an all zero matrix. The eigenvalues of the Jacobian are then determined by $\det(\mathbf{J}-\mu\mathbf{I}_7) = \det(\mathbf{J}_{11}-\mu\mathbf{I}_3)\det(\mathbf{J}_{22}-\mu\mathbf{I}_4)$. We know that the characteristic polynomial $\det(\mathbf{J}_{22} - \mu\mathbf{I}_4)$ has only real eigenvalues. Moreover, when $\bar{y}(0,1) = \bar{y}(0,2) = \bar{y}(1,2) = 0$ and $\bar{y}(1,1) = \frac{1}{3}$ we have that $\det(\mathbf{J}_{11} - \mu\mathbf{I}_3) = -\frac{1}{3}(\mu + \xi + \chi)(3\alpha(2\gamma+\mu+\xi) + 3(\gamma+\lambda+\mu+\xi)(2\gamma+2\lambda+\mu+\xi) + 2\beta n x_1(2\gamma+\mu+\xi))$. The roots give us the eigenvalues $\mu_{1,2} = \frac{1}{6}(-3\alpha - 9\gamma - 9\lambda \pm \sqrt{9\lambda^2 + 18\lambda(3\alpha+\gamma+2\beta n x_1) + (3\alpha - 3\gamma + 2\beta n x_1)^2} - 2\beta n x_1 - 6\xi)$ and $\mu_3 = -\xi - \chi$. Since μ_3 is real and negative by definition, and the the term under the square root in the expression for $\mu_{1,2}$ cannot be negative, we find that the eigenvalues are all real. □

H. Total Output, Efficiency and Inequality

In this appendix we study total output from a social planners perspective. We show that total output is increasing with competition, and we find that this effect is stronger the higher is the variance in the stocks of knowledge relative to the average stock of knowledge. We then show that more centralized structures increase total output, and derive conditions under which nested split graphs,[56] a particular class of core-periphery networks, are output maximizing.

[56] See Appendix E for a formal definition and characterization.

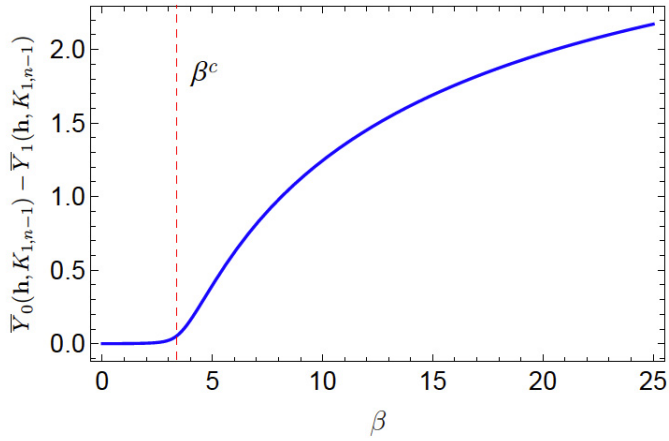


Figure H.1: The net output gain from multi-product competition in the star network $K_{1,n-1}$ with the same parameter values as in Figure I.1 in Appendix I. The threshold $\beta^c = (\lambda + \gamma)/\sqrt{n-1}$ is indicated with a dashed line.

H.1. Output Gains from Competition

Net total output, $\overline{Y}$, is given by total output minus the cost of production (see Equation (27) in Appendix B), R&D collaboration and entry:

$$\overline{Y}_\theta(\mathbf{h}, G) = \Gamma \sum_{i=1}^{n_a} \left(\theta + (1-\theta) \frac{A_i}{\frac{1}{n_a}\sum_{j=1}^{n_a} A_j} \right) A_i - 2mc - n_a\kappa, \tag{62}$$

for $\theta \in \{0, 1\}$, where m is the number of links in the network, A_i is a function of the technology portfolio $\mathbf{h}_i$ of firm i as defined in Equation (2), a constant $\Gamma > 0$, and $n_a = n\sum_{s=1}^{N}\binom{N}{s}\bar{x}(s)$ is the number of active firms. In the case of independent markets ($\theta = 1$) we obtain $\overline{Y}_1(\mathbf{h}, G) = \Gamma\sum_{i=1}^{n_a} A_i - 2mc - n_a\kappa$. In this case net output is increasing with the total productivity $A = \sum_{i=1}^{n_a} A_i$ in the economy and decreasing with the number of links m in the network. In the following we will normalize $\Gamma = 1$. Using the fact that $A_i(\mathbf{h}_i) = a + b|\,\mathsf{S}(\mathbf{h}_i)|$ from Equation (2) and setting also $a = 1$ we obtain for the case of single-product monopolies

$$\overline{Y}_1(\mathbf{h}, G) = n_a(1 + b\bar{h} - c\bar{d} - \kappa), \tag{63}$$

where $\bar{h}$ is the average stock of knowledge and $\bar{d}$ the average degree. In contrast, in the case of multi-product competition ($\theta = 0$) net output can be written as

$$\overline{Y}_0(\mathbf{h}, G) = \frac{\sum_{i=1}^{n_a} A_i^2}{\frac{1}{n_a}\sum_{i=1}^{n_a} A_i} - 2mc - n_a\kappa.$$

Gross total output is thus proportional to the ratio of the second to the first sample moment of the productivity distribution, and net output is decreasing with the number of links m in the network and the number of active firms. Similarly, for $a = 1$ we get

$$\overline{Y}_0(\mathbf{h}, G) = n_a \left(1 + b\bar{h} + \frac{b^2\sigma_h^2}{1 + b\bar{h}} - c\bar{d} - \kappa \right), \tag{64}$$

where σ_h^2 is the variance in the stocks of knowledge. Note that when the network is exogenously given, then the distribution of the stocks of knowledge is the same irrespective of whether we consider the case with or without competition. This is because competition affects only the formation of collaborations (or entry), but not the diffusion of technologies. With net output in the case of single-product monopolies from Equation (63), multi-product competition in Equation (64) and assuming that there is no entry and exit ($\chi = \xi = 0$), such that $n_a = n$, we then can write the output gains from multi-product competition simply as

$$\overline{Y}_0(\mathbf{h}, G) - \overline{Y}_1(\mathbf{h}, G) = \frac{nb^2\sigma_h^2}{1 + b\bar{h}}. \tag{65}$$

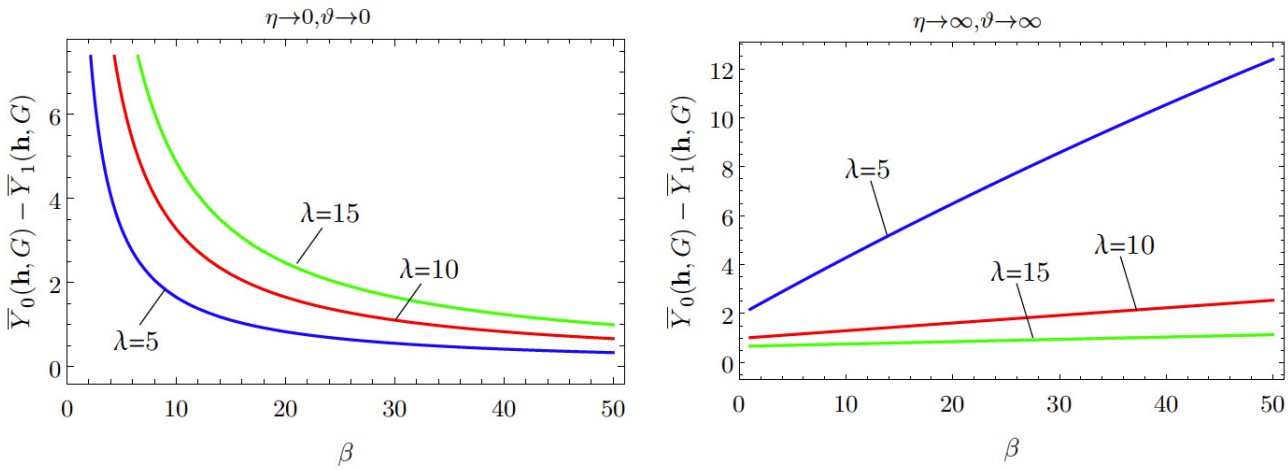


Figure H.2: (Left panel) The output gain from competition across multiple products for different values of the knowledge obsolescence rate λ with $\alpha = 0.01$, $\gamma = 0$, $n = 100$ and $b = 1$ when $\eta, \vartheta \to 0$ (high uncertainty). (Right panel) The output gain from competition for different values of the knowledge obsolescence rate λ with $\alpha = 0$, $n = 100$ and $b = 1$ when $\eta, \vartheta \to \infty$ (low uncertainty).

It follows that, when the network and the technology portfolios are exogenously given, output in the multi-product competition case ($\theta = 0$) is higher than in the case of independent markets, and this effect is stronger the higher is the variance in the stocks of knowledge relative to the average stock of knowledge (cf. the coefficient of variation $CV_h = \sigma_h/\bar{h}$). These results also hold in the case of endogenous networks and technology portfolios.[57] The output gains are due to the fact that multi-product competition leads to reallocation and the replacement of less productive firms with more productive ones.

By considering the two polar opposite cases of single-product monopolies ($\theta = 1$) and multi-product competition ($\theta = 0$) we can investigate whether competition has a conducive or detrimental effect on innovation and output [cf. e.g. Aghion et al., 2005]. As in Section H.2 we assume that there is no entry and exit in the following.

We first consider specific graphs where we can compute (asymptotic) net output explicitly.[58] The output gains from introducing product competition for a given network have been derived in Equation (65). For the complete graph, K_n, the variance in the knowledge stocks is zero, so that net output in both cases, $\theta = 1$ and $\theta = 0$ is the same, and there are no output gains from competition. However, output gains can be obtained from multi-product competition in the case of the star $K_{1,n-1}$. These gains are shown in Figure H.1. We see that the gains are increasing with increasing values of β, but only after β exceeds the threshold β^c. Moreover, one can show that net output for the star $K_{1,n-1}$ is always higher than in the complete graph K_n,[59] indicating that centralization has a conducive effect on output. This resembles previous welfare studies for R&D networks which did not consider any technological dynamics [cf. Westbrock, 2010].

Next, we consider endogenously formed networks in the case of $N = 2$ without R&D uncertainty ($\eta, \vartheta \to \infty$) as well as with strong uncertainty ($\eta, \vartheta \to 0$).[60] We know from Propositions 4 that in these cases the stationary average knowledge stocks, $\bar{h}$, and average degree, $\bar{d}$, are identical for the single-product monopolies case ($\theta = 1$) and the multi-product competition case ($\theta = 0$). It then

[57] To illustrate this, in Appendix **??** we also derive the output gains from multi-product competition for specific graphs, as well as for endogenously formed networks and technology portfolios within specific parameter ranges, and compare the effect on output of different levels of spillovers from R&D collaborations.

[58] This is done for the star network, $K_{1,n-1}$ and the complete network, K_n in Appendix I. In particular, Figure I.1 in Appendix I shows net output for the complete graph K_n and the star for $\theta = 1$ and $\theta = 0$ as a function of β for different values of the linking cost $c \in \{0.01, 0.02, 0.03\}$. For both cases, $\theta = 1$ and $\theta = 0$, net output is increasing with decreasing cost c.

[59] See Appendix I.

[60] Similar results hold for the case of $N = 1$.

follows that the output gains from multi-product competition are given by Equation (65).[61] Figure H.2 shows the output gains from multi-product competition for different values of the knowledge obsolescence rate λ for both cases without R&D uncertainty and with strong uncertainty as a function of the spillover parameter β. The two cases can show starkly different behavior, as we find, for example, that the output gains are decreasing with strong uncertainty, but increasing with vanishing uncertainty. The explanation for this result is that with strong uncertainty links are formed mostly at randomly, while with vanishing uncertainty, only the most profitable links are formed. In the latter case, firms with larger technology portfolios attract more links, the network becomes more centralized (with improved technology diffusion properties), and this increases net output.

The next proposition computes the output gains from competition across sectors in the case of vanishing ($\eta \to \infty$) and large noise ($\eta \to 0$) when there are two technology categories ($N = 2$), assuming that there is no entry and exit ($\xi = \chi = 0$).

Proposition 5. *In the case of $N = 2$ and $\rho \to \infty$ the value function $\lim_{r\to 0} V_{1,r}$ in the competitive case with $\theta = 0$ is higher than in the non-competitive case with $\theta = 1$ for both strong shocks when $\eta \to 0$ and vanishing shocks when $\eta \to \infty$. In particular, in the limit of large ρ the following holds:*

(i) In the case of $\eta \to 0$ and $\gamma = 0$ the output gain from competition is given by

$$\overline{Y}_0(\mathbf{h}, G) - \overline{Y}_1(\mathbf{h}, G) = \frac{b^2\left(25\alpha^3 + \beta^2\lambda n^2 + 5\alpha^2(20\lambda + \beta n) + 5\alpha\lambda(20\lambda + 3\beta n) - S\left(10\alpha^2 + 20\alpha\lambda + \alpha\beta n + \beta\lambda n\right) + \alpha S^2\right)}{\alpha\beta(b(5\alpha + 10\lambda - \beta n - S) + \beta n)}, \tag{66}$$

where we have denoted by $S \equiv \sqrt{100\alpha\lambda + (5\alpha + \beta n)^2}$.

(ii) In the case of $\eta \to \infty$ the output gain from competition is given by

$$\begin{aligned}\overline{Y}_0(\mathbf{h}, G) - \overline{Y}_1(\mathbf{h}, G) = {} & \frac{1}{\beta\gamma^2\left(-3b(\gamma+\lambda)\left(\alpha\gamma + (\gamma+\lambda)^2\right) - (2b+1)\beta\gamma^2 n + b\gamma S + b\lambda S\right)} \\ & \times \Big(b^2\left(18\alpha^2\gamma^2(\gamma+\lambda)^2 + 3\alpha\gamma\left(12(\gamma+\lambda)^4 + \beta\gamma^2 n(\gamma+2\lambda) - 2S(\gamma+\lambda)^2\right)\right. \\ & \left. + \beta\gamma^2 n\left(3(\gamma+\lambda)^2(\gamma+6\lambda) - S(\gamma+2\lambda)\right) - 6(\gamma+\lambda)^4\left(S - 3(\gamma+\lambda)^2\right)\right)\Big),\end{aligned} \tag{67}$$

where we have denoted by $S \equiv \sqrt{9\left(\alpha\gamma + (\gamma+\lambda)^2\right)^2 + 12\beta\gamma^2\lambda n}$.

Proof of Proposition 5. In order to determine net output, we need to compute the following quantities from the endogenous variables $\bar{x}(s)$ and $\bar{\xi}(s, s')$, $s, s' \in \{0, \ldots, N\}$: $\bar{h} = \sum_{s=0}^{N} s\binom{N}{s}\bar{x}(s)$ and $\sigma_h^2 = \sum_{s=0}^{N}(s - \bar{h})^2\binom{N}{s}\bar{x}(s)$. In the case of $N = 2$ we have that $\bar{h} = 2(\bar{x}(1) + \bar{x}(2))$, and $\sigma_h^2 = \bar{h}^2(1 - 2\bar{x}(1) - \bar{x}(2)) + 2(1 - \bar{h})^2\bar{x}(1) + (2 - \bar{h})^2\bar{x}(2)$. From Proposition 4 we know that in the limit of $\eta \to \infty$ both cases $\theta = 1$ and $\theta = 0$ are identical, and in particular the average degree, $\bar{d}$, as well as the average knowledge stock, $\bar{h}$, are the same, so that the output gain can be computed from $\overline{Y}_0(\mathbf{h}, G) - \overline{Y}_1(\mathbf{h}, G) = \frac{nb^2\sigma_h^2}{1+b\bar{h}}$, where we insert $\bar{x}(1)$ and $\bar{x}(2)$ as stated in Proposition 4 into the above expressions for $\bar{h}$ and σ_h^2 delivers Equation (67). The same observation can be made for the case of strong shocks as stated in Proposition 4, where $\bar{x}(1)$ and $\bar{x}(2)$ can be obtained for the case of $\gamma = 0$. Inserting into the output gain gives Equation (66). □

H.2. Efficient Networks

The planner's objective is to maximize the stationary net output, $\lim_{t\to\infty} \overline{Y}_\theta(\mathbf{h}_t, G)$, by choosing a network $G \in \mathcal{G}(n)$ and knowing that the technology dynamics $\mathbf{h}_t$ depend on the network structure

[61] An explicit expression for these net output gains can be found in Proposition 5 in Appendix **??**.

G of collaborating firms (as discussed in Section 3.2). Hence, we can write the planner's problem as follows

$$\max_{G\in\mathcal{G}(n)} V_\theta(G) = \max_{G\in\mathcal{G}(n)} \lim_{t\to\infty} \mathbb{E}(\overline{Y}_\theta(\mathbf{h}_t, G)|G), \tag{68}$$

subject to Equation (10). We denote by G^* the efficient network solving the above optimization problem. To simplify our analysis, in the following, we assume that there is no entry and exit ($\chi = \xi = 0$), such that $n_a = n$.[62] This planner is concerned with the long-run outcome in the economy. The question then is whether the long-run outcome of the decentralized economy coincides with the one that the planner would prefer. With the above characterization of the planner's value function we then can derive the following proposition.

Proposition 6. *Assume that $\alpha = 0$, $\xi = \chi = 0$ (no entry or exit), and $h_{ik,0} = 0$ for all $i = 1, \ldots, n$ and $k = 1, \ldots, N$. Define the Katz-Bonacich centrality vector as $\mathbf{b}(G, \varphi) \equiv (\mathbf{I}_n - \varphi\mathbf{A})^{-1}\mathbf{u}$ for $0 \le \varphi \le 1/\mu_{\max}(G)$, with $\mathbf{u}$ being a vector of ones and $\mu_{\max}(G)$ being the largest real eigenvalue of the adjacency matrix of G. Moreover, let $\mathcal{G}(n, m)$ denote the class of graphs with n nodes and m links. Then, under the assumption of pairwise independence, i.e. $\lim_{t\to\infty} \mathrm{Cov}(h_{ik,t}, h_{jk,t}|G) = 0$ for all $i \neq j$, we have the following:*

(i) The asymptotic knowledge stocks, $|\mathsf{S}(\mathbf{h}_i)|$, follow a binomial distribution, $\mathrm{Binom}(N, p_i)$, with success probability $p_i = \mathbb{E}(h_{ik})$, mean $\mu_i = Np_i$ and variance $\sigma_i^2 = Np_i(1 - p_i)$, where

$$p_i \equiv \lim_{t\to\infty} \mathbb{E}\left(h_{ik,t}|\, G\right) = \begin{cases} 1, & \text{if } \beta \gg \lambda, \\ \frac{\gamma}{\lambda+\gamma} b_i\left(G, \frac{\beta}{\lambda+\gamma}\right), & \text{if } \beta \ll \lambda. \end{cases} \tag{69}$$

(ii) In the case of $\theta = 1$ (single-product monopolies) the value function can be written as

$$V_1(G) = n\left(1 + \frac{Nb}{n}\sum_{i=1}^{n} p_i\right) - 2mc. \tag{70}$$

Moreover, the graph G^* maximizing the value function $\lim_{r\to 0} V_{1,r}$ in the case of β being much smaller than λ is a nested split graph.

(iii) In the case of $\theta = 0$ (multi-product competition) the value function can be written as

$$V_0(G) = n\left(\frac{n + Nb(2+b)\sum_{i=1}^{n} p_i + N(N-1)b^2\sum_{i=1}^{n} p_i^2}{n + Nb\sum_{i=1}^{n} p_i}\right) + O(1) - 2mc. \tag{71}$$

The importance of the Katz-Bonacich centrality in Equation (69) in relation to equilibria and aggregate outcomes in network games has been prominently studied in Ballester et al. [2006]. Moreover, nested split graphs,[63] which appear in part (ii) of Proposition 6, have a core-periphery structure with a densely connected core of firms and a periphery of sparsely connected firms attached to the core [cf. Hojman and Szeidl, 2008]. These highly centralized network structures have good technology diffusion properties while minimizing on the number of links [cf. König et al., 2011], and thus they are optimal from an efficiency perspective.

The characterization of the efficient network in the multi-product competitive case ($\theta = 0$) is more involved due to the nonlinearity of Equation (71), and remains an unresolved problem. However, net output in the competitive case can be computed for particular classes of graphs, and we analyze two examples, the star and the regular graph, in Appendix I. We find that for both cases, with and without competition, the star generates higher net output than the complete graph. This indicates that, in the absence of entry or exit, more centralized network structures generate

[62] In Section 7 we study net output in the case with entry and exit.
[63] See Appendix E for a formal definition and characterization.

higher output [cf. König et al., 2011; Westbrock, 2010]. The importance of centralization will be discussed further in the policy Section 7 below, where we also consider the effect of entry and exit.

Before we can proceed with the proof of Proposition 6 we need to state and prove two lemmas related to the asymptotic knowledges stocks that appear in the value function of Equation (68):

$$\lim_{t\to\infty} \mathbb{E}(\overline{Y}_\theta(\mathbf{h}_t, G)|G) = n \times \begin{cases} \left(1 + b\lim_{t\to\infty} \mathbb{E}\left(\bar{h}_t \middle| G\right) - c\bar{d}\right), & \text{if } \theta = 1, \\ \lim_{t\to\infty} \mathbb{E}\left(1 + b\bar{h}_t + \frac{b^2\sigma^2_{h,t}}{1+b\bar{h}_t} - c\bar{d}\middle| G\right), & \text{if } \theta = 0. \end{cases}$$

In particular, we next analyze the dynamics of the stock of knowledge of a firm when $\alpha = 0$, assuming that pair-correlations of the form $\mathrm{Cov}(h_{ik,t}, h_{jk,t})$ can be neglected.[64] The following lemma describes the evolution of the stocks of knowledge at early times starting from the initial condition $h_{ik,0} = 0$ for all $i = 1, \ldots, n$ and $k = 1, \ldots, N$.

Lemma 2. *Consider a given network G and let $\alpha = 0$, $h_{ik,0} = 0$ for all $i = 1, \ldots, n$ and $k = 1, \ldots, N$, and assume that pair-correlations can be neglected, i.e. $\mathrm{Cov}(h_{ik,t}, h_{jk,t}|G) = 0$, then in the limit of small t the expected stock of knowledge is given by $\mathbb{E}\left(|\,\mathsf{S}(\mathbf{h}_{i,t})|\,|\,G\right) = N\sum_{j=1}^n \frac{\gamma\langle \mathbf{u}, \mathbf{v}_j\rangle^2}{\gamma+\lambda-\beta\mu_j}\left(1 - e^{-(\gamma+\lambda-\beta\mu_j)t}\right)$, where $\mathbf{v}_j$ is the eigenvector associated with the j-th eigenvalue μ_j of $\mathbf{A}$, i.e. $\mathbf{A}\mathbf{v}_j = \mu_j\mathbf{v}_j$ and $\langle \mathbf{v}_i, \mathbf{v}_j\rangle = \delta_{ij}$ for all $k = 1, \ldots, N$ and $i, j = 1, \ldots, n$.*

Proof of Lemma 2. When $\alpha = 0$, the dynamics of the individual knowledge categories $k = 1, \ldots, N$ become independent. Then consider the time dependent random variable $X_i(t) = \mathbb{1}_{\{h_{ik,t}=1\}}$, where we have dropped the index k. Given the current state $X_i(t)$, if $h_{ik,t} = 1$ then $X_i(t)$ can change from 1 to 0 at a rate λ. If $h_{ik,t} = 0$ then $X_i(t)$ can change from 0 to 1 at a rate $\gamma + \beta\sum_{j=1}^n a_{ij}X_j(t)$. The expected change in $h_{ik,t}$ for a sufficiently small time interval $[t, t+\Delta t)$, conditional on the current state $X_i(t)$ and G, is then given by

$$\mathbb{E}(X_i(t+\Delta t)|X_i(t), G) - X_i(t) = \left(\gamma + \beta\sum_{j=1}^n a_{ij}X_j(t)\right)(1 - X_i(t))\Delta t - \lambda X_i(t)\Delta t + o(\Delta t).$$

Taking the expectation on both sides, dividing by Δt and denoting by $y_i(t) \equiv \mathbb{E}(X_i(t)|G)$, where $y_i(t+\Delta t) = \mathbb{E}(\mathbb{E}(X_i(t+\Delta t)|X_i(t), G)|G)$ by the law of iterated expectation, we obtain

$$\frac{y_i(t+\Delta t) - y_i(t)}{\Delta t} = \gamma - (\lambda+\gamma)y_i(t) + \beta\sum_{j=1}^n a_{ij}y_j(t) - \beta\sum_{j=1}^n a_{ij}\mathbb{E}(X_i(t)X_j(t)|G) + o(1).$$

Observe that the last term can be written as follows $\mathbb{E}(X_i(t)X_j(t)|G) = \mathbb{E}(\mathbb{1}_{\{h_{ik,t}=1\}}\mathbb{1}_{\{h_{jk,t}=1\}}|G)$ $= \mathbb{P}(h_{ik,t} = 1, h_{jk,t} = 1|G) = \mathbb{P}(h_{jk,t} = 1|h_{ik,t} = 1, G)\mathbb{P}(h_{ik,t} = 1|G) = \mathbb{P}(h_{jk,t} = 1|h_{ik,t} = 1, G)\mathbb{E}(\mathbb{1}_{\{h_{ik,t}=1\}}|G) = \mathbb{P}(h_{jk,t} = 1|h_{ik,t} = 1, G)y_i(t)$. Hence, in the limit of $\Delta t \downarrow 0$ we obtain

$$\frac{dy_i(t)}{dt} = \gamma - (\lambda+\gamma)y_i(t) + \beta\sum_{j=1}^n a_{ij}y_j(t) - \beta\sum_{j=1}^n a_{ij}\mathbb{P}(h_{jk,t} = 1|h_{ik,t} = 1, G)y_i(t).$$

In the following we make the pairwise independence assumption $\mathbb{P}(h_{jk,t} = 1, h_{ik,t} = 1|G) = \mathbb{P}(h_{ik,t} = 1|G)\mathbb{P}(h_{jk,t} = 1|G)$, so that $\mathbb{P}(h_{jk,t} = 1|h_{ik,t} = 1, G) = \mathbb{P}(h_{jk,t} = 1|G)$, and we obtain the following system of ODEs

$$\frac{dy_i(t)}{dt} = \gamma - (\lambda+\gamma)y_i(t) + \beta\sum_{j=1}^n a_{ij}y_j(t) - \beta\sum_{j=1}^n a_{ij}y_i(t)y_j(t). \tag{72}$$

When we start from the initial condition $h_{ik,0} = 0$ for all $i = 1, \ldots, n$ and $k = 1, \ldots, N$ at early

[64] We assume that $\mathbb{E}(X_i(t)X_j(t)|G) = \mathbb{E}(X_i(t)|G)\mathbb{E}(X_j(t)|G)$.

times t quadratic terms in $O(y_i(t)y_j(t))$ are negligible, and we get from Equation (72) that

$$\frac{dy_i(t)}{dt} = \gamma - (\lambda+\gamma)y_i(t) + \beta\sum_{j=1}^{n} a_{ij}y_j(t). \tag{73}$$

In vector-matrix form this is $\frac{d\mathbf{y}(t)}{dt} = \gamma\mathbf{u} - (\lambda+\gamma)\mathbf{y}(t) + \beta\mathbf{A}\mathbf{y}(t)$. We can write $\mathbf{y}(t)$ as a linear combination of the eigenvectors $\{\mathbf{v}_k\}_{k=1}^n$ associated with the eigenvalues $\{\mu_k\}_{k=1}^n$ of $\mathbf{A}$, that is $\mathbf{y}(t) = \sum_{k=1}^n c_k(t)\mathbf{v}_k$. Inserting into Equation (73) yields $\sum_{k=1}^n \frac{dc_k(t)}{dt}\mathbf{v}_k = \sum_{k=1}^n(\gamma\langle\mathbf{u},\mathbf{v}_k\rangle - (\lambda+\gamma)c_k(t) + \beta\mu_k c_k(t))\mathbf{v}_k$. Using the othonormality condition of the eigenvectors $\langle\mathbf{v}_j,\mathbf{v}_k\rangle = \delta_{jk}$ we get $\frac{dc_k(t)}{dt} = \gamma\langle\mathbf{u},\mathbf{v}_k\rangle - (\lambda+\gamma+\beta\mu_k)c_k$. The solution of this ODE is given by

$$c_k(t) = \frac{1}{\lambda+\gamma-\beta\mu_k}\left[(\gamma+\lambda-\beta\mu_k)c_k(0) + \gamma\langle\mathbf{u},\mathbf{v}_k\rangle\left(e^{(\gamma+\lambda-\beta\mu_k)t}-1\right)\right]e^{-(\gamma+\lambda-\beta\mu_k)t}.$$

From the initial condition it follows that $c_k(0) = \langle(0,\dots,0)^\top,\mathbf{v}_k\rangle = 0$, so that we get $c_k(t) = \frac{\gamma\langle\mathbf{u},\mathbf{v}_k\rangle}{\gamma+\lambda-\beta\mu_k}\left(1-e^{-(\gamma+\lambda-\beta\mu_k)t}\right)$. Consequently, it follows that $\mathbf{y}(t) = \sum_{k=1}^n \frac{\gamma\langle\mathbf{u},\mathbf{v}_k\rangle}{\gamma+\lambda-\beta\mu_k}\left(1-e^{-(\gamma+\lambda-\beta\mu_k)t}\right)\mathbf{v}_k$. Because for $\alpha = 0$ the different knowledge categories are independent, we then have that $\mathbb{E}(|\,\mathsf{S}(\mathbf{h}_{it})||G) = N\mathbb{E}(h_{ik,t}|G) = Ny_i(t)$. □

From the Lemma 2 we see that the assumption of weak correlations implies a solution for the stock of knowledge of the firm that is bounded only if $\beta\mu_1 < \gamma+\lambda$. This means that there exists a critical value given by the inverse of the largest eigenvalue $1/\mu_1$ such that if $\beta/(\lambda+\gamma) > 1/\mu_1$ then there is a rapid diffusion of knowledge such that all firms quickly attain $h_{ik,t} = 1$ for all $i = 1,\dots,n$ and $k = 1,\dots,N$. In contrast, if $\beta/(\lambda+\gamma) < 1/\mu_1$, the average size of the knowledge portfolios will be much smaller and given by $N\sum_{j=1}^n \gamma\langle\mathbf{u},\mathbf{v}_j\rangle^2/(\gamma+\lambda-\beta\mu_j)$ which is determined by the spectral decomposition of $\mathbf{A}$.

A discussion for the derivation without the assumption that $\mathrm{Cov}(h_{ik,t},h_{jk,t}|G) = 0$ using a moment closure method at the level of two variables averages can be found in Newman [2010, Chap. 17.10.1], and its interpretation of a *mean field approximation* is given in Van Mieghem et al. [2009].

The next lemma derives the asymptotic knowledge stocks that appear in the value function of Equation (68) in the case of $\alpha = 0$ and a given network $G \in \mathcal{G}(n)$.

Lemma 3. *Consider a given network G, let $\alpha = 0$, assume that asymptotically pair-correlations can be neglected, i.e. $\lim_{t\to\infty}\mathrm{Cov}(h_{ik,t},h_{jk,t}|G) = 0$ for all $i = 1,\dots,n$ and $k = 1,\dots,N$. Then*

(i) the stationary stocks of knowledge can be computed from the continued fraction expansion

$$\lim_{t\to\infty}\mathbb{E}\left(\sum_{i=1}^n|\,\mathsf{S}(\mathbf{h}_{i,t})|\,\middle|\,G\right) = N\left(1-\cfrac{\frac{\lambda}{\lambda+\gamma}}{1+\frac{\beta}{\lambda+\gamma}d_i-\cfrac{\frac{\beta\lambda}{(\lambda+\gamma)^2}\sum_{j=1}^n\cfrac{a_{ij}}{1+\frac{\beta}{\lambda+\gamma}d_j-\ddots}}}\right), \tag{74}$$

(ii) we have the bounds

$$0 \le \lim_{t\to\infty}\mathbb{E}\left(\sum_{i=1}^n|\,\mathsf{S}(\mathbf{h}_{i,t})|\,\middle|\,G\right) \le N\left(1-\frac{\frac{\lambda}{\lambda+\gamma}}{1+\frac{\beta}{\lambda+\gamma}d_i}\right),$$

(iii) and asymptotically, we have that

$$\lim_{t\to\infty}\mathbb{E}\left(\sum_{i=1}^n|\,\mathsf{S}(\mathbf{h}_{i,t})|\,\middle|\,G\right) = \begin{cases} Nn, & \text{if } \beta \gg \lambda,\\ \frac{\gamma N}{\lambda+\gamma}\left\langle\mathbf{u},\mathbf{b}\left(G,\frac{\beta}{\lambda+\gamma}\right)\right\rangle & \text{if } \beta \ll \lambda,\\ \frac{Nn\gamma}{\lambda+\gamma} & \text{if } \beta \to 0,\end{cases}$$

where $\mathbf{b}(G,\cdot)$ *is the Katz-Bonacich centrality vector defined as* $\mathbf{b}(G,\varphi) \equiv (\mathbf{I}_n - \varphi\mathbf{A})^{-1}\mathbf{u}$ *for* $0 \le \varphi \le 1/\mu_{\max}(G)$ *and* $\mathbf{u}$ *being a vector of ones.*

Proof of Lemma 3. In the stationary state $\frac{dy_i(t)}{dt} = 0$ we obtain from Equation (72) that $0 = \gamma - (\lambda+\gamma)y_i + \beta\sum_{j=1}^n a_{ij}y_j - \beta\sum_{j=1}^n a_{ij}y_iy_j$. This can be written as $y_i = \frac{\gamma+\beta\sum_{j=1}^n a_{ij}y_j}{\lambda+\gamma+\beta\sum_{j=1}^n a_{ij}y_j}$. For $\beta \gg \lambda+\gamma$ we immediately see that $y_i = 1$. In contrast, for $\beta \ll \lambda+\gamma$ we get $y_i = \frac{\gamma}{\lambda+\gamma} + \frac{\beta}{\lambda+\gamma}\sum_{j=1}^n a_{ij}y_j$, which can be written as $\mathbf{y} = \frac{\gamma}{\lambda+\gamma}\mathbf{u} + \frac{\beta}{\lambda+\gamma}\mathbf{A}\mathbf{y}$. If $\frac{\beta}{\lambda+\gamma} < \frac{1}{\mu_{\max}(G)}$, where $\mu_{\max}(G)$ is the largest eigenvalue of $\mathbf{A}$, respectively, G, the matrix $\mathbf{I}_n - \frac{\beta}{\lambda+\gamma}\mathbf{A}$ is invertible, and we obtain $\mathbf{y} = \frac{\gamma}{\lambda+\gamma}\left(\mathbf{I}_n - \frac{\beta}{\lambda+\gamma}\mathbf{A}\right)^{-1}\mathbf{u} = \frac{\gamma}{\lambda+\gamma}\mathbf{b}\left(G, \frac{\beta}{\lambda+\gamma}\right)$. We have introduced the Bonacich centrality vector defined by $\mathbf{b}(G,\phi) = \sum_{k=1}^n \phi^k\mathbf{A}^k\mathbf{u}$ [cf. Bonacich, 1987], for $\phi < 1/\mu_{\max}(G)$. The Boncaich centrality can also be written as $b_i(G,\phi) = 1 + \phi d_i + \phi^2\sum_{j=1}^n a_{ij}d_j + O(\phi^3)$. For $\beta \to 0$ we then find that $y_i = \frac{\gamma}{\lambda+\gamma}$, which corresponds to the steady state values of a pure birth death process with birth rate γ and death rate λ.

We further find that the steady state values for y_i can be written as a continued fraction expansion

$$y_i = 1 - \frac{\frac{\lambda}{\lambda+\gamma}}{1 + \frac{\beta}{\lambda+\gamma}\sum_{j=1}^n a_{ij}y_j} = 1 - \frac{\frac{\lambda}{\lambda+\gamma}}{1 + \frac{\beta}{\lambda+\gamma}d_i - \frac{\beta\lambda}{(\lambda+\gamma)^2}\sum_{j=1}^n \frac{a_{ij}}{1+\frac{\beta}{\lambda+\gamma}d_j - \ddots}}. \tag{75}$$

Equation (75) gives us an upper bound on y_i given by $0 \le y_i \le 1 - \frac{\lambda}{\lambda+\gamma}/\left(1 + \frac{\beta}{\lambda+\gamma}d_i\right)$. This concludes the proof. □

For example, using Equation (75), a third-order continued fraction approximation to the stationary knowledge stocks is given by

$$\lim_{t\to\infty}\mathbb{E}\left(\sum_{i=1}^n |\mathsf{S}(\mathbf{h}_{i,t})| \middle| G\right) \approx N\left(1 - \frac{\frac{\lambda}{\gamma+\lambda}}{1 + \frac{\beta}{\gamma+\lambda}\sum_{j=1}^n a_{ij}\left(1 - \frac{\frac{\lambda}{\gamma+\lambda}}{1+\frac{\beta}{\gamma+\lambda}\sum_{k=1}^n a_{jk}}\left(1 - \frac{\frac{\lambda}{\gamma+\lambda}}{1+\frac{\beta}{\gamma+\lambda}d_k}\right)\right)}\right). \tag{76}$$

With the results of Lemma 3 we are now able to give a proof of Proposition 6.

Proof of Proposition 6. Part (i) of the proposition is a direct consequence of Lemma 3 when $\alpha = 0$ and the individual technology categories $h_{ik,t}$, $h_{il,t}$, $k \ne l$, become independent.

The proof of the first statement in part (ii) of the proposition follows from noting that $\lim_{t\to\infty}\mathbb{E}\left(\bar{h}_t \middle| G\right) = \frac{1}{n}\sum_{i=1}^n \lim_{t\to\infty}\mathbb{E}\left(|\mathsf{S}(\mathbf{h}_i)|_t \middle| G\right) = Np_i$, and inserting this expression into $\lim_{t\to\infty}\mathbb{E}(\overline{Y}_1(\mathbf{h}_t,G)|G) = n(1 + b\lim_{t\to\infty}\mathbb{E}\left(\bar{h}_t \middle| G\right) - c\bar{d})$ delivers Equation (70). In order to give a proof of the second statement in part (ii) of the proposition, note that the value function is increasing in the sum of the Bonacich centralities, $\sum_{i=1}^n b_i$, and decreasing in the number of links, m, times the linking cost, c. This optimization problem is equivalent to the one analyzed in Belhaj et al. [2016],[65] where it is shown that the optimal network must be a nested split graph.

Next, we give a proof of part (iii) of the proposition. In the case of competition ($\theta = 0$), we can write net output as follows:

$$\begin{aligned}\overline{Y}_0(\mathbf{h},G) &= \frac{n + 2b\sum_{i=1}^n |\mathsf{S}(\mathbf{h}_i)| + b^2\sum_{i=1}^n |\mathsf{S}(\mathbf{h}_i)|^2}{1 + \frac{1}{n}\sum_{i=1}^{n_a} |\mathsf{S}(\mathbf{h}_i)|} - 2mc \\ &= n\left(\frac{1 + 2b\frac{1}{n}\sum_{i=1}^n |\mathsf{S}(\mathbf{h}_i)| + b^2\frac{1}{n}\sum_{i=1}^n |\mathsf{S}(\mathbf{h}_i)|^2}{1 + b\frac{1}{n}\sum_{i=1}^n |\mathsf{S}(\mathbf{h}_i)|} - c\bar{d}\right).\end{aligned}$$

[65] See in particular Theorem 1 and Remark 1 in Belhaj et al. [2016].

Observe that

$$\mathbb{E}\left(\frac{1+2b\frac{1}{n}\sum_{i=1}^{n}|\,\mathsf{S}(\mathbf{h}_i)|+b^2\frac{1}{n}|\,\mathsf{S}(\mathbf{h}_i)|^2}{1+b\frac{1}{n}\sum_{i=1}^{n}|\,\mathsf{S}(\mathbf{h}_i)|}\right)=n\left(\frac{\mathbb{E}\left(1+2b\frac{1}{n}\sum_{i=1}^{n}|\,\mathsf{S}(\mathbf{h}_i)|+b^2\frac{1}{n}|\,\mathsf{S}(\mathbf{h}_i)|^2\right)}{\mathbb{E}\left(n+b\sum_{i=1}^{n}|\,\mathsf{S}(\mathbf{h}_i)|\right)}+O\left(\frac{1}{n^2}\right)\right),$$

where we have used the fact that $\mathbb{E}\left(\frac{X}{Y}\right)=\frac{\mathbb{E}(X)}{\mathbb{E}(Y)}+O\left(\frac{1}{\mathbb{E}(Y)^2}\right)$ (see also the proof of Proposition 1). We then can write

$$\mathbb{E}\left(\overline{Y}_0(\mathbf{h},G)\big|\,G\right)=n\frac{1+2b\frac{1}{n}\sum_{i=1}^{n}\mathbb{E}\left(|\,\mathsf{S}(\mathbf{h}_i)|\,|\,G\right)+b^2\frac{1}{n}\sum_{i=1}^{n}\mathbb{E}\left(|\,\mathsf{S}(\mathbf{h}_i)|^2\big|\,G\right)}{1+b\frac{1}{n}\sum_{i=1}^{n}\mathbb{E}\left(|\,\mathsf{S}(\mathbf{h}_i)|\,|\,G\right)}+O(1)-2mc.$$

By part (i) of the proposition, we know that the asymptotic knowledge stocks, $|\,\mathsf{S}(\mathbf{h}_i)|$, follow a binomial distribution, $\text{Binom}(N,p_i)$, with success probability $p_i=\mathbb{E}(h_{ik}|\,G)$ and mean Np_i. Hence, we have that $\mathbb{E}\left(|\,\mathsf{S}(\mathbf{h}_i)|\,|\,G\right)=Np_i$ and $\mathbb{E}\left(|\,\mathsf{S}(\mathbf{h}_i)^2\big|\,G\right)=Np_i(1+(N-1)p_i)$. Inserting yields

$$\mathbb{E}\left(\overline{Y}_0(\mathbf{h},G)\big|\,G\right)=n\left(\frac{1+b(2+b)\frac{N}{n}\sum_{i=1}^{n}p_i+b^2\frac{N(N-1)}{n}\sum_{i=1}^{n}p_i^2}{1+b\frac{N}{n}\sum_{i=1}^{n}p_i}\right)+O(1)-2mc.$$

With the above expression for expected net welfare we then get Equation (71). This concludes the proof of the proposition. □

H.3. Inequality

It is often assumed that knowledge spillovers reduce inequality, and that through spillovers followers will eventually catch up with technological leaders. Recent work, however, has stressed the role that technological progress plays in generating inequality, and that knowledge spillovers might actually promote inequality [cf. Eeckhout and Jovanovic, 2002; Jovanovic, 2009]. For example, while innovation is not the only source of the increasing inequality in the U.S., Aghion et al. [2015] show that about 17% of the overall increase in inequality can be attributed to innovations.

Here we investigate the role of technological progress on inequality through in-house R&D and spillovers among R&D collaborating firms.[66] Figure H.3 shows the variance in the knowledge stocks, σ_h^2 and the number of active firms, n_a, in the stationary state for $N=2$ technologies as a function of the spillover parameter β for different values of the in-house innovate rate γ. We observe that both, an increase in the in-house R&D success probability γ as well as an increase in the technology spillover probability β, raise the variance σ_h^2. Inequality, measured by the variance σ_h^2, is thus increased through both γ and β.[67] Similarly, the number of active firms, n_a, falls with both β and γ. This is because potential entrants expect lower returns from entry through more intense competition stemming from a higher average knowledge stock $\bar{h}$ of incumbent firms. Therefore, both, an increase in γ and β, lead to more inequality of incumbents over potential entrants.

As discussed in Section 7, higher inequality and the appearance of a few dominant incumbent firms, however, makes the economy less resilient to shocks (leading to the exit of firms), and thus can have negative effects on overall production and output.

[66] The average and the variance in the stocks of knowledge is given by $\bar{h}=\mathbb{E}(|\,\mathsf{S}(\mathbf{h}_i)|)$ and $\sigma_h^2=\frac{1}{n}\sum_{i=1}^{N}\left(|\,\mathsf{S}(\mathbf{h}_i)|-\bar{h}\right)^2=\mathbb{E}(|\,\mathsf{S}(\mathbf{h}_i)|^2)-\mathbb{E}(|\,\mathsf{S}(\mathbf{h}_i)|)^2$, respectively (considering here only the stationary state). With the endogenous variables $\bar{x}(s)$ we can compute the average and the variance of the knowledge stocks as $\bar{h}=\sum_{s=0}^{N}s\binom{N}{s}\bar{x}(s)$ and $\sigma_h^2=\sum_{s=0}^{N}(s-\bar{h})^2\binom{N}{s}\bar{x}(s)$. In particular, in the case of $N=2$ we have that $\bar{h}=2(\bar{x}(1)+\bar{x}(2))$ and $\sigma_h^2=\bar{h}^2(1-2\bar{x}(1)-\bar{x}(2))+2(1-\bar{h})^2\bar{x}(1)+(2-\bar{h})^2\bar{x}(2)$.

[67] However, the coefficient of variation, defined as the standard deviation over the mean, $\text{CV}_h=\sigma_h/\bar{h}$, is reduced both, with increasing γ and β.

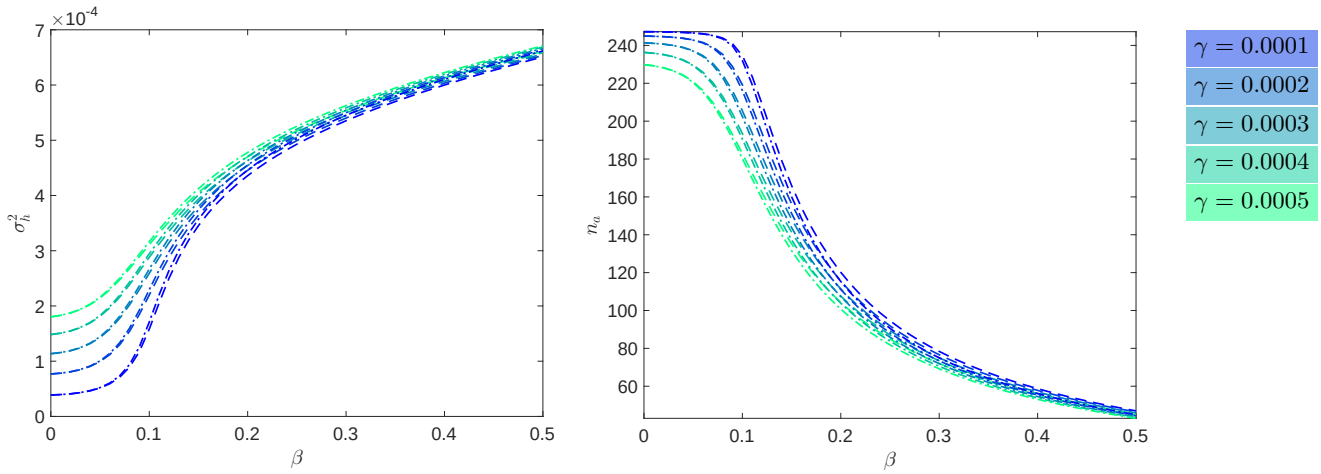


Figure H.3: (Left panel) The variance in the knowledge stocks, σ_h^2, as a function of β for different values of γ. (Right panel) The number of active firms, n_a, as a function of β for different values of γ. The number of technologies is assumed to be $N = 2$. The parameters used correspond to Figure 8. Dashed lines correspond to the large ρ approximation of Section G.2, while dashed-dotted lines represent the pair approximation of Appendix J.2.

I. Asymptotic Knowledge Stocks for Specific Graphs

In the following we provide two examples for which we explicitly compute the asymptotic knowledge stocks. First, we consider a k-regular graph.[68]

Corollary 1. *Assume that $\alpha = 0$, $\beta > 0$ and that pair-correlations can be neglected, i.e. $\lim_{t\to\infty} \mathrm{Cov}(h_{ik,t}, h_{jk,t}|G) = 0$ for all $i = 1, \dots, n$ and $k = 1, \dots, N$. Then in the k-regular graph the asymptotic knowledge stocks are given by*

$$\lim_{t\to\infty} \mathbb{E}\left(|\,\mathsf{S}(\mathbf{h}_{i,t})|\,|\, k\text{–reg. } G\right) = N\frac{\sqrt{2\lambda(\gamma-\beta k)+(\gamma+\beta k)^2+\lambda^2}-1+\beta k(\gamma+\lambda)}{2\beta k}.$$

Proof of Corollary 1. From Equation (74) and the symmetry implied by a k-regular graph we know that the

$$y = 1 - \frac{\frac{\lambda}{\lambda+\gamma}}{1+\frac{\beta}{\lambda+\gamma}ky},$$

for $0 \le k \le n-1$, where we have denoted by $y = \lim_{t\to\infty} \mathbb{E}(X_i(t)|G)$ and $X_i(t) = \mathbb{1}_{\{h_{ik,t}=1\}}$. Solving this equation delivers

$$y = \frac{1}{2} + \frac{\gamma+\lambda}{2\beta k}\left(\sqrt{\frac{2\lambda(\gamma-\beta k)+(\gamma+\beta k)^2+\lambda^2}{(\gamma+\lambda)^2}}-1\right).$$

□

Next, we consider the star $K_{1,n-1}$ in the following corollary.

Corollary 2. *Assume that $\alpha = 0$ and that pair-correlations can be neglected, i.e. $\lim_{t\to\infty} \mathrm{Cov}(h_{ik,t}, h_{jk,t}|G) = 0$ for all $i = 1, \dots, n$ and $k = 1, \dots, N$. Then in the star network $K_{1,n-1}$ the asymptotic knowledge stocks are given by*

$$\lim_{t\to\infty} \mathbb{E}\left(|\,\mathsf{S}(\mathbf{h}_{1,t})|\,|\, G\right) = \frac{N(\gamma+\lambda)^2}{2\beta(\gamma+\lambda+\beta(n-1))}\left(A+\frac{\beta(\beta(n-1)-\gamma(n-2))}{(\gamma+\lambda)^2}-1\right)$$
$$\lim_{t\to\infty} \mathbb{E}\left(|\,\mathsf{S}(\mathbf{h}_{j\neq 1,t})|\,|\, G\right) = \frac{N(\gamma+\lambda)^2}{2\beta(n-1)(\beta+\gamma+\lambda)}\left(A+\frac{\beta(\beta(n-1)+\gamma(n-2))}{(\gamma+\lambda)^2}-1\right),$$

[68] A discussion of the approximation with vanishing correlations and the exact analysis for the star and the complete graph considered here can be found in Cator and Van Mieghem [2013].

where

$$A \equiv \frac{\sqrt{4\gamma\lambda^3 + \lambda^4 + 2\lambda^2\left(3\gamma^2 + \beta^2(-(n-1)) + \beta\gamma n\right) + 4\gamma\lambda(\beta+\gamma)(\gamma+\beta(n-1)) + (\beta+\gamma)^2(\gamma+\beta(n-1))^2}}{(\gamma+\lambda)^2}.$$

Proof of Corollary 2. In the star $K_{1,n-1}$ we have two types of firms, the one in the center and the ones in the periphery. W.l.o.g. we denote by y_1 the asymptotic probability of the central firm to have knowledge of the technology, and by y_2 the corresponding probability of a firm in the periphery. From Equation (74) it then follows that $y_1 = 1 - \frac{\frac{\lambda}{\lambda+\gamma}}{1+\frac{\beta}{\lambda+\gamma}(n-1)y_2}$, $y_2 = 1 - \frac{\frac{\lambda}{\lambda+\gamma}}{1+\frac{\beta}{\lambda+\gamma}y_1}$. The solution is given by

$$y_1 = \frac{(\gamma+\lambda)^2}{2\beta(\gamma+\lambda+\beta(n-1))}\left(\frac{\beta(\beta(n-1)-\gamma(n-2))}{(\gamma+\lambda)^2} - 1\right.$$
$$\left. + \sqrt{\frac{4\gamma\lambda^3 + \lambda^4 + 2\lambda^2\left(3\gamma^2 + \beta^2(-(n-1)) + \beta\gamma n\right) + 4\gamma\lambda(\beta+\gamma)(\gamma+\beta(n-1)) + (\beta+\gamma)^2(\gamma+\beta(n-1))^2}{(\gamma+\lambda)^4}}\right)$$
$$y_2 = \frac{(\gamma+\lambda)^2}{2\beta(n-1)(\beta+\gamma+\lambda)}\left(\frac{\beta(\beta(n-1)+\gamma(n-2))}{(\gamma+\lambda)^2} - 1\right.$$
$$\left. + \sqrt{\frac{4\gamma\lambda^3 + \lambda^4 + 2\lambda^2\left(3\gamma^2 + \beta^2(-(n-1)) + \beta\gamma n\right) + 4\gamma\lambda(\beta+\gamma)(\gamma+\beta(n-1)) + (\beta+\gamma)^2(\gamma+\beta(n-1))^2}{(\gamma+\lambda)^4}}\right).$$

□

An illustration for the asymptotic knowledge stocks is given in Figure I.1 for the complete graph K_n and the star $K_{1,n-1}$. The figures shows that in both cases, with and without competition, the star, $K_{1,n-1}$, generates higher net output than the complete graph, K_n.

J. Pair Approximation

In this section we provide a complete derivation of the dynamics for an arbitrary number N of technology categories without making the assumption that ρ is large, so that we need to take into account the remainder term of the order of $o(\rho)$ in Equation (11) in Theorem 1.

Proposition 7. *Consider the parameters as in Theorem 1. Let the probability that a firm with technology vector* $\mathbf{h}$ *is connected to a firm with technology vector* $\mathbf{h}'$ *be denoted by* $y_t(\mathbf{h}, \mathbf{h}') \equiv \mathbb{P}(a_{ij,t} = 1 | \mathbf{h}_{it} = \mathbf{h}, \mathbf{h}_{jt} = \mathbf{h}')$*, and let the fraction of firms with knowledge vector* $\mathbf{h}$ *be* $x_t(\mathbf{h}) \equiv \mathbb{P}(\mathbf{h}_{it} = \mathbf{h})$*. Then, under the pair approximation of Equation (12),* $x_t(\mathbf{h})$ *converges in probability to the solution*

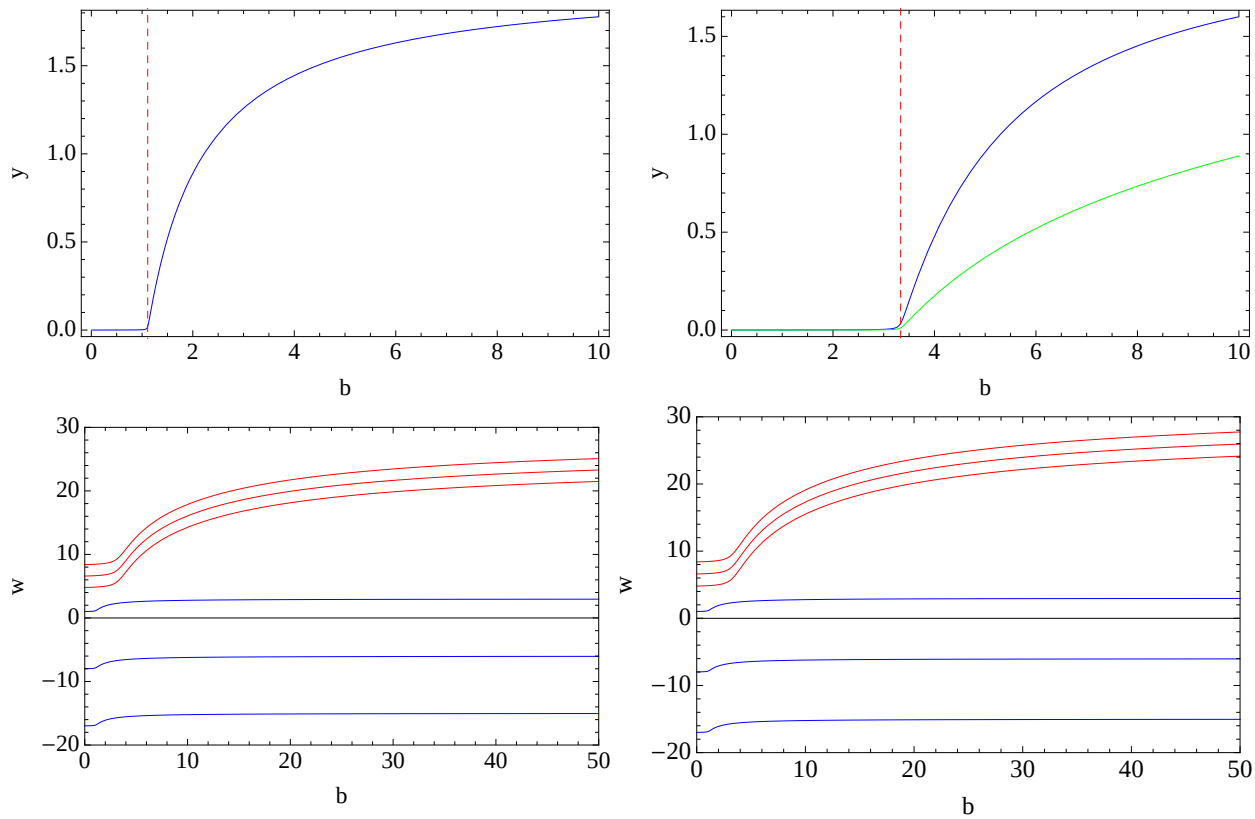


Figure I.1: (Top left panel) The asymptotic knowledge stock of the firms in the complete graph K_n for varying values of β with $\gamma = 0.001$, $\lambda = 10$, $N = 2$ and $n = 10$. The critical value for the spillover parameter β is $\beta^c = \frac{\gamma+\lambda}{\mu_1(K_n)} = \frac{\gamma+\lambda}{n-1}$. (Top right panel) The asymptotic knowledge stock of the firms in the star $K_{1,n-1}$ for varying values of β with $\gamma = 0.001$, $\lambda = 10$, $N = 2$ and $n = 10$. The critical value for the spillover parameter β is $\beta^c = \frac{\gamma+\lambda}{\mu_1(K_{1,n-1})} = \frac{\gamma+\lambda}{\sqrt{n-1}}$. (Bottom left panel) Asymptotic net output for the complete graph K_n (blue) and the star $K_{1,n-1}$ (red) as a function of β for different values of the linking cost $c = 0.01, 0.02, 0.03$ when $\theta = 1$. (Bottom right panel) Asymptotic net output for the complete graph K_n (blue) and the star $K_{1,n-1}$ (red) as a function of β for different values of the linking cost $c \in \{0.01, 0.02, 0.03\}$ when $\theta = 0$.

of Equation (10), and $y_t(\mathbf{h},\mathbf{h}')$ converges in probability to the solution of the ODE

$$
\begin{aligned}
\frac{dy_t(\mathbf{h},\mathbf{h}')}{dt} &= \rho g^{\eta}_{\theta,\tau}(\mathbf{h},\mathbf{h}';\bar{h}_t)\left(1-y_t(\mathbf{h},\mathbf{h}')\right) - \rho\, y_t(\mathbf{h},\mathbf{h}') \\
&+ \sum_{k=1}^{N} \mathbb{1}_{\{h_k=1\}}\left(\gamma+\alpha\langle \mathbf{h}-\mathbf{e}_k,\mathbf{u}\rangle\right)\frac{x_t(\mathbf{h}-\mathbf{e}_k)}{x_t(\mathbf{h})}\left(y_t(\mathbf{h}-\mathbf{e}_k,\mathbf{h}')-y_t(\mathbf{h},\mathbf{h}')\right) \\
&+ \lambda\sum_{k=1}^{N} \mathbb{1}_{\{h_k=0\}}\frac{x_t(\mathbf{h}+\mathbf{e}_k)}{x_t(\mathbf{h})}\left(y_t(\mathbf{h}+\mathbf{e}_k,\mathbf{h}')-y_t(\mathbf{h},\mathbf{h}')\right) \\
&+ n\beta\sum_{k=1}^{N} \mathbb{1}_{\{h_k=1\}}\left(y_t(\mathbf{h}-\mathbf{e}_k,\mathbf{h}')-y_t(\mathbf{h},\mathbf{h}')\right)\sum_{\mathbf{h}''\in\mathcal{H}^N:h''_k=1} y_t(\mathbf{h}-\mathbf{e}_k,\mathbf{h}'')\frac{x_t(\mathbf{h}-\mathbf{e}_k)x_t(\mathbf{h}'')}{x_t(\mathbf{h})} \\
&+ \sum_{k=1}^{N} \mathbb{1}_{\{h'_k=1\}}\left(\gamma+\alpha\langle \mathbf{h}'-\mathbf{e}_k,\mathbf{u}\rangle\right)\frac{x_t(\mathbf{h}'-\mathbf{e}_k)}{x_t(\mathbf{h}')}\left(y_t(\mathbf{h},\mathbf{h}'-\mathbf{e}_k)-y_t(\mathbf{h},\mathbf{h}')\right) \\
&+ \lambda\sum_{k=1}^{N} \mathbb{1}_{\{h'_k=0\}}\frac{x_t(\mathbf{h}'+\mathbf{e}_k)}{x_t(\mathbf{h}')}\left(y_t(\mathbf{h},\mathbf{h}'+\mathbf{e}_k)-y_t(\mathbf{h},\mathbf{h}')\right) \\
&+ n\beta\sum_{k=1}^{N} \mathbb{1}_{\{h'_k=1\}}\left(y_t(\mathbf{h}'-\mathbf{e}_k,\mathbf{h})-y_t(\mathbf{h},\mathbf{h}')\right)\sum_{\mathbf{h}''\in\mathcal{H}^N:h''_k=1} y_t(\mathbf{h}'-\mathbf{e}_k,\mathbf{h}'')\frac{x_t(\mathbf{h}'-\mathbf{e}_k)x_t(\mathbf{h}'')}{x_t(\mathbf{h}')} \\
&- 2\xi x_t(\mathbf{h})x_t(\mathbf{h}')y_t(\mathbf{h},\mathbf{h}'). \qquad (77)
\end{aligned}
$$

Proof of Proposition 7. In the following we also consider contributions to $m_t(\mathbf{h},\mathbf{h}')$ originating from changes in the technologies $\mathbf{h}$ and $\mathbf{h}'$ of the firms that are of the order of $o(\rho)$ in Equation (31):

$$
\begin{aligned}
&n^2\mathsf{F}_z(\mathbf{h},\mathbf{h}') \equiv \\
&\left.n_t(\mathbf{h})n_t(\mathbf{h}')\rho\, g^{\eta}_{\theta,\tau}(\mathbf{h},\mathbf{h}';\bar{h}_t)\left(1-\frac{m_t(\mathbf{h},\mathbf{h}')}{n_t(\mathbf{h})n_t(\mathbf{h}')}\right)-\rho\, m_t(\mathbf{h},\mathbf{h}')\right\} A \\
&\left.\begin{aligned}
&+\sum_{k=1}^{N}\mathbb{1}_{\{h_k=1\}}\left(\gamma+\alpha|\,\mathsf{S}(\mathbf{h}-\mathbf{e}_k)|\right)n_t(\mathbf{h}-\mathbf{e}_k)\frac{m_t(\mathbf{h}-\mathbf{e}_k,\mathbf{h}')}{n_t(\mathbf{h}-\mathbf{e}_k)}-\sum_{k=1}^{N}\mathbb{1}_{\{h_k=0\}}\left(\gamma+\alpha|\,\mathsf{S}(\mathbf{h})|\right)n_t(\mathbf{h})\frac{m_t(\mathbf{h},\mathbf{h}')}{n_t(\mathbf{h})} \\
&+\lambda\sum_{k=1}^{N}\mathbb{1}_{\{h_k=0\}}n_t(\mathbf{h}+\mathbf{e}_k)\frac{m_t(\mathbf{h}+\mathbf{e}_k,\mathbf{h}')}{n_t(\mathbf{h}+\mathbf{e}_k)}-\lambda\sum_{k=1}^{N}\mathbb{1}_{\{h_k=1\}}n_t(\mathbf{h})\frac{m_t(\mathbf{h},\mathbf{h}')}{n_t(\mathbf{h})}
\end{aligned}\right\} B \\
&\left.\begin{aligned}
&+\beta\sum_{k=1}^{N}\mathbb{1}_{\{h_k=1\}}\sum_{\mathbf{h}''\in\mathcal{H}^N:h''_k=1} m_t(\mathbf{h}-\mathbf{e}_k,\mathbf{h}'')\left[\mathbb{1}_{\{\mathbf{h}''\neq\mathbf{h}'\}}\frac{\Delta_t(\mathbf{h}'',\mathbf{h}-\mathbf{e}_k,\mathbf{h}')}{m_t(\mathbf{h}-\mathbf{e}_k,\mathbf{h}'')}+\mathbb{1}_{\{\mathbf{h}''=\mathbf{h}'\}}\left(1+\frac{\Delta_t(\mathbf{h}',\mathbf{h}-\mathbf{e}_k,\mathbf{h}')}{m_t(\mathbf{h}-\mathbf{e}_k,\mathbf{h}')}\right)\right] \\
&-\beta\sum_{k=1}^{N}\mathbb{1}_{\{h_k=0\}}\sum_{\mathbf{h}''\in\mathcal{H}^N:h''_k=1} m_t(\mathbf{h},\mathbf{h}'')\left[\mathbb{1}_{\{\mathbf{h}''\neq\mathbf{h}'\}}\frac{\Delta_t(\mathbf{h}'',\mathbf{h},\mathbf{h}')}{m_t(\mathbf{h},\mathbf{h}'')}+\mathbb{1}_{\{\mathbf{h}''=\mathbf{h}'\}}\left(1+\frac{\Delta_t(\mathbf{h}',\mathbf{h},\mathbf{h}')}{m_t(\mathbf{h},\mathbf{h}')}\right)\right]
\end{aligned}\right\} C
\end{aligned}
$$

$$
\left.\begin{aligned}
&+\sum_{k=1}^{N}\mathbb{1}_{\{h'_k=1\}}\left(\gamma+\alpha|\,\mathsf{S}(\mathbf{h}'-\mathbf{e}_k)|\right)n_t(\mathbf{h}'-\mathbf{e}_k)\frac{m_t(\mathbf{h},\mathbf{h}'-\mathbf{e}_k)}{n_t(\mathbf{h}'-\mathbf{e}_k)}-\sum_{k=1}^{N}\mathbb{1}_{\{h'_k=0\}}\left(\gamma+\alpha|\,\mathsf{S}(\mathbf{h}')|\right)n_t(\mathbf{h}')\frac{m_t(\mathbf{h},\mathbf{h}')}{n_t(\mathbf{h}')}\\
&+\lambda\sum_{k=1}^{N}\mathbb{1}_{\{h'_k=0\}}n_t(\mathbf{h}'+\mathbf{e}_k)\frac{m_t(\mathbf{h},\mathbf{h}'+\mathbf{e}_k)}{n_t(\mathbf{h}'+\mathbf{e}_k)}-\lambda\sum_{k=1}^{N}\mathbb{1}_{\{h'_k=1\}}n_t(\mathbf{h}')\frac{m_t(\mathbf{h},\mathbf{h}')}{n_t(\mathbf{h}')}
\end{aligned}\right\}B'
$$

$$
\left.\begin{aligned}
&+\beta\sum_{k=1}^{N}\mathbb{1}_{\{h'_k=1\}}\sum_{\mathbf{h}''\in\mathcal{H}^N:h''_k=1}m_t(\mathbf{h}'-\mathbf{e}_k,\mathbf{h}'')\left[\mathbb{1}_{\{\mathbf{h}''\neq\mathbf{h}\}}\frac{m_t(\mathbf{h},\mathbf{h}'-\mathbf{e}_k)}{n_t(\mathbf{h}'-\mathbf{e}_k)}+\mathbb{1}_{\{\mathbf{h}''=\mathbf{h}\}}\left(1+\frac{\Delta_t(\mathbf{h},\mathbf{h}'-\mathbf{e}_k,\mathbf{h})}{m_t(\mathbf{h}'-\mathbf{e}_k,\mathbf{h})}\right)\right]\\
&-\beta\sum_{k=1}^{N}\mathbb{1}_{\{h'_k=0\}}\sum_{\mathbf{h}''\in\mathcal{H}^N:h''_k=1}m_t(\mathbf{h}',\mathbf{h}'')\left[\mathbb{1}_{\{\mathbf{h}''\neq\mathbf{h}\}}\frac{m_t(\mathbf{h}',\mathbf{h})}{n_t(\mathbf{h}')}+\mathbb{1}_{\{\mathbf{h}''=\mathbf{h}\}}\left(1+\frac{\Delta_t(\mathbf{h},\mathbf{h}',\mathbf{h})}{m_t(\mathbf{h}',\mathbf{h})}\right)\right]
\end{aligned}\right\}C'
$$

$$
-\xi n_t(\mathbf{h})\frac{m_t(\mathbf{h},\mathbf{h}')}{n_t(\mathbf{h})}-\xi n_t(\mathbf{h}')\frac{m_t(\mathbf{h},\mathbf{h}')}{n_t(\mathbf{h}')}\Bigg\}D \tag{78}
$$

We now explain each of the terms on the RHS in Equation (78). Part A takes into account the contribution due to link creation or removal. The rate at which links between firms with technologies $\mathbf{h}$ and $\mathbf{h}'$ decay is given by $\rho n_t(\mathbf{h})n_t(\mathbf{h}')\frac{m_t(\mathbf{h},\mathbf{h}')}{n_t(\mathbf{h})n_t(\mathbf{h}')}$, where $n_t(\mathbf{h})n_t(\mathbf{h}')$ is the expected number of pairs of firms with technologies $\mathbf{h}$ and $\mathbf{h}'$ that are selected, and $\frac{m_t(\mathbf{h},\mathbf{h}')}{n_t(\mathbf{h})n_t(\mathbf{h}')}$ is the probability that a link exists between them. Similarly, the rate at which such links are created is given by $\rho n_t(\mathbf{h})n_t(\mathbf{h}')g^{\eta}_{\theta,\tau}(\mathbf{h},\mathbf{h}';\bar{h}_t)\left(1-\frac{m_t(\mathbf{h},\mathbf{h}')}{n_t(\mathbf{h})n_t(\mathbf{h}')}\right)$, where $1-\frac{m_t(\mathbf{h},\mathbf{h}')}{n_t(\mathbf{h})n_t(\mathbf{h}')}$ is the probability that a link does not exist between the firms with technologies $\mathbf{h}$ and $\mathbf{h}'$, and $g^{\eta}_{\theta,\tau}(\mathbf{h},\mathbf{h}';\bar{h}_t)$ is the probability that they want to form a link when they have the opportunity.

The remaining parts, B, C, B' and C' capture contributions stemming from changes in the technologies $\mathbf{h}$ and $\mathbf{h}'$ of the firms. Moreover, part D incorporates the effect of firm exit.

First, we consider part B which captures either gains due to the discovery of $\mathbf{h}$ by a firm with technology $\mathbf{h}-\mathbf{e}_k$, gains through obsolescence of idea k of a firm with technology $\mathbf{h}+\mathbf{e}_k$, losses due to successful in-house R&D of a firm with technology $\mathbf{h}$, or losses due to obsolescence of an idea of a firm with technology $\mathbf{h}$. The rate at which the first happens through in-house R&D is given by $(\gamma+\alpha|\,\mathsf{S}(\mathbf{h}-\mathbf{e}_k)|)\,n_t(\mathbf{h}-\mathbf{e}_k)$. Moreover, the expected number of links to firms with technologies $\mathbf{h}'$ in which a firm with technology $\mathbf{h}-\mathbf{e}_k$ is involved is given by $\frac{m_t(\mathbf{h}-\mathbf{e}_k,\mathbf{h}')}{n_t(\mathbf{h}-\mathbf{e}_k)}$. Summation over all $k=1,\ldots,N$ gives the first equation in B. In the same way, the second term captures the rate of decline through a firm with technology $\mathbf{h}$ learning about a new idea k it does not possess yet, i.e. $h_k=0$, which happens at a rate $(\gamma+\alpha|\,\mathsf{S}(\mathbf{h})|)\,n_t(\mathbf{h})$ and the expected number of links to firms with technology $\mathbf{h}'$ involving the firm with technology $\mathbf{h}$ given by $\frac{m_t(\mathbf{h},\mathbf{h}')}{n_t(\mathbf{h})}$. The third term captures the loss of an idea of a firm with technology $\mathbf{h}+\mathbf{e}_k$, which happens at a rate $\lambda n_t(\mathbf{h}+\mathbf{e}_k)$, and the expected number of links to firms with technology $\mathbf{h}'$ involving a firm with technology $\mathbf{h}$ being given by $\frac{m_t(\mathbf{h}+\mathbf{e}_k,\mathbf{h}')}{n_t(\mathbf{h}+\mathbf{e}_k)}$. Finally, we need to consider the loss of an idea k by a firm with technology $\mathbf{h}$, which happens at a rate $\lambda n_t(\mathbf{h})$, and the expected number of links to firms with technology $\mathbf{h}'$ involving a firm with technology $\mathbf{h}$ being given by $\frac{m_t(\mathbf{h},\mathbf{h}')}{n_t(\mathbf{h})}$. Summation over all $k=1,\ldots,N$ gives the last equation in B.

Part C captures contributions due to technology spillovers. The first equation corresponds to a firm with technology $\mathbf{h}-\mathbf{e}_k$ learning about the idea k from linked firms with technology $\mathbf{h}''$ which have the idea k, i.e. $h''_k=1$. The rate at which this happens is $\beta m_t(\mathbf{h}-\mathbf{e}_k,\mathbf{h}'')$. We then need to consider two cases. First, assume that $\mathbf{h}''\neq\mathbf{h}'$. Then the expected number of links created is given by the expected number of links to firms with technology $\mathbf{h}'$ involving a firm with technology $\mathbf{h}-\mathbf{e}_k$, which is $\frac{m_t(\mathbf{h}-\mathbf{e}_k,\mathbf{h}')}{n_t(\mathbf{h}-\mathbf{e}_k)}$. Second, assume that $\mathbf{h}''=\mathbf{h}'$. Then (at least) one link between a firm with technology $\mathbf{h}$ and technology $\mathbf{h}'$ is created. Additional links are created if the firm with technology $\mathbf{h}-\mathbf{e}_k$, which has learned from the firm with technology $\mathbf{h}'$, has other neighbors with technology $\mathbf{h}'$. The number of such neighbors for each link between a firm with technology $\mathbf{h}-\mathbf{e}_k$ and a firm with technology $\mathbf{h}'$ is given by $\frac{\Delta_t(\mathbf{h}',\mathbf{h}-\mathbf{e}_k,\mathbf{h}')}{m_t(\mathbf{h}-\mathbf{e}_k,\mathbf{h}')}$. Summation over all $k=1,\ldots,N$ and technologies $\mathbf{h}''$ with $h''_k=1$ gives the first equation in part C.

The second equation in part C captures the losses from a firm with technology $\mathbf{h}$ which is connected to a firm with technology $\mathbf{h}'$ that learns about a new idea k (that is $h_k = 0$) from a linked firm with technology $\mathbf{h}''$ with $h''_k = 1$. Similar to the discussion in the previous paragraph, the rate at which this happens is $\beta m_t(\mathbf{h}, \mathbf{h}'')$ times the expected number of links to firms with technology $\mathbf{h}'$ involving a firm with technology $\mathbf{h}$, which is $\frac{m_t(\mathbf{h},\mathbf{h}')}{n_t(\mathbf{h})}$ for all $\mathbf{h}'' \neq \mathbf{h}'$. Moreover, when $\mathbf{h}'' = \mathbf{h}'$ additional links are created if the firm with technology $\mathbf{h}$, which has learned from the firm with technology $\mathbf{h}'$, has other neighbors with technology $\mathbf{h}'$. The number of such neighbors for each link between a firm with technology $\mathbf{h}$ and a firm with technology $\mathbf{h}'$ is given by $\frac{\Delta_t(\mathbf{h}',\mathbf{h},\mathbf{h}')}{m_t(\mathbf{h},\mathbf{h}')}$. Summation over all $k = 1, \ldots, N$ and technologies $\mathbf{h}''$ with $h''_k = 1$ gives the second equation in part C.

Part B' is identical to part B but with the roles of $\mathbf{h}$ and $\mathbf{h}'$ exchanged. Similarly, part C' is identical to C but with $\mathbf{h}$ and $\mathbf{h}'$ exchanged.

Part D captures the exit of a firm with technology portfolio $\mathbf{h}$, which happens at a rate ξ, and involves $\frac{m_t(\mathbf{h},\mathbf{h}')}{n_t(\mathbf{h})}$ number of links to firms with technology portfolios $\mathbf{h}'$ in expectation, and the exit of a firm with technology portfolio $\mathbf{h}'$, which happens at a rate ξ, and involves $\frac{m_t(\mathbf{h},\mathbf{h}')}{n_t(\mathbf{h}')}$ number of links to firms with technology portfolios $\mathbf{h}$ in expectation.

In the following we make a *pair approximation* as in Gross et al. [2006]; Keeling and Eames [2005]; Miller and Kiss [2014]; Rogers et al. [2012] for the number of triplets, $\Delta_t(\mathbf{h}, \mathbf{h}', \mathbf{h}'')$, where a firm with technology $\mathbf{h}$ is linked to a firm with technology $\mathbf{h}'$, and this firm is linked to a firm with technology $\mathbf{h}''$:

$$\Delta_t(\mathbf{h}, \mathbf{h}', \mathbf{h}'') \approx \frac{m_t(\mathbf{h}, \mathbf{h}')m_t(\mathbf{h}', \mathbf{h}'')}{n_t(\mathbf{h}')}.$$

In particular, we then obtain

$$\frac{\Delta_t(\mathbf{h}, \mathbf{h}', \mathbf{h})}{m_t(\mathbf{h}', \mathbf{h})} \approx \frac{m_t(\mathbf{h}, \mathbf{h}')}{n_t(\mathbf{h}')} = n\frac{z_t(\mathbf{h}, \mathbf{h}')}{x_t(\mathbf{h}')}.$$

Dividing by n^2 we then can write Equation (78) as

$$
\begin{aligned}
\mathsf{F}_z(\mathbf{h},\mathbf{h}') &= \rho\, x_t(\mathbf{h})x_t(\mathbf{h}')\left[g^{\eta}_{\theta,\tau}(\mathbf{h},\mathbf{h}';\bar{h}_t)\left(1-\frac{z_t(\mathbf{h},\mathbf{h}')}{x_t(\mathbf{h})x_t(\mathbf{h}')}\right)-\frac{z_t(\mathbf{h},\mathbf{h}')}{x_t(\mathbf{h})x_t(\mathbf{h}')}\right]\\
&+\sum_{k=1}^{N}\mathbb{1}_{\{h_k=1\}}\left(\gamma+\alpha|\,\mathsf{S}(\mathbf{h}-\mathbf{e}_k)|\right)x_t(\mathbf{h}-\mathbf{e}_k)\frac{z_t(\mathbf{h}-\mathbf{e}_k,\mathbf{h}')}{x_t(\mathbf{h}-\mathbf{e}_k)}-\sum_{k=1}^{N}\mathbb{1}_{\{h_k=0\}}\left(\gamma+\alpha|\,\mathsf{S}(\mathbf{h})|\right)x_t(\mathbf{h})\frac{z_t(\mathbf{h},\mathbf{h}')}{x_t(\mathbf{h})}\\
&+\lambda\sum_{k=1}^{N}\mathbb{1}_{\{h_k=0\}}x_t(\mathbf{h}+\mathbf{e}_k)\frac{z_t(\mathbf{h}+\mathbf{e}_k,\mathbf{h}')}{x_t(\mathbf{h}+\mathbf{e}_k)}-\lambda\sum_{k=1}^{N}\mathbb{1}_{\{h_k=1\}}x_t(\mathbf{h})\frac{z_t(\mathbf{h},\mathbf{h}')}{x_t(\mathbf{h})}\\
&+n\beta\sum_{k=1}^{N}\mathbb{1}_{\{h_k=1\}}\sum_{\mathbf{h}''\in\mathcal{H}^N:h''_k=1}z_t(\mathbf{h}-\mathbf{e}_k,\mathbf{h}'')\left[\mathbb{1}_{\{\mathbf{h}''\neq\mathbf{h}'\}}\frac{z_t(\mathbf{h}-\mathbf{e}_k,\mathbf{h}')}{x_t(\mathbf{h}-\mathbf{e}_k)}+\mathbb{1}_{\{\mathbf{h}''=\mathbf{h}'\}}\left(\frac{1}{n}+\frac{z_t(\mathbf{h}',\mathbf{h}-\mathbf{e}_k)}{x_t(\mathbf{h}-\mathbf{e}_k)}\right)\right]\\
&-n\beta\sum_{k=1}^{N}\mathbb{1}_{\{h_k=0\}}\sum_{\mathbf{h}''\in\mathcal{H}^N:h''_k=1}z_t(\mathbf{h},\mathbf{h}'')\left[\mathbb{1}_{\{\mathbf{h}''\neq\mathbf{h}'\}}\frac{z_t(\mathbf{h},\mathbf{h}')}{x_t(\mathbf{h})}+\mathbb{1}_{\{\mathbf{h}''=\mathbf{h}'\}}\left(\frac{1}{n}+\frac{z_t(\mathbf{h},\mathbf{h}')}{x_t(\mathbf{h})}\right)\right]\\
&+\sum_{k=1}^{N}\mathbb{1}_{\{h'_k=1\}}\left(\gamma+\alpha|\,\mathsf{S}(\mathbf{h}'-\mathbf{e}_k)|\right)x_t(\mathbf{h}'-\mathbf{e}_k)\frac{z_t(\mathbf{h},\mathbf{h}'-\mathbf{e}_k)}{x_t(\mathbf{h}'-\mathbf{e}_k)}-\sum_{k=1}^{N}\mathbb{1}_{\{h'_k=0\}}\left(\gamma+\alpha|\,\mathsf{S}(\mathbf{h}')|\right)x_t(\mathbf{h}')\frac{z_t(\mathbf{h},\mathbf{h}')}{x_t(\mathbf{h}')}\\
&+\lambda\sum_{k=1}^{N}\mathbb{1}_{\{h'_k=0\}}x_t(\mathbf{h}'+\mathbf{e}_k)\frac{z_t(\mathbf{h},\mathbf{h}'+\mathbf{e}_k)}{x_t(\mathbf{h}'+\mathbf{e}_k)}-\lambda\sum_{k=1}^{N}\mathbb{1}_{\{h'_k=1\}}x_t(\mathbf{h}')\frac{z_t(\mathbf{h},\mathbf{h}')}{x_t(\mathbf{h}')}\\
&+n\beta\sum_{k=1}^{N}\mathbb{1}_{\{h'_k=1\}}\sum_{\mathbf{h}''\in\mathcal{H}^N:h''_k=1}z_t(\mathbf{h}'-\mathbf{e}_k,\mathbf{h}'')\left[\mathbb{1}_{\{\mathbf{h}''\neq\mathbf{h}\}}\frac{z_t(\mathbf{h},\mathbf{h}'-\mathbf{e}_k)}{x_t(\mathbf{h}'-\mathbf{e}_k)}+\mathbb{1}_{\{\mathbf{h}''=\mathbf{h}\}}\left(\frac{1}{n}+\frac{z_t(\mathbf{h},\mathbf{h}'-\mathbf{e}_k)}{x_t(\mathbf{h}'-\mathbf{e}_k)}\right)\right]\\
&-n\beta\sum_{k=1}^{N}\mathbb{1}_{\{h'_k=0\}}\sum_{\mathbf{h}''\in\mathcal{H}^N:h''_k=1}z_t(\mathbf{h}',\mathbf{h}'')\left[\mathbb{1}_{\{\mathbf{h}''\neq\mathbf{h}\}}\frac{z_t(\mathbf{h}',\mathbf{h})}{x_t(\mathbf{h}')}+\mathbb{1}_{\{\mathbf{h}''=\mathbf{h}\}}\left(\frac{1}{n}+\frac{z_t(\mathbf{h},\mathbf{h}')}{x_t(\mathbf{h}')}\right)\right]\\
&-\xi x_t(\mathbf{h})\frac{z_t(\mathbf{h},\mathbf{h}')}{x_t(\mathbf{h})}-\xi x_t(\mathbf{h}')\frac{z_t(\mathbf{h},\mathbf{h}')}{x_t(\mathbf{h}')}. \qquad (79)
\end{aligned}
$$

Using the fact that $\sum_{k=1}^{N} \mathbb{1}_{\{h_k=0\}}\mathbb{1}_{\{h'_k=1\}} = \langle \mathbf{h}^c, \mathbf{h}' \rangle$, and $\sum_{k=1}^{N} \mathbb{1}_{\{h'_k=0\}}\mathbb{1}_{\{h_k=1\}} = \langle \mathbf{h}'^c, \mathbf{h} \rangle$, we obtain

$$
\begin{aligned}
\mathsf{F}_z(\mathbf{h},\mathbf{h}') = \rho\, x_t(\mathbf{h})x_t(\mathbf{h}') &\left[g^{\eta}_{\theta,\tau}(\mathbf{h},\mathbf{h}';\bar{h}_t)\left(1 - \frac{z_t(\mathbf{h},\mathbf{h}')}{x_t(\mathbf{h})x_t(\mathbf{h}')}\right) - \frac{z_t(\mathbf{h},\mathbf{h}')}{x_t(\mathbf{h})x_t(\mathbf{h}')}\right] \\
&+ \frac{z_t(\mathbf{h},\mathbf{h}')}{x_t(\mathbf{h})}\frac{dx_t(\mathbf{h})}{dt} + \sum_{k=1}^{N} \mathbb{1}_{\{h_k=1\}}\left(\gamma + \alpha\langle \mathbf{h}-\mathbf{e}_k, \mathbf{u}\rangle\right) x_t(\mathbf{h}-\mathbf{e}_k)\left(\frac{z_t(\mathbf{h}-\mathbf{e}_k,\mathbf{h}')}{x_t(\mathbf{h}-\mathbf{e}_k)} - \frac{z_t(\mathbf{h},\mathbf{h}')}{x_t(\mathbf{h})}\right) \\
&+ \lambda \sum_{k=1}^{N} \mathbb{1}_{\{h_k=0\}} x_t(\mathbf{h}+\mathbf{e}_k)\left(\frac{z_t(\mathbf{h}+\mathbf{e}_k,\mathbf{h}')}{x_t(\mathbf{h}+\mathbf{e}_k)} - \frac{z_t(\mathbf{h},\mathbf{h}')}{x_t(\mathbf{h})}\right) \\
&+ n\beta \sum_{k=1}^{N} \mathbb{1}_{\{h_k=1\}}\left(\frac{z_t(\mathbf{h}-\mathbf{e}_k,\mathbf{h}')}{x_t(\mathbf{h}-\mathbf{e}_k)} - \frac{z_t(\mathbf{h},\mathbf{h}')}{x_t(\mathbf{h})}\right) \sum_{\mathbf{h}''\in\mathcal{H}^N : h''_k=1} z_t(\mathbf{h}-\mathbf{e}_k,\mathbf{h}'') \\
&+ \frac{z_t(\mathbf{h},\mathbf{h}')}{x_t(\mathbf{h}')}\frac{dx_t(\mathbf{h}')}{dt} + \sum_{k=1}^{N} \mathbb{1}_{\{h'_k=1\}}\left(\gamma + \alpha\langle \mathbf{h}'-\mathbf{e}_k, \mathbf{u}\rangle\right) x_t(\mathbf{h}'-\mathbf{e}_k)\left(\frac{z_t(\mathbf{h},\mathbf{h}'-\mathbf{e}_k)}{x_t(\mathbf{h}'-\mathbf{e}_k)} - \frac{z_t(\mathbf{h},\mathbf{h}')}{x_t(\mathbf{h}')}\right) \\
&+ \lambda \sum_{k=1}^{N} \mathbb{1}_{\{h'_k=0\}} x_t(\mathbf{h}'+\mathbf{e}_k)\left(\frac{z_t(\mathbf{h},\mathbf{h}'+\mathbf{e}_k)}{x_t(\mathbf{h}'+\mathbf{e}_k)} - \frac{z_t(\mathbf{h},\mathbf{h}')}{x_t(\mathbf{h}')}\right) \\
&+ n\beta \sum_{k=1}^{N} \mathbb{1}_{\{h'_k=1\}}\left(\frac{z_t(\mathbf{h}'-\mathbf{e}_k,\mathbf{h})}{x_t(\mathbf{h}'-\mathbf{e}_k)} - \frac{z_t(\mathbf{h},\mathbf{h}')}{x_t(\mathbf{h}')}\right) \sum_{\mathbf{h}''\in\mathcal{H}^N : h''_k=1} z_t(\mathbf{h}'-\mathbf{e}_k,\mathbf{h}'') \\
&+ \beta \sum_{k=1}^{N} \mathbb{1}_{\{h_k=1\}}\mathbb{1}_{\{h'_k=1\}}\left(z_t(\mathbf{h}-\mathbf{e}_k,\mathbf{h}') + z_t(\mathbf{h}'-\mathbf{e}_k,\mathbf{h})\right) - \beta z_t(\mathbf{h},\mathbf{h}')(\langle \mathbf{h}^c,\mathbf{h}'\rangle + \langle \mathbf{h}'^c,\mathbf{h}\rangle) \\
&- \xi x_t(\mathbf{h})\frac{z_t(\mathbf{h},\mathbf{h}')}{x_t(\mathbf{h})} - \xi x_t(\mathbf{h}')\frac{z_t(\mathbf{h},\mathbf{h}')}{x_t(\mathbf{h}')}. \qquad (80)
\end{aligned}
$$

Using the fact that $\operatorname*{plim}_{n\to\infty} \frac{dz_t(\mathbf{h},\mathbf{h}')}{dt} = \mathsf{F}_z(\mathbf{h},\mathbf{h}')$,[69] and dropping terms of the order $O\left(\frac{1}{n}\right)$ in Equation

[69] Note that the RHS of Equation (80) is Lipschitz in $x_t(\cdot)$ and $z_t(\cdot,\cdot)$, as it is composed of either linear terms, or has derivatives that are products of (conditional) probabilities, which are all bounded. It then follows that Kurtz's Theorem can be applied to establish convergence in probability to the mean dynamic in Equation (81). This is also confirmed in numerical simulations.

(80) yields

$$
\begin{aligned}
\frac{dz_t(\mathbf{h},\mathbf{h}')}{dt} &= x_t(\mathbf{h})x_t(\mathbf{h}')\left[\rho\, g^{\eta}_{\theta,\tau}(\mathbf{h},\mathbf{h}';\bar{h}_t)\left(1-\frac{z_t(\mathbf{h},\mathbf{h}')}{x_t(\mathbf{h})x_t(\mathbf{h}')}\right)-\rho\,\frac{z_t(\mathbf{h},\mathbf{h}')}{x_t(\mathbf{h})x_t(\mathbf{h}')}\right] \\
&+ \frac{z_t(\mathbf{h},\mathbf{h}')}{x_t(\mathbf{h})}\frac{dx_t(\mathbf{h})}{dt} + \sum_{k=1}^{N}\mathbb{1}_{\{h_k=1\}}\left(\gamma+\alpha\langle\mathbf{h}-\mathbf{e}_k,\mathbf{u}\rangle\right)x_t(\mathbf{h}-\mathbf{e}_k)\left(\frac{z_t(\mathbf{h}-\mathbf{e}_k,\mathbf{h}')}{x_t(\mathbf{h}-\mathbf{e}_k)}-\frac{z_t(\mathbf{h},\mathbf{h}')}{x_t(\mathbf{h})}\right) \\
&+ \lambda\sum_{k=1}^{N}\mathbb{1}_{\{h_k=0\}}x_t(\mathbf{h}+\mathbf{e}_k)\left(\frac{z_t(\mathbf{h}+\mathbf{e}_k,\mathbf{h}')}{x_t(\mathbf{h}+\mathbf{e}_k)}-\frac{z_t(\mathbf{h},\mathbf{h}')}{x_t(\mathbf{h})}\right) \\
&+ n\beta\sum_{k=1}^{N}\mathbb{1}_{\{h_k=1\}}\left(\frac{z_t(\mathbf{h}-\mathbf{e}_k,\mathbf{h}')}{x_t(\mathbf{h}-\mathbf{e}_k)}-\frac{z_t(\mathbf{h},\mathbf{h}')}{x_t(\mathbf{h})}\right)\sum_{\mathbf{h}''\in\mathcal{H}^N:h''_k=1} z_t(\mathbf{h}-\mathbf{e}_k,\mathbf{h}'') \\
&+ \frac{z_t(\mathbf{h},\mathbf{h}')}{x_t(\mathbf{h}')}\frac{dx_t(\mathbf{h}')}{dt} + \sum_{k=1}^{N}\mathbb{1}_{\{h'_k=1\}}\left(\gamma+\alpha\langle\mathbf{h}'-\mathbf{e}_k,\mathbf{u}\rangle\right)x_t(\mathbf{h}'-\mathbf{e}_k)\left(\frac{z_t(\mathbf{h},\mathbf{h}'-\mathbf{e}_k)}{x_t(\mathbf{h}'-\mathbf{e}_k)}-\frac{z_t(\mathbf{h},\mathbf{h}')}{x_t(\mathbf{h}')}\right) \\
&+ \lambda\sum_{k=1}^{N}\mathbb{1}_{\{h'_k=0\}}x_t(\mathbf{h}'+\mathbf{e}_k)\left(\frac{z_t(\mathbf{h},\mathbf{h}'+\mathbf{e}_k)}{x_t(\mathbf{h}'+\mathbf{e}_k)}-\frac{z_t(\mathbf{h},\mathbf{h}')}{x_t(\mathbf{h}')}\right) \\
&+ n\beta\sum_{k=1}^{N}\mathbb{1}_{\{h'_k=1\}}\left(\frac{z_t(\mathbf{h}'-\mathbf{e}_k,\mathbf{h})}{x_t(\mathbf{h}'-\mathbf{e}_k)}-\frac{z_t(\mathbf{h},\mathbf{h}')}{x_t(\mathbf{h}')}\right)\sum_{\mathbf{h}''\in\mathcal{H}^N:h''_k=1} z_t(\mathbf{h}'-\mathbf{e}_k,\mathbf{h}'') \\
&- \xi x_t(\mathbf{h})\frac{z_t(\mathbf{h},\mathbf{h}')}{x_t(\mathbf{h})} - \xi x_t(\mathbf{h}')\frac{z_t(\mathbf{h},\mathbf{h}')}{x_t(\mathbf{h}')}. \qquad (81)
\end{aligned}
$$

Equations (29) and (81) now represent a closed system of ODEs for the variables $x_t(\mathbf{h})$ and $z_t(\mathbf{h},\mathbf{h}')$. We next introduce the variable

$$
y_t(\mathbf{h},\mathbf{h}') \equiv \frac{z_t(\mathbf{h},\mathbf{h}')}{x_t(\mathbf{h})x_t(\mathbf{h}')},
$$

for which we have that

$$
\frac{dy_t(\mathbf{h},\mathbf{h}')}{dt} = \frac{1}{x_t(\mathbf{h})x_t(\mathbf{h}')}\frac{dz_t(\mathbf{h},\mathbf{h}')}{dt} - \frac{z_t(\mathbf{h},\mathbf{h}')}{x_t(\mathbf{h})x_t(\mathbf{h}')}\left(\frac{1}{x_t(\mathbf{h})}\frac{dx_t(\mathbf{h})}{dt}+\frac{1}{x_t(\mathbf{h}')}\frac{dx_t(\mathbf{h}')}{dt}\right).
$$

Using Equation (81) we then get

$$
\begin{aligned}
\frac{dy_t(\mathbf{h},\mathbf{h}')}{dt} &= \rho g^{\eta}_{\theta,\tau}(\mathbf{h},\mathbf{h}';\bar{h}_t)\left(1-y_t(\mathbf{h},\mathbf{h}')\right)-\rho\, y_t(\mathbf{h},\mathbf{h}') \\
&+ \sum_{k=1}^{N}\mathbb{1}_{\{h_k=1\}}\left(\gamma+\alpha\langle\mathbf{h}-\mathbf{e}_k,\mathbf{u}\rangle\right)\frac{x_t(\mathbf{h}-\mathbf{e}_k)}{x_t(\mathbf{h})}\left(y_t(\mathbf{h}-\mathbf{e}_k,\mathbf{h}')-y_t(\mathbf{h},\mathbf{h}')\right) \\
&+ \lambda\sum_{k=1}^{N}\mathbb{1}_{\{h_k=0\}}\frac{x_t(\mathbf{h}+\mathbf{e}_k)}{x_t(\mathbf{h})}\left(y_t(\mathbf{h}+\mathbf{e}_k,\mathbf{h}')-y_t(\mathbf{h},\mathbf{h}')\right)
\end{aligned}
$$

$$
\begin{aligned}
&+ n\beta \sum_{k=1}^{N} \mathbb{1}_{\{h_k=1\}} \left(y_t(\mathbf{h}-\mathbf{e}_k,\mathbf{h}') - y_t(\mathbf{h},\mathbf{h}')\right) \sum_{\mathbf{h}''\in\mathcal{H}^N : h''_k=1} y_t(\mathbf{h}-\mathbf{e}_k,\mathbf{h}'') \frac{x_t(\mathbf{h}-\mathbf{e}_k)x_t(\mathbf{h}'')}{x_t(\mathbf{h})} \\
&+ \sum_{k=1}^{N} \mathbb{1}_{\{h'_k=1\}} \left(\gamma + \alpha\langle \mathbf{h}'-\mathbf{e}_k, \mathbf{u}\rangle\right) \frac{x_t(\mathbf{h}'-\mathbf{e}_k)}{x_t(\mathbf{h}')} \left(y_t(\mathbf{h},\mathbf{h}'-\mathbf{e}_k) - y_t(\mathbf{h},\mathbf{h}')\right) \\
&+ \lambda \sum_{k=1}^{N} \mathbb{1}_{\{h'_k=0\}} \frac{x_t(\mathbf{h}'+\mathbf{e}_k)}{x_t(\mathbf{h}')} \left(y_t(\mathbf{h},\mathbf{h}'+\mathbf{e}_k) - y_t(\mathbf{h},\mathbf{h}')\right) \\
&+ n\beta \sum_{k=1}^{N} \mathbb{1}_{\{h'_k=1\}} \left(y_t(\mathbf{h}'-\mathbf{e}_k,\mathbf{h}) - y_t(\mathbf{h},\mathbf{h}')\right) \sum_{\mathbf{h}''\in\mathcal{H}^N : h''_k=1} y_t(\mathbf{h}'-\mathbf{e}_k,\mathbf{h}'') \frac{x_t(\mathbf{h}'-\mathbf{e}_k)x_t(\mathbf{h}'')}{x_t(\mathbf{h}')} \\
&- 2\xi x_t(\mathbf{h})x_t(\mathbf{h}')y_t(\mathbf{h},\mathbf{h}'). \qquad (82)
\end{aligned}
$$

□

The next proposition characterizes the dynamics of the technology stocks and the number of links between firms with different technology stocks.

Proposition 8. *Consider the parameters as in Theorem 1. Let the fraction of firms with a stock of knowledge of s be denoted by $\bar{x}_t(s)$ and let the probability of a link between a firm with knowledge stock s and a firm with s' be $\bar{y}_t(s,s')$ for any $0 \le s, s' \le N$ defined as in Equations (13) and (14). Then, under the pair approximation of Equation (12), $\bar{x}_t(s)$ converges in probability to the solution of Equation (36), and $\bar{y}_t(s,s')$ converges in probability to the solution of the ODE*

$$
\begin{aligned}
\frac{d\bar{y}_t(s,s')}{dt} &= \rho\bar{g}^{\eta}_{\theta,\tau}(s,s';\bar{h}_t) - \rho(1+\bar{g}^{\eta}_{\theta,\tau}(s,s';\bar{h}_t))\bar{y}_t(s,s') - 2\xi\bar{x}_t(s)\bar{x}_t(s')\bar{y}_t(s,s') \\
&+ \frac{\bar{x}_t(s-1)}{\bar{x}_t(s)} \left(\bar{y}_t(s-1,s') - \bar{y}_t(s,s')\right) s \left[(\gamma+\alpha(s-1)) + \beta \sum_{s''=1}^{N} \binom{N}{s''-1} \bar{y}_t(s-1,s'')\bar{x}_t(s'')\right] \\
&+ \frac{\bar{x}_t(s'-1)}{\bar{x}_t(s')} \left(\bar{y}_t(s'-1,s) - \bar{y}_t(s',s)\right) s' \left[\left(\gamma+\alpha(s'-1)\right) + \beta \sum_{s''=1}^{N} \binom{N}{s''-1} \bar{y}_t(s'-1,s'')\bar{x}_t(s'')\right] \\
&+ \lambda\frac{\bar{x}_t(s+1)}{\bar{x}_t(s)}(N-s)\left(\bar{y}_t(s+1,s') - \bar{y}_t(s,s')\right) + \lambda\frac{\bar{x}_t(s'+1)}{\bar{x}_t(s')}(N-s')\left(\bar{y}_t(s'+1,s) - \bar{y}_t(s',s)\right). \qquad (83)
\end{aligned}
$$

Proof of Proposition 8. Summation over all $\mathbf{h} \in \mathcal{H}^N$ with the property that $|\mathsf{S}(\mathbf{h})| = s$ and

$\mathbf{h}' \in \mathcal{H}^N$ with $|\,\mathsf{S}(\mathbf{h}')| = s'$ and inserting the definition in Equation (42) into Equation (77) gives

$$
\begin{aligned}
\frac{d\bar{y}_t(s,s')}{dt} &= \rho \bar{g}^{\eta}_{\theta,\tau}(s,s';\bar{h}_t) - \rho(1+\bar{g}^{\eta}_{\theta,\tau}(s,s';\bar{h}_t))\bar{y}_t(s,s') \\
&+ \frac{1}{\binom{N}{s}}\frac{1}{\binom{N}{s'}} \sum_{\substack{\mathbf{h}\in\mathcal{H}^N:|\,\mathsf{S}(\mathbf{h})|=s \\ \mathbf{h}'\in\mathcal{H}^N:|\,\mathsf{S}(\mathbf{h})|=s'}} \sum_{k=1}^{N} \mathbb{1}_{\{h_k=1\}} \left(\gamma + \alpha\langle \mathbf{h}-\mathbf{e}_k, \mathbf{u}\rangle\right) \frac{x_t(\mathbf{h}-\mathbf{e}_k)}{x_t(\mathbf{h})} \left(y_t(\mathbf{h}-\mathbf{e}_k,\mathbf{h}') - y_t(\mathbf{h},\mathbf{h}')\right) \\
&+ \frac{1}{\binom{N}{s}}\frac{1}{\binom{N}{s'}} \sum_{\substack{\mathbf{h}\in\mathcal{H}^N:|\,\mathsf{S}(\mathbf{h})|=s \\ \mathbf{h}'\in\mathcal{H}^N:|\,\mathsf{S}(\mathbf{h})|=s'}} \sum_{k=1}^{N} \mathbb{1}_{\{h'_k=1\}} \left(\gamma + \alpha\langle \mathbf{h}'-\mathbf{e}_k, \mathbf{u}\rangle\right) \frac{x_t(\mathbf{h}'-\mathbf{e}_k)}{x_t(\mathbf{h}')} \left(y_t(\mathbf{h},\mathbf{h}'-\mathbf{e}_k) - y_t(\mathbf{h},\mathbf{h}')\right) \\
&+ \lambda\frac{1}{\binom{N}{s}}\frac{1}{\binom{N}{s'}} \sum_{\substack{\mathbf{h}\in\mathcal{H}^N:|\,\mathsf{S}(\mathbf{h})|=s \\ \mathbf{h}'\in\mathcal{H}^N:|\,\mathsf{S}(\mathbf{h})|=s'}} \sum_{k=1}^{N} \mathbb{1}_{\{h_k=0\}} \frac{x_t(\mathbf{h}+\mathbf{e}_k)}{x_t(\mathbf{h})} \left(y_t(\mathbf{h}+\mathbf{e}_k,\mathbf{h}') - y_t(\mathbf{h},\mathbf{h}')\right) \\
&+ \lambda\frac{1}{\binom{N}{s}}\frac{1}{\binom{N}{s'}} \sum_{\substack{\mathbf{h}\in\mathcal{H}^N:|\,\mathsf{S}(\mathbf{h})|=s \\ \mathbf{h}'\in\mathcal{H}^N:|\,\mathsf{S}(\mathbf{h})|=s'}} \sum_{k=1}^{N} \mathbb{1}_{\{h'_k=0\}} \frac{x_t(\mathbf{h}'+\mathbf{e}_k)}{x_t(\mathbf{h}')} \left(y_t(\mathbf{h},\mathbf{h}'+\mathbf{e}_k) - y_t(\mathbf{h},\mathbf{h}')\right) \\
&+ n\beta\frac{1}{\binom{N}{s}}\frac{1}{\binom{N}{s'}} \sum_{\substack{\mathbf{h}\in\mathcal{H}^N:|\,\mathsf{S}(\mathbf{h})|=s \\ \mathbf{h}'\in\mathcal{H}^N:|\,\mathsf{S}(\mathbf{h})|=s'}} \sum_{k=1}^{N} \mathbb{1}_{\{h_k=1\}} \left(y_t(\mathbf{h}-\mathbf{e}_k,\mathbf{h}') - y_t(\mathbf{h},\mathbf{h}')\right) \sum_{\mathbf{h}''\in\mathcal{H}^N:h''_k=1} y_t(\mathbf{h}-\mathbf{e}_k,\mathbf{h}'')\frac{x_t(\mathbf{h}-\mathbf{e}_k)x_t(\mathbf{h}'')}{x_t(\mathbf{h})} \\
&+ n\beta\frac{1}{\binom{N}{s}}\frac{1}{\binom{N}{s'}} \sum_{\substack{\mathbf{h}\in\mathcal{H}^N:|\,\mathsf{S}(\mathbf{h})|=s \\ \mathbf{h}'\in\mathcal{H}^N:|\,\mathsf{S}(\mathbf{h})|=s'}} \sum_{k=1}^{N} \mathbb{1}_{\{h'_k=1\}} \left(y_t(\mathbf{h}'-\mathbf{e}_k,\mathbf{h}) - y_t(\mathbf{h},\mathbf{h}')\right) \sum_{\mathbf{h}''\in\mathcal{H}^N:h''_k=1} y_t(\mathbf{h}'-\mathbf{e}_k,\mathbf{h}'')\frac{x_t(\mathbf{h}'-\mathbf{e}_k)x_t(\mathbf{h}'')}{x_t(\mathbf{h}')} \\
&- 2\xi\frac{1}{\binom{N}{s}}\frac{1}{\binom{N}{s'}} \sum_{\substack{\mathbf{h}\in\mathcal{H}^N:|\,\mathsf{S}(\mathbf{h})|=s \\ \mathbf{h}'\in\mathcal{H}^N:|\,\mathsf{S}(\mathbf{h})|=s'}} x_t(\mathbf{h})x_t(\mathbf{h}')y_t(\mathbf{h},\mathbf{h}'). \qquad (84)
\end{aligned}
$$

Note that $\langle \mathbf{h}-\mathbf{e}_k, \mathbf{u}\rangle = |\,\mathsf{S}(\mathbf{h}-\mathbf{e}_k)|$ and that

$$
\begin{aligned}
&\sum_{\mathbf{h}'\in\mathcal{H}^N:|\,\mathsf{S}(\mathbf{h})|=s'} \sum_{\mathbf{h}\in\mathcal{H}^N:|\,\mathsf{S}(\mathbf{h})|=s} \sum_{k\in\mathsf{S}(\mathbf{h})} \left(\gamma+\alpha|\,\mathsf{S}(\mathbf{h}-\mathbf{e}_k)|\right) \frac{x_t(\mathbf{h}-\mathbf{e}_k)}{x_t(\mathbf{h})} y_t(\mathbf{h}-\mathbf{e}_k,\mathbf{h}') \\
&= \sum_{\mathbf{h}'\in\mathcal{H}^N:|\,\mathsf{S}(\mathbf{h})|=s'} \sum_{\mathbf{h}\in\mathcal{H}^N:|\,\mathsf{S}(\mathbf{h})|=s-1} \sum_{k\in\mathsf{S}(\mathbf{h}^c)} \left(\gamma+\alpha|\,\mathsf{S}(\mathbf{h})|\right) \frac{x_t(\mathbf{h})}{x_t(\mathbf{h}+\mathbf{e}_k)} y_t(\mathbf{h},\mathbf{h}') \\
&= \sum_{\mathbf{h}'\in\mathcal{H}^N:|\,\mathsf{S}(\mathbf{h})|=s'} \sum_{\mathbf{h}\in\mathcal{H}^N:|\,\mathsf{S}(\mathbf{h})|=s-1} \left(\gamma+\alpha|\,\mathsf{S}(\mathbf{h})|\right) x_t(\mathbf{h})y_t(\mathbf{h},\mathbf{h}') \sum_{k\in\mathsf{S}(\mathbf{h}^c)} \frac{1}{x_t(\mathbf{h}+\mathbf{e}_k)} \\
&= \sum_{\mathbf{h}'\in\mathcal{H}^N:|\,\mathsf{S}(\mathbf{h})|=s'} \sum_{\mathbf{h}\in\mathcal{H}^N:|\,\mathsf{S}(\mathbf{h})|=s-1} \left(\gamma+\alpha(s-1)\right) \frac{\bar{x}_t(s-1)\bar{y}_t(s-1,s')}{\bar{x}_t(s)} \sum_{k\in\mathsf{S}(\mathbf{h}^c)} \\
&= \sum_{\mathbf{h}'\in\mathcal{H}^N:|\,\mathsf{S}(\mathbf{h})|=s'} \sum_{\mathbf{h}\in\mathcal{H}^N:|\,\mathsf{S}(\mathbf{h})|=s-1} \left(\gamma+\alpha(s-1)\right) \frac{\bar{x}_t(s-1)\bar{y}_t(s-1,s')}{\bar{x}_t(s)} (N-s+1) \\
&= \left(\gamma+\alpha(s-1)\right) \frac{\bar{x}_t(s-1)\bar{y}_t(s-1,s')}{\bar{x}_t(s)} (N-s+1) \binom{N}{s-1}\binom{N}{s'}.
\end{aligned}
$$

It then follows that

$$\frac{1}{\binom{N}{s}}\frac{1}{\binom{N}{s'}}\sum_{\mathbf{h}'\in\mathcal{H}^N:|\,\mathsf{S}(\mathbf{h})|=s'}\sum_{\mathbf{h}\in\mathcal{H}^N:|\,\mathsf{S}(\mathbf{h})|=s}\sum_{k\in\mathsf{S}(\mathbf{h})}(\gamma+\alpha|\,\mathsf{S}(\mathbf{h}-\mathbf{e}_k)|)\,\frac{x_t(\mathbf{h}-\mathbf{e}_k)}{x_t(\mathbf{h})}y_t(\mathbf{h}-\mathbf{e}_k,\mathbf{h}')$$
$$=s\,(\gamma+\alpha(s-1))\,\frac{\bar{x}_t(s-1)\bar{y}_t(s-1,s')}{\bar{x}_t(s)}.$$

Similarly, we have that

$$\begin{aligned}
&\sum_{\mathbf{h}'\in\mathcal{H}^N:|\,\mathsf{S}(\mathbf{h})|=s'}\sum_{\mathbf{h}\in\mathcal{H}^N:|\,\mathsf{S}(\mathbf{h})|=s}\sum_{k\in\mathsf{S}(\mathbf{h})}(\gamma+\alpha|\,\mathsf{S}(\mathbf{h}-\mathbf{e}_k)|)\,\frac{x_t(\mathbf{h}-\mathbf{e}_k)}{x_t(\mathbf{h})}y_t(\mathbf{h},\mathbf{h}')\\
&=\sum_{\mathbf{h}'\in\mathcal{H}^N:|\,\mathsf{S}(\mathbf{h})|=s'}\sum_{\mathbf{h}\in\mathcal{H}^N:|\,\mathsf{S}(\mathbf{h})|=s-1}\sum_{k\in\mathsf{S}(\mathbf{h}^c)}(\gamma+\alpha|\,\mathsf{S}(\mathbf{h})|)\,\frac{x_t(\mathbf{h})}{x_t(\mathbf{h}+\mathbf{e}_k)}y_t(\mathbf{h}+\mathbf{e}_k,\mathbf{h}')\\
&=\sum_{\mathbf{h}'\in\mathcal{H}^N:|\,\mathsf{S}(\mathbf{h})|=s'}\sum_{\mathbf{h}\in\mathcal{H}^N:|\,\mathsf{S}(\mathbf{h})|=s-1}(\gamma+\alpha|\,\mathsf{S}(\mathbf{h})|)\,x_t(\mathbf{h})\sum_{k\in\mathsf{S}(\mathbf{h}^c)}\frac{y_t(\mathbf{h},\mathbf{h}')}{x_t(\mathbf{h}+\mathbf{e}_k)}\\
&=\sum_{\mathbf{h}'\in\mathcal{H}^N:|\,\mathsf{S}(\mathbf{h})|=s'}\sum_{\mathbf{h}\in\mathcal{H}^N:|\,\mathsf{S}(\mathbf{h})|=s-1}(\gamma+\alpha(s-1))\,\frac{\bar{x}_t(s-1)\bar{y}_t(s,s')}{\bar{x}_t(s)}\sum_{k\in\mathsf{S}(\mathbf{h}^c)}\\
&=\sum_{\mathbf{h}'\in\mathcal{H}^N:|\,\mathsf{S}(\mathbf{h})|=s'}\sum_{\mathbf{h}\in\mathcal{H}^N:|\,\mathsf{S}(\mathbf{h})|=s-1}(\gamma+\alpha(s-1))\,\frac{\bar{x}_t(s-1)\bar{y}_t(s,s')}{\bar{x}_t(s)}(N-s+1)\\
&=(\gamma+\alpha(s-1))\,\frac{\bar{x}_t(s-1)\bar{y}_t(s,s')}{\bar{x}_t(s)}(N-s+1)\binom{N}{s-1}\binom{N}{s'}.
\end{aligned}$$

Hence, we get

$$\frac{1}{\binom{N}{s}}\frac{1}{\binom{N}{s'}}\sum_{\mathbf{h}'\in\mathcal{H}^N:|\,\mathsf{S}(\mathbf{h})|=s'}\sum_{\mathbf{h}\in\mathcal{H}^N:|\,\mathsf{S}(\mathbf{h})|=s}\sum_{k\in\mathsf{S}(\mathbf{h})}(\gamma+\alpha|\,\mathsf{S}(\mathbf{h}-\mathbf{e}_k)|)\,\frac{x_t(\mathbf{h}-\mathbf{e}_k)}{x_t(\mathbf{h})}y_t(\mathbf{h},\mathbf{h}')$$
$$=s\,(\gamma+\alpha(s-1))\,\frac{\bar{x}_t(s-1)\bar{y}_t(s,s')}{\bar{x}_t(s)}.$$

Next, we have that

$$\begin{aligned}
&\sum_{\mathbf{h}'\in\mathcal{H}^N:|\,\mathsf{S}(\mathbf{h})|=s'}\sum_{\mathbf{h}\in\mathcal{H}^N:|\,\mathsf{S}(\mathbf{h})|=s}\sum_{k\in\mathsf{S}(\mathbf{h}^c)}\frac{x_t(\mathbf{h}+\mathbf{e}_k)}{x_t(\mathbf{h})}y_t(\mathbf{h}+\mathbf{e}_k,\mathbf{h}')\\
&=\sum_{\mathbf{h}'\in\mathcal{H}^N:|\,\mathsf{S}(\mathbf{h})|=s'}\sum_{\mathbf{h}\in\mathcal{H}^N:|\,\mathsf{S}(\mathbf{h})|=s+1}\sum_{k\in\mathsf{S}(\mathbf{h})}\frac{x_t(\mathbf{h})}{x_t(\mathbf{h}-\mathbf{e}_k)}y_t(\mathbf{h},\mathbf{h}')\\
&=\sum_{\mathbf{h}'\in\mathcal{H}^N:|\,\mathsf{S}(\mathbf{h})|=s'}\sum_{\mathbf{h}\in\mathcal{H}^N:|\,\mathsf{S}(\mathbf{h})|=s+1}x_t(\mathbf{h})y_t(\mathbf{h},\mathbf{h}')\sum_{k\in\mathsf{S}(\mathbf{h})}\frac{1}{x_t(\mathbf{h}-\mathbf{e}_k)}\\
&=\sum_{\mathbf{h}'\in\mathcal{H}^N:|\,\mathsf{S}(\mathbf{h})|=s'}\sum_{\mathbf{h}\in\mathcal{H}^N:|\,\mathsf{S}(\mathbf{h})|=s+1}\frac{\bar{x}_t(s+1)\bar{y}_t(s+1,s')}{\bar{x}_t(s)}\sum_{k\in\mathsf{S}(\mathbf{h})}\\
&=\sum_{\mathbf{h}'\in\mathcal{H}^N:|\,\mathsf{S}(\mathbf{h})|=s'}\sum_{\mathbf{h}\in\mathcal{H}^N:|\,\mathsf{S}(\mathbf{h})|=s+1}\frac{\bar{x}_t(s+1)\bar{y}_t(s+1,s')}{\bar{x}_t(s)}(s+1)\\
&=\frac{\bar{x}_t(s+1)\bar{y}_t(s+1,s')}{\bar{x}_t(s)}(s+1)\binom{N}{s+1}\binom{N}{s'}.
\end{aligned}$$

Thus we get

$$\frac{1}{\binom{N}{s}}\frac{1}{\binom{N}{s'}}\sum_{\mathbf{h}'\in\mathcal{H}^N:|\,\mathsf{S}(\mathbf{h})|=s'}\sum_{\mathbf{h}\in\mathcal{H}^N:|\,\mathsf{S}(\mathbf{h})|=s}\sum_{k\in\mathsf{S}(\mathbf{h}^c)}\frac{x_t(\mathbf{h}+\mathbf{e}_k)}{x_t(\mathbf{h})}y_t(\mathbf{h}+\mathbf{e}_k,\mathbf{h}')=\frac{\bar{x}_t(s+1)\bar{y}_t(s+1,s')}{\bar{x}_t(s)}(N-s)$$

Similarly, we have that

$$\begin{aligned}
&\sum_{\mathbf{h}'\in\mathcal{H}^N:|\,\mathsf{S}(\mathbf{h})|=s'}\sum_{\mathbf{h}\in\mathcal{H}^N:|\,\mathsf{S}(\mathbf{h})|=s}\sum_{k\in\mathsf{S}(\mathbf{h}^c)}\frac{x_t(\mathbf{h}+\mathbf{e}_k)}{x_t(\mathbf{h})}y_t(\mathbf{h},\mathbf{h}')\\
&=\sum_{\mathbf{h}'\in\mathcal{H}^N:|\,\mathsf{S}(\mathbf{h})|=s'}\sum_{\mathbf{h}\in\mathcal{H}^N:|\,\mathsf{S}(\mathbf{h})|=s+1}\sum_{k\in\mathsf{S}(\mathbf{h})}\frac{x_t(\mathbf{h})}{x_t(\mathbf{h}-\mathbf{e}_k)}y_t(\mathbf{h}-\mathbf{e}_k,\mathbf{h}')\\
&=\sum_{\mathbf{h}'\in\mathcal{H}^N:|\,\mathsf{S}(\mathbf{h})|=s'}\sum_{\mathbf{h}\in\mathcal{H}^N:|\,\mathsf{S}(\mathbf{h})|=s+1}x_t(\mathbf{h})\sum_{k\in\mathsf{S}(\mathbf{h})}\frac{y_t(\mathbf{h}-\mathbf{e}_k,\mathbf{h}')}{x_t(\mathbf{h}-\mathbf{e}_k)}\\
&=\sum_{\mathbf{h}'\in\mathcal{H}^N:|\,\mathsf{S}(\mathbf{h})|=s'}\sum_{\mathbf{h}\in\mathcal{H}^N:|\,\mathsf{S}(\mathbf{h})|=s+1}\frac{\bar{x}_t(s+1)\bar{y}_t(s,s')}{\bar{x}_t(s)}\sum_{k\in\mathsf{S}(\mathbf{h})}\\
&=\sum_{\mathbf{h}'\in\mathcal{H}^N:|\,\mathsf{S}(\mathbf{h})|=s'}\sum_{\mathbf{h}\in\mathcal{H}^N:|\,\mathsf{S}(\mathbf{h})|=s+1}\frac{\bar{x}_t(s+1)\bar{y}_t(s,s')}{\bar{x}_t(s)}(s+1)\\
&=\frac{\bar{x}_t(s+1)\bar{y}_t(s,s')}{\bar{x}_t(s)}(s+1)\binom{N}{s'}\binom{N}{s+1}.
\end{aligned}$$

We then get

$$\frac{1}{\binom{N}{s}}\frac{1}{\binom{N}{s'}}\sum_{\mathbf{h}'\in\mathcal{H}^N:|\,\mathsf{S}(\mathbf{h})|=s'}\sum_{\mathbf{h}\in\mathcal{H}^N:|\,\mathsf{S}(\mathbf{h})|=s}\sum_{k\in\mathsf{S}(\mathbf{h}^c)}\frac{x_t(\mathbf{h}+\mathbf{e}_k)}{x_t(\mathbf{h})}y_t(\mathbf{h},\mathbf{h}')=\frac{\bar{x}_t(s+1)\bar{y}_t(s,s')}{\bar{x}_t(s)}(N-s).$$

Moreover, we have that

$$\begin{aligned}
&\sum_{\mathbf{h}'\in\mathcal{H}^N:|\,\mathsf{S}(\mathbf{h})|=s'}\sum_{\mathbf{h}\in\mathcal{H}^N:|\,\mathsf{S}(\mathbf{h})|=s}\sum_{k\in\mathsf{S}(\mathbf{h})}\left(y_t(\mathbf{h}-\mathbf{e}_k,\mathbf{h}')-y_t(\mathbf{h},\mathbf{h}')\right)\frac{x_t(\mathbf{h}-\mathbf{e}_k)}{x_t(\mathbf{h})}\sum_{\mathbf{h}''\in\mathcal{H}^N:h''_k=1}y_t(\mathbf{h}-\mathbf{e}_k,\mathbf{h}'')x_t(\mathbf{h}'')\\
&=\sum_{\mathbf{h}'\in\mathcal{H}^N:|\,\mathsf{S}(\mathbf{h})|=s'}\sum_{\mathbf{h}\in\mathcal{H}^N:|\,\mathsf{S}(\mathbf{h})|=s-1}\sum_{k\in\mathsf{S}(\mathbf{h}^c)}\left(y_t(\mathbf{h},\mathbf{h}')-y_t(\mathbf{h}+\mathbf{e}_k,\mathbf{h}')\right)\frac{x_t(\mathbf{h})}{x_t(\mathbf{h}+\mathbf{e}_k)}\sum_{s''=1}^{N}\sum_{\mathbf{h}''\in\mathcal{H}^N:h''_k=1,|\,\mathsf{S}(\mathbf{h}'')|=s''}y_t(\mathbf{h},\mathbf{h}'')\\
&=\sum_{\mathbf{h}'\in\mathcal{H}^N:|\,\mathsf{S}(\mathbf{h})|=s'}\sum_{\mathbf{h}\in\mathcal{H}^N:|\,\mathsf{S}(\mathbf{h})|=s-1}\sum_{k\in\mathsf{S}(\mathbf{h}^c)}\left(\bar{y}_t(s-1,s')-\bar{y}_t(s,s')\right)\frac{\bar{x}_t(s-1)}{\bar{x}_t(s)}\sum_{s''=1}^{N}\binom{N}{s''-1}\bar{y}_t(s-1,s'')\bar{x}_t(s'')\\
&=\sum_{\mathbf{h}'\in\mathcal{H}^N:|\,\mathsf{S}(\mathbf{h})|=s'}\sum_{\mathbf{h}\in\mathcal{H}^N:|\,\mathsf{S}(\mathbf{h})|=s-1}(N-s+1)\left(\bar{y}_t(s-1,s')-\bar{y}_t(s,s')\right)\frac{\bar{x}_t(s-1)}{\bar{x}_t(s)}\sum_{s''=1}^{N}\binom{N}{s''-1}\bar{y}_t(s-1,s'')\bar{x}_t(s'')\\
&=\binom{N}{s-1}\binom{N}{s'}(N-s+1)\left(\bar{y}_t(s-1,s')-\bar{y}_t(s,s')\right)\frac{\bar{x}_t(s-1)}{\bar{x}_t(s)}\sum_{s''=1}^{N}\binom{N}{s''-1}\bar{y}_t(s-1,s'')\bar{x}_t(s'').
\end{aligned}$$

We then get

$$\begin{aligned}
&\frac{1}{\binom{N}{s}}\frac{1}{\binom{N}{s'}}\sum_{\mathbf{h}'\in\mathcal{H}^N:|\,\mathsf{S}(\mathbf{h})|=s'}\sum_{\mathbf{h}\in\mathcal{H}^N:|\,\mathsf{S}(\mathbf{h})|=s}\sum_{k\in\mathsf{S}(\mathbf{h})}\left(y_t(\mathbf{h}-\mathbf{e}_k,\mathbf{h}')-y_t(\mathbf{h},\mathbf{h}')\right)\frac{x_t(\mathbf{h}-\mathbf{e}_k)}{x_t(\mathbf{h})}\sum_{\mathbf{h}''\in\mathcal{H}^N:h''_k=1}y_t(\mathbf{h}-\mathbf{e}_k,\mathbf{h}'')x_t(\mathbf{h}'')\\
&=s\left(\bar{y}_t(s-1,s')-\bar{y}_t(s,s')\right)\frac{\bar{x}_t(s-1)}{\bar{x}_t(s)}\sum_{s''=1}^{N}\binom{N}{s''-1}\bar{y}_t(s-1,s'')\bar{x}_t(s'').
\end{aligned}$$

Finally, we have that

$$\frac{1}{\binom{N}{s}}\frac{1}{\binom{N}{s'}}\sum_{\substack{\mathbf{h}\in\mathcal{H}^N:|\mathsf{S}(\mathbf{h})|=s\\ \mathbf{h}'\in\mathcal{H}^N:|\mathsf{S}(\mathbf{h})|=s'}} x_t(\mathbf{h})x_t(\mathbf{h}')y_t(\mathbf{h},\mathbf{h}') = \bar{x}_t(s)\bar{x}_t(s')\bar{y}_t(s,s').$$

Collecting the above terms in Equation (84) delivers

$$\begin{aligned}
\frac{d\bar{y}_t(s,s')}{dt} &= \rho\bar{g}^{\eta}_{\theta,\tau}(s,s';\bar{h}_t) - \rho(1+\bar{g}^{\eta}_{\theta,\tau}(s,s';\bar{h}_t))\bar{y}_t(s,s') - 2\xi\bar{x}_t(s)\bar{x}_t(s')\bar{y}_t(s,s')\\
&+ \frac{\bar{x}_t(s-1)}{\bar{x}_t(s)}\left(\bar{y}_t(s-1,s')-\bar{y}_t(s,s')\right)s\left[(\gamma+\alpha(s-1))+\beta\sum_{s''=1}^{N}\binom{N}{s''-1}\bar{y}_t(s-1,s'')\bar{x}_t(s'')\right]\\
&+ \frac{\bar{x}_t(s'-1)}{\bar{x}_t(s')}\left(\bar{y}_t(s'-1,s)-\bar{y}_t(s',s)\right)s'\left[\left(\gamma+\alpha(s'-1)\right)+\beta\sum_{s''=1}^{N}\binom{N}{s''-1}\bar{y}_t(s'-1,s'')\bar{x}_t(s'')\right]\\
&+ \lambda\frac{\bar{x}_t(s+1)}{\bar{x}_t(s)}(N-s)\left(\bar{y}_t(s+1,s')-\bar{y}_t(s,s')\right)+\lambda\frac{\bar{x}_t(s'+1)}{\bar{x}_t(s')}(N-s')\left(\bar{y}_t(s'+1,s)-\bar{y}_t(s',s)\right).
\end{aligned} \tag{85}$$

Equations (43) and (85) provide a complete system of ODEs to describe the time evolution of $\bar{x}_t(s)$ and $\bar{y}_t(s,s')$. □

J.1. A Single Technology ($N = 1$)

In the case of $N = 1$ where $s \in \{0,1\}$ we obtain from Equation (43) that

$$\begin{aligned}
\frac{d\bar{x}_t(0)}{dt} &= -\gamma\bar{x}_t(0)+\lambda\bar{x}_t(1)-n\beta\bar{y}_t(0,1)\bar{x}_t(0)\bar{x}_t(1)-\xi\bar{x}_t(0)+\chi f^{\vartheta}_{\theta,\kappa}(\bar{x}_t(1))(1-\bar{x}_t(0)-\bar{x}_t(1))\\
\frac{d\bar{x}_t(1)}{dt} &= \gamma\bar{x}_t(0)-\lambda\bar{x}_t(1)+n\beta\bar{y}_t(0,1)\bar{x}_t(0)\bar{x}_t(1)-\xi\bar{x}_t(1)\\
\frac{d\bar{y}_t(0,1)}{dt} &= \rho g^{\eta}_{\theta,\tau}(0,1;\bar{x}_t(1))-\rho(1+g^{\eta}_{\theta,\tau}(0,1;\bar{x}_t(1)))\bar{y}_t(0,1)-2\xi\bar{x}_t(0)\bar{x}_t(1)\bar{y}_t(0,1)\\
&+\frac{x_t(0)}{x_t(1)}(y_t(0,0)-y_t(0,1))\left(\gamma+n\beta y_t(0,1)x_t(1)\right)+\lambda\frac{x_t(1)}{x_t(0)}(y_t(1,1)-y_t(0,1))\\
\frac{d\bar{y}_t(0,0)}{dt} &= \rho g^{\eta}_{\theta,\tau}(0,0;\bar{x}_t(1))-\rho(1+g^{\eta}_{\theta,\tau}(0,0;\bar{x}_t(1)))\bar{y}_t(0,1)-2\xi\bar{x}_t(0)^2\bar{y}_t(0,0)+2\lambda\frac{x_t(1)}{x_t(0)}(y_t(1,0)-y_t(0,0))\\
\frac{d\bar{y}_t(1,1)}{dt} &= \rho g^{\eta}_{\theta,\tau}(1,1;\bar{x}_t(1))-\rho(1+g^{\eta}_{\theta,\tau}(1,1;\bar{x}_t(1)))\bar{y}_t(0,1)-2\xi\bar{x}_t(1)^2\bar{y}_t(1,1)\\
&+2\frac{x_t(0)}{x_t(1)}(y_t(0,1)-y_t(1,1))\left(\gamma+n\beta y_t(0,1)x_t(1)\right),
\end{aligned} \tag{86}$$

where $\bar{h}_t = \bar{x}_t(1)$, and from Equation (8) we know that in the competitive case with $\theta = 0$

$$\begin{aligned}
g^{\eta}_{0,\tau}(0,1;\bar{x}_t(1)) &= \frac{e^{\eta\left(\frac{2\beta\tau b}{1+b\bar{x}_t(1)}-c\right)}}{1+e^{\eta\left(\frac{2\beta\tau b}{1+b\bar{x}_t(1)}-c\right)}}\frac{e^{-\eta c}}{1+e^{-\eta c}},\\
g^{\eta}_{0,\tau}(0,0;\bar{x}_t(1)) &= \frac{e^{-\eta c}}{1+e^{-\eta c}}\frac{e^{-\eta c}}{1+e^{-\eta c}},\\
g^{\eta}_{0,\tau}(1,1;\bar{x}_t(1)) &= \frac{e^{-\eta c}}{1+e^{-\eta c}}\frac{e^{-\eta c}}{1+e^{-\eta c}},
\end{aligned}$$

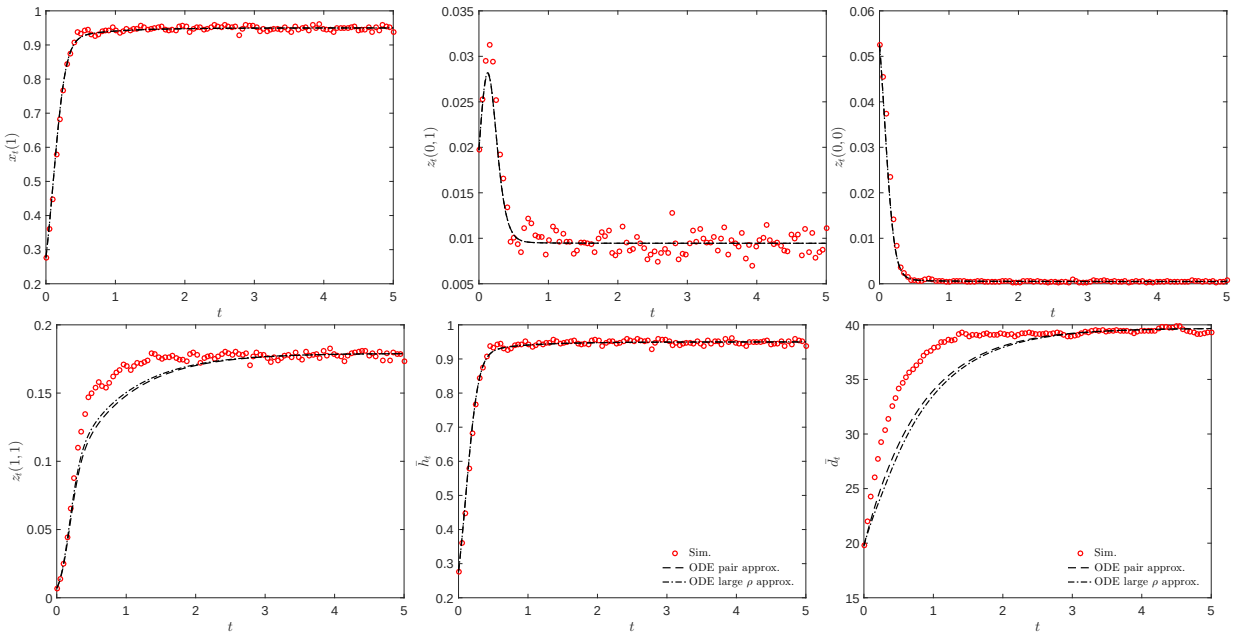


Figure J.1: Comparison of the agent based simulations for $N = 1$ and $n = 200$ with the numerical solutions of the ODEs using the large ρ approximation and the pair approximation for $\rho = 1$.

while from Equation (9) we have for the competitive case with $\theta = 0$ that

$$f^{\vartheta}_{0,\kappa}(\bar{x}_t(1)) = \frac{e^{\vartheta\left(\frac{1}{1+b\bar{x}_t(1)}-\kappa\right)}}{1+e^{\vartheta\left(\frac{1}{1+b\bar{x}_t(1)}-\kappa\right)}}.$$

Note that $\lim_{\eta\to 0}\bar{g}(0,0) = \lim_{\eta\to 0}\bar{g}(1,1) = \frac{1}{4}$ and $\lim_{\eta\to\infty}\bar{g}(0,0) = \lim_{\eta\to\infty}\bar{g}(1,1) = 0$. Further note that $\bar{x}_t(1) = 1 - \bar{x}_t(0)$. Observe that in the case of $N = 1$ the parameter α does not affect the dynamics. Also note that $\lim_{\eta\to\infty}\bar{g}(0,1) = 0$ and $\lim_{\eta\to 0}\bar{g}(0,1) = \frac{1}{4}$.

An example of a numerical simulation of the stochastic process introduced in in Sections 3.2, 3.3 and 3.4 using the "next reaction method" for simulating a continuous time Markov chain [cf. Anderson, 2012; Gibson and Bruck, 2000], and the solution of the ODEs in Equation (86) superimposed is shown in Figure J.1.

J.2. Two Technologies ($N = 2$)

When $N = 2$ with $s \in \{0,1,2\}$ we obtain from Equation (43)

$$\begin{aligned}
\frac{d\bar{x}_t(0)}{dt} &= 2\left(\lambda\bar{x}_t(1) - \gamma\bar{x}_t(0) - \beta(\bar{y}_t(0,1)\bar{x}_t(0)\bar{x}_t(1) + \bar{y}_t(0,2)\bar{x}_t(0)\bar{x}_t(2))\right) - \xi\bar{x}_t(0) \\
&\quad + \chi f^{\vartheta}_{\theta,\kappa}(\bar{h}_t)\left(1 - \sum_{s'=0}^{2}\binom{2}{s'}\bar{x}_t(s')\right), \\
\frac{d\bar{x}_t(1)}{dt} &= \gamma\bar{x}_t(0) + \lambda\bar{x}_t(2) - (\lambda+\gamma+\alpha)\bar{x}_t(1) + \beta\left(\bar{y}_t(0,1)\bar{x}_t(0)\bar{x}_t(1) - \bar{y}_t(1,1)\bar{x}_t(1)^2\right. \\
&\quad \left.+\bar{y}_t(0,2)\bar{x}_t(0)\bar{x}_t(2) - \bar{y}_t(1,2)\bar{x}_t(1)\bar{x}_t(2)\right) - \xi\bar{x}_t(1), \\
\frac{d\bar{x}_t(2)}{dt} &= 2\left((\gamma+\alpha)\bar{x}_t(1) - \lambda\bar{x}_t(2) + \beta\left(\bar{y}_t(1,1)\bar{x}_t(1)^2 + \bar{y}_t(1,2)\bar{x}_t(1)\bar{x}_t(2)\right)\right) - \xi\bar{x}_t(2),
\end{aligned} \tag{87}$$

where the average stock of knowledge is given by $\bar{h}_t(\mathbf{x}) = 2(\bar{x}_t(1) + \bar{x}_t(2))$. From Equation (85) we obtain

$$\frac{d\bar{y}_t(0,1)}{dt} = \frac{1}{2}\rho\tilde{g}^{\eta}_{\theta,\tau}(0,1;\bar{h}_t) - \rho\left(1 + \frac{1}{2}\tilde{g}^{\eta}_{\theta,\tau}(0,1;\bar{h}_t)\right)\bar{y}_t(0,1) - 2\xi\bar{x}_t(0)\bar{x}_t(1)\bar{y}_t(0,1)$$
$$+ \frac{\bar{x}_t(0)}{\bar{x}_t(1)}\left(\bar{y}_t(0,0) - \bar{y}_t(1,0)\right)\left[\gamma + \beta\left(\bar{y}_t(0,1)\bar{x}_t(1) + 2\bar{y}_t(0,2)\bar{x}_t(2)\right)\right]$$
$$+ \lambda\frac{\bar{x}_t(1)}{\bar{x}_t(0)}2\left(\bar{y}_t(1,1) - \bar{y}_t(0,1)\right) + \lambda\frac{\bar{x}_t(2)}{\bar{x}_t(1)}\left(\bar{y}_t(2,0) - \bar{y}_t(1,0)\right),$$

$$\frac{d\bar{y}_t(0,2)}{dt} = \rho\tilde{g}^{\eta}_{\theta,\tau}(0,2;\bar{h}_t) - \rho\left(1 + \tilde{g}^{\eta}_{\theta,\tau}(0,2;\bar{h}_t)\right)\bar{y}_t(0,2) - 2\xi\bar{x}_t(0)\bar{x}_t(2)\bar{y}_t(0,2)$$
$$+ 2\frac{\bar{x}_t(1)}{\bar{x}_t(2)}\left(\bar{y}_t(1,0) - \bar{y}_t(2,0)\right)\left[\gamma + \alpha + \beta\left(\bar{y}_t(1,1)\bar{x}_t(1) + 2\bar{y}_t(1,2)\bar{x}_t(2)\right)\right]$$
$$+ 2\lambda\frac{\bar{x}_t(1)}{\bar{x}_t(0)}\left(\bar{y}_t(1,2) - \bar{y}_t(0,2)\right),$$

$$\frac{d\bar{y}_t(1,1)}{dt} = \frac{1}{4}\rho\tilde{g}^{\eta}_{\theta,\tau}(1,1;\bar{h}_t) - \rho\left(1 + \frac{1}{4}\tilde{g}^{\eta}_{\theta,\tau}(1,1;\bar{h}_t)\right)\bar{y}_t(1,1) - 2\xi\bar{x}_t(1)^2\bar{y}_t(1,1)$$
$$+ 2\frac{\bar{x}_t(0)}{\bar{x}_t(1)}\left(\bar{y}_t(0,1) - \bar{y}_t(1,1)\right)\left[\gamma + \beta\left(\bar{y}_t(0,1)\bar{x}_t(1) + 2\bar{y}_t(0,2)\bar{x}_t(2)\right)\right]$$
$$+ 2\lambda\frac{\bar{x}_t(2)}{\bar{x}_t(1)}\left(\bar{y}_t(2,1) - \bar{y}_t(1,1)\right),$$

$$\frac{d\bar{y}_t(1,2)}{dt} = \frac{1}{2}\rho\tilde{g}^{\eta}_{\theta,\tau}(1,2;\bar{h}_t) - \rho\left(1 + \frac{1}{2}\tilde{g}^{\eta}_{\theta,\tau}(1,2;\bar{h}_t)\right)\bar{y}_t(1,2) - 2\xi\bar{x}_t(1)\bar{x}_t(2)\bar{y}_t(1,2)$$
$$+ \frac{\bar{x}_t(0)}{\bar{x}_t(1)}\left(\bar{y}_t(0,2) - \bar{y}_t(1,2)\right)\left[\gamma + \beta\left(\bar{y}_t(0,1)\bar{x}_t(1) + 2\bar{y}_t(0,2)\bar{x}_t(2)\right)\right]$$
$$+ 2\frac{\bar{x}_t(1)}{\bar{x}_t(2)}\left(\bar{y}_t(1,1) - \bar{y}_t(2,1)\right)\left[\gamma + \alpha + \beta\left(\bar{y}_t(1,1)\bar{x}_t(1) + 2\bar{y}_t(1,2)\bar{x}_t(2)\right)\right]$$
$$+ \lambda\frac{\bar{x}_t(2)}{\bar{x}_t(1)}\left(\bar{y}_t(2,2) - \bar{y}_t(1,2)\right),$$

$$\frac{d\bar{y}_t(0,0)}{dt} = \rho\bar{g}^{\eta}_{\theta,\tau}(0,0';\bar{h}_t) - 2\xi\bar{x}_t(0)^2\bar{y}_t(0,0) - \rho(1 + \bar{g}^{\eta}_{\theta,\tau}(0,0';\bar{h}_t))\tilde{y}_t(0,0') + 4\lambda\frac{\bar{x}_t(1)}{\bar{x}_t(0)}\left(\bar{y}_t(1,0) - \bar{y}_t(0,0)\right),$$

$$\frac{d\bar{y}_t(2,2)}{dt} = \rho\bar{g}^{\eta}_{\theta,\tau}(2,2;\bar{h}_t) - \rho(1 + \bar{g}^{\eta}_{\theta,\tau}(2,2;\bar{h}_t))\tilde{y}_t(2,2) - 2\xi\bar{x}_t(2)^2\bar{y}_t(2,2)$$
$$+ 4\frac{\bar{x}_t(1)}{\bar{x}_t(2)}\left(\bar{y}_t(1,2) - \bar{y}_t(2,2)\right)\left[\gamma + \alpha + \beta\left(\bar{y}_t(1,1)\bar{x}_t(1) + 2\bar{y}_t(1,2)\bar{x}_t(2)\right)\right]. \tag{88}$$

From the definitions in Equations (8) and (15) we find that

$$\tilde{g}^{\eta}_{\theta,\tau}(0,1;\bar{h}_t) = \sum_{\substack{\mathbf{h}=(0,0)^\top \\ \mathbf{h}'\in\{(0,1)^\top,(1,0)^\top\}}} \frac{e^{\eta(\beta g_{\theta,\tau}(\mathbf{x})(1+b|\,\mathsf{S}(\mathbf{h})|)^{1-\theta}\langle \mathbf{h}^c,\mathbf{h}'\rangle - c)}}{1+e^{\eta(\beta g_{\theta,\tau}(\mathbf{x})(1+b|\,\mathsf{S}(\mathbf{h})|)^{1-\theta}\langle \mathbf{h}^c,\mathbf{h}'\rangle - c)}} \frac{e^{\eta(\beta g_{\theta,\tau}(\mathbf{x})(1+b|\,\mathsf{S}(\mathbf{h}')|)^{1-\theta}\langle \mathbf{h}'^c,\mathbf{h}\rangle - c)}}{1+e^{\eta(\beta g_{\theta,\tau}(\mathbf{x})(1+b|\,\mathsf{S}(\mathbf{h}')|)^{1-\theta}\langle \mathbf{h}'^c,\mathbf{h}\rangle - c)}}$$

$$= \begin{cases} 2\frac{e^{-\eta c}}{1+e^{-\eta c}}\frac{e^{\eta(\beta g_{1,\tau}-c)}}{1+e^{\eta(\beta g_{1,\tau}-c)}} = 2\frac{e^{-\eta c}}{1+e^{-\eta c}}\frac{e^{\eta(\beta b\tau - c)}}{1+e^{\eta(\beta b\tau - c)}} & \text{if } \theta = 1, \\ 2\frac{e^{-\eta c}}{1+e^{-\eta c}}\frac{e^{\eta(\beta g_{0,\tau}(\mathbf{x})-c)}}{1+e^{\eta(\beta g_{0,\tau}(\mathbf{x})-c)}} = 2\frac{e^{-\eta c}}{1+e^{-\eta c}}\frac{e^{\eta(\beta\frac{2\tau b}{1+bh_t(\mathbf{x})}-c)}}{1+e^{\eta(\beta\frac{2\tau b}{1+bh_t(\mathbf{x})}-c)}} & \text{if } \theta = 0, \end{cases}$$

$$\tilde{g}^{\eta}_{\theta,\tau}(0,2;\bar{h}_t) = \sum_{\substack{\mathbf{h}=(0,0)^\top \\ \mathbf{h}'=(1,1)^\top}} \frac{e^{\eta(\beta g_{\theta,\tau}(\mathbf{x})(1+b|\,\mathsf{S}(\mathbf{h})|)^{1-\theta}\langle \mathbf{h}^c,\mathbf{h}'\rangle - c)}}{1+e^{\eta(\beta g_{\theta,\tau}(\mathbf{x})(1+b|\,\mathsf{S}(\mathbf{h})|)^{1-\theta}\langle \mathbf{h}^c,\mathbf{h}'\rangle - c)}} \frac{e^{\eta(\beta g_{\theta,\tau}(\mathbf{x})(1+b|\,\mathsf{S}(\mathbf{h}')|)^{1-\theta}\langle \mathbf{h}'^c,\mathbf{h}\rangle - c)}}{1+e^{\eta(\beta g_{\theta,\tau}(\mathbf{x})(1+b|\,\mathsf{S}(\mathbf{h}')|)^{1-\theta}\langle \mathbf{h}'^c,\mathbf{h}\rangle - c)}}$$

$$= \begin{cases} \frac{e^{-\eta c}}{1+e^{-\eta c}}\frac{e^{\eta(2\beta g^{n}_{1,\tau}-c)}}{1+e^{\eta(2\beta g^{n}_{1,\tau}-c)}} = \frac{e^{-\eta c}}{1+e^{-\eta c}}\frac{e^{\eta(2\beta b\tau - c)}}{1+e^{\eta(2\beta b\tau - c)}} & \text{if } \theta = 1, \\ \frac{e^{-\eta c}}{1+e^{-\eta c}}\frac{e^{\eta(2\beta g_{0,\tau}(\mathbf{x})-c)}}{1+e^{\eta(2\beta g_{0,\tau}(\mathbf{x})-c)}} = \frac{e^{-\eta c}}{1+e^{-\eta c}}\frac{e^{\eta(2\beta\frac{2\tau b}{1+bh_t(\mathbf{x})}-c)}}{1+e^{\eta(2\beta\frac{2\tau b}{1+bh_t(\mathbf{x})}-c)}} & \text{if } \theta = 0, \end{cases}$$

$$\tilde{g}^{\eta}_{\theta,\tau}(1,1;\bar{h}_t) = \sum_{\substack{\mathbf{h}\in\{(0,1)^\top,(1,0)^\top\} \\ \mathbf{h}'\in\{(0,1)^\top,(1,0)^\top\}}} \frac{e^{\eta(\beta g_{\theta,\tau}(\mathbf{x})(1+b|\,\mathsf{S}(\mathbf{h})|)^{1-\theta}\langle \mathbf{h}^c,\mathbf{h}'\rangle - c)}}{1+e^{\eta(\beta g_{\theta,\tau}(\mathbf{x})(1+b|\,\mathsf{S}(\mathbf{h})|)^{1-\theta}\langle \mathbf{h}^c,\mathbf{h}'\rangle - c)}} \frac{e^{\eta(\beta g_{\theta,\tau}(\mathbf{x})(1+b|\,\mathsf{S}(\mathbf{h}')|)^{1-\theta}\langle \mathbf{h}'^c,\mathbf{h}\rangle - c)}}{1+e^{\eta(\beta g_{\theta,\tau}(\mathbf{x})(1+b|\,\mathsf{S}(\mathbf{h}')|)^{1-\theta}\langle \mathbf{h}'^c,\mathbf{h}\rangle - c)}}$$

$$= \begin{cases} 2\left(\frac{e^{-2\eta c}}{(1+e^{-\eta c})^2} + \frac{e^{2\eta(\beta g_{1,\tau}-c)}}{(1+e^{\eta(\beta g_{1,\tau}-c)})^2}\right) = 2\left(\frac{e^{-2\eta c}}{(1+e^{-\eta c})^2} + \frac{e^{2\eta(\beta b\tau - c)}}{(1+e^{\eta(\beta b\tau - c)})^2}\right) & \text{if } \theta = 1, \\ 2\left(\frac{e^{-2\eta c}}{(1+e^{-\eta c})^2} + \frac{e^{2\eta((1+b)\beta g_{0,\tau}(\mathbf{x})-c)}}{(1+e^{\eta((1+b)\beta g_{0,\tau}(\mathbf{x})-c)})^2}\right) = 2\left(\frac{e^{-2\eta c}}{(1+e^{-\eta c})^2} + \frac{e^{2\eta((1+b)\beta\frac{2\tau b}{1+bh_t(\mathbf{x})}-c)}}{(1+e^{\eta((1+b)\beta\frac{2\tau b}{1+bh_t(\mathbf{x})}-c)})^2}\right) & \text{if } \theta = 0, \end{cases}$$

$$\tilde{g}^{\eta}_{\theta,\tau}(1,2;\bar{h}_t) = \sum_{\substack{\mathbf{h}\in\{(0,1)^\top,(1,0)^\top\} \\ \mathbf{h}'=(1,1)^\top}} \frac{e^{\eta(\beta g_{\theta,\tau}(\mathbf{x})(1+b|\,\mathsf{S}(\mathbf{h})|)^{1-\theta}\langle \mathbf{h}^c,\mathbf{h}'\rangle - c)}}{1+e^{\eta(\beta g_{\theta,\tau}(\mathbf{x})(1+b|\,\mathsf{S}(\mathbf{h})|)^{1-\theta}\langle \mathbf{h}^c,\mathbf{h}'\rangle - c)}} \frac{e^{\eta(\beta g_{\theta,\tau}(\mathbf{x})(1+b|\,\mathsf{S}(\mathbf{h}')|)^{1-\theta}\langle \mathbf{h}'^c,\mathbf{h}\rangle - c)}}{1+e^{\eta(\beta g_{\theta,\tau}(\mathbf{x})(1+b|\,\mathsf{S}(\mathbf{h}')|)^{1-\theta}\langle \mathbf{h}'^c,\mathbf{h}\rangle - c)}}$$

$$= \begin{cases} 2\frac{e^{-\eta c}}{1+e^{-\eta c}}\frac{e^{\eta(\beta g_{0,\tau}-c)}}{1+e^{\eta(\beta g_{0,\tau}-c)}} = 2\frac{e^{-\eta c}}{1+e^{-\eta c}}\frac{e^{\eta(\beta b\tau - c)}}{1+e^{\eta(\beta b\tau - c)}} & \text{if } \theta = 1, \\ 2\frac{e^{-\eta c}}{1+e^{-\eta c}}\frac{e^{\eta((1+b)\beta g_{0,\tau}(\mathbf{x})-c)}}{1+e^{\eta((1+b)\beta g_{0,\tau}(\mathbf{x})-c)}} = 2\frac{e^{-\eta c}}{1+e^{-\eta c}}\frac{e^{\eta((1+b)\beta\frac{2\tau b}{1+bh_t(\mathbf{x})}-c)}}{1+e^{\eta((1+b)\beta\frac{2\tau b}{1+bh_t(\mathbf{x})}-c)}} & \text{if } \theta = 0, \end{cases}$$

$$\tilde{g}^{\eta}_{\theta,\tau}(0,0;\bar{h}_t) = \sum_{\substack{\mathbf{h}=(0,0)^\top, \\ \mathbf{h}'=(0,0)^\top}} \frac{e^{\eta(\beta g_{\theta,\tau}(\mathbf{x})(1+b|\,\mathsf{S}(\mathbf{h})|)^{1-\theta}\langle \mathbf{h}^c,\mathbf{h}'\rangle - c)}}{1+e^{\eta(\beta g_{\theta,\tau}(\mathbf{x})(1+b|\,\mathsf{S}(\mathbf{h})|)^{1-\theta}\langle \mathbf{h}^c,\mathbf{h}'\rangle - c)}} \frac{e^{\eta(\beta g_{\theta,\tau}(\mathbf{x})(1+b|\,\mathsf{S}(\mathbf{h}')|)^{1-\theta}\langle \mathbf{h}'^c,\mathbf{h}\rangle - c)}}{1+e^{\eta(\beta g_{\theta,\tau}(\mathbf{x})(1+b|\,\mathsf{S}(\mathbf{h}')|)^{1-\theta}\langle \mathbf{h}'^c,\mathbf{h}\rangle - c)}}$$

$$= \begin{cases} \frac{e^{-2\eta c}}{(1+e^{-\eta c})^2} & \text{if } \theta = 1, \\ \frac{e^{-2\eta c}}{(1+e^{-\eta c})^2} & \text{if } \theta = 0, \end{cases}$$

$$\tilde{g}^{\eta}_{\theta,\tau}(2,2;\bar{h}_t) = \sum_{\substack{\mathbf{h}=(1,1)^\top, \\ \mathbf{h}'=(1,1)^\top}} \frac{e^{\eta(\beta g_{\theta,\tau}(\mathbf{x})(1+b|\,\mathsf{S}(\mathbf{h})|)^{1-\theta}\langle \mathbf{h}^c,\mathbf{h}'\rangle - c)}}{1+e^{\eta(\beta g_{\theta,\tau}(\mathbf{x})(1+b|\,\mathsf{S}(\mathbf{h})|)^{1-\theta}\langle \mathbf{h}^c,\mathbf{h}'\rangle - c)}} \frac{e^{\eta(\beta g_{\theta,\tau}(\mathbf{x})(1+b|\,\mathsf{S}(\mathbf{h}')|)^{1-\theta}\langle \mathbf{h}'^c,\mathbf{h}\rangle - c)}}{1+e^{\eta(\beta g_{\theta,\tau}(\mathbf{x})(1+b|\,\mathsf{S}(\mathbf{h}')|)^{1-\theta}\langle \mathbf{h}'^c,\mathbf{h}\rangle - c)}}$$

$$= \begin{cases} \frac{e^{-2\eta c}}{(1+e^{-\eta c})^2} & \text{if } \theta = 1, \\ \frac{e^{-2\eta c}}{(1+e^{-\eta c})^2} & \text{if } \theta = 0. \end{cases}$$

Finally, we from Equation (9) have that

$$f^{\vartheta}_{\theta,\kappa}(\bar{h}_t) \equiv \frac{e^{\vartheta\left(\theta+\frac{1-\theta}{1+b\bar{h}_t}-\kappa\right)}}{1+e^{\vartheta\left(\theta+\frac{1-\theta}{1+b\bar{h}_t}-\kappa\right)}} = \begin{cases} \frac{e^{\vartheta(1-\kappa)}}{1+e^{\vartheta(1-\kappa)}} & \text{if } \theta = 1, \\ \frac{e^{\vartheta\left(\frac{1}{1+2b(\bar{x}_t(1)+\bar{x}_t(2))}-\kappa\right)}}{1+e^{\vartheta\left(\frac{1}{1+2b(\bar{x}_t(1)+\bar{x}_t(2))}-\kappa\right)}} & \text{if } \theta = 0. \end{cases}$$

An example of a numerical simulation of the stochastic process introduced in in Sections 3.2, 3.3 and 3.4 using the "next reaction method" for simulating a continuous time Markov chain [cf. Anderson, 2012; Gibson and Bruck, 2000], and the solution of the ODEs in Equations (87)–(88) superimposed is shown in Figure J.2.

A comparison of the agent based simulation for $N = 2$ and $n = 200$ with the numerical solutions of the ODEs using the large ρ approximation of Section G.2 and the pair approximation of Appendix J.2 is shown in Figure J.3. We observe that the pair approximation predicts the correct behavior of the time evolution of the average stock of knowledge and the average degree for a value of $\rho = 1$, while for $\rho = 50$, both, the pair approximation and the large ρ approximation predict the right behavior and yield identical results.

K. Data

We obtain information on R&D collaborations from the Thomson Securities Data Company (SDC) alliance database. SDC collects data from the U. S. Securities and Exchange Commission (SEC) filings (and their international counterparts), trade publications, wires, and news sources. We include only alliances from SDC which are classified as (i) alliances that imply the transmission of an existing technology from one partner to another or to the alliance, (ii) alliances that imply the cross-transfer of existing technologies between two or more partners or between these and the alliance, and (iii) alliances that include the undertaking of R&D activities. This gives us a total of $21,478$ firms in our sample.

K.1. Country Composition

The number of firms in each country is shown in Figure K.1 while Table K.1 shows the 25 countries with the largest numbers of firms. The dominant role of the U.S. with 11,801 collaborations making up 55.18% of the total number of collaborations is clearly visible. The second largest country in terms of R&D collaborations is Japan with 1,331, which comprises 6.22% of all collaborations.

K.2. Mergers and Acquisitions

Some firms might be acquired by other firms due to mergers and acquisitions (M&A) over time, and this will impact the R&D collaboration network [cf. Hanaki et al., 2010]. To get a comprehensive picture of the M&A activities of the firms in our dataset, we use two extensive data sources to obtain information about M&As. The first is the Thomson Reuters' Securities Data Company (SDC) M&A database, which has historically been the most widely used database for empirical research in the field of M&As. The second database with information about M&As is Bureau van Dijk's (BvD) Zephyr database, which is a recent alternative to the SDC M&As database. We merged the SDC and Zephyr databases (using a name matching algorithm; see also Atalay et al. [2011]; Trajtenberg et al. [2009]) to obtain information on M&As of $116,641$ unique firms. Using the same name matching algorithm we could identify roughly 40% of the firms in the SDC alliance database that also appear in the combined SDC-Zephyr M&As database. We then account for the M&A activities of these matched firms when constructing the R&D collaboration network by assuming that an acquiring firm in a M&A inherits all the R&D collaborations of the target firm, and we remove the target firm form from the network.

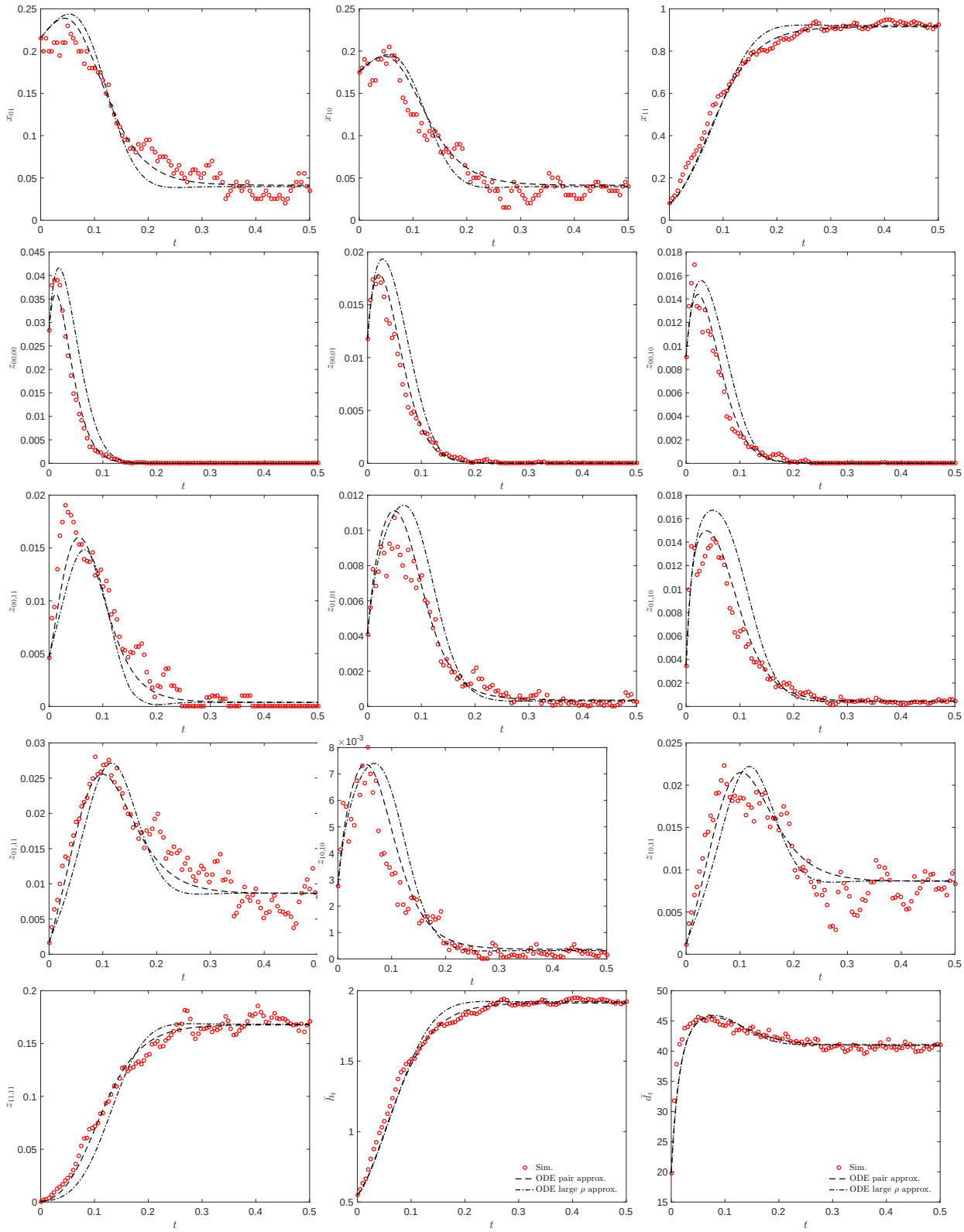


Figure J.2: Comparison of the agent based simulations for $N = 2$ and $n = 200$ using the numerical solutions of the ODEs using the large ρ approximation and the pair approximation for $\rho = 50$.

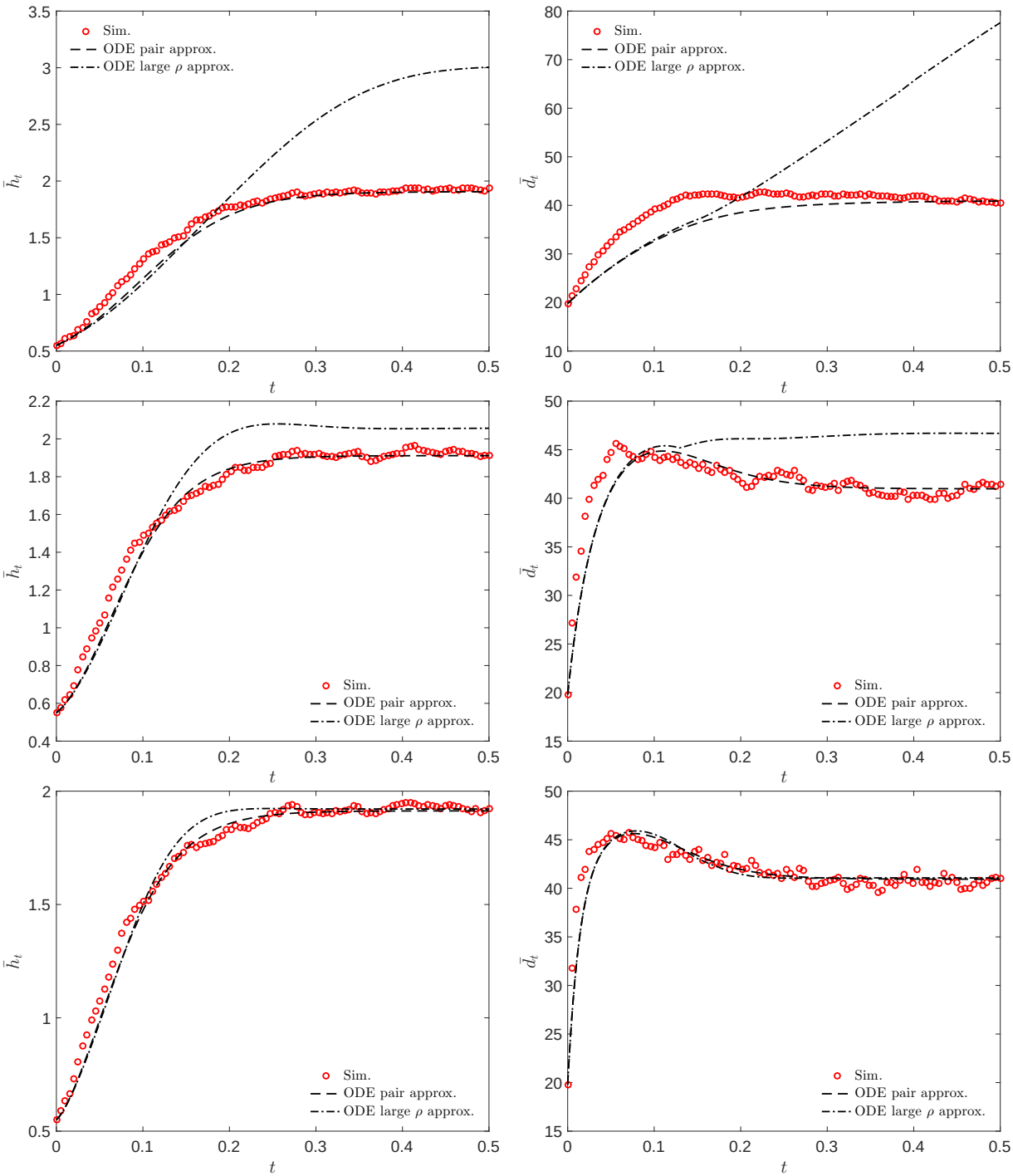


Figure J.3: Comparison of the agent based simulation for $N = 2$ and $n = 200$ with the numerical solutions of the ODEs using the large ρ approximation of Section G.2 and the pair approximation of Appendix J.2. The left panels show the average stock of knowledge, $\bar{h}_t$, and the right panels the average degree $\bar{d}_t$. The top panels correspond to $\rho = 5$, the middle panels to $\rho = 25$ while the bottom panels correspond to $\rho = 50$.

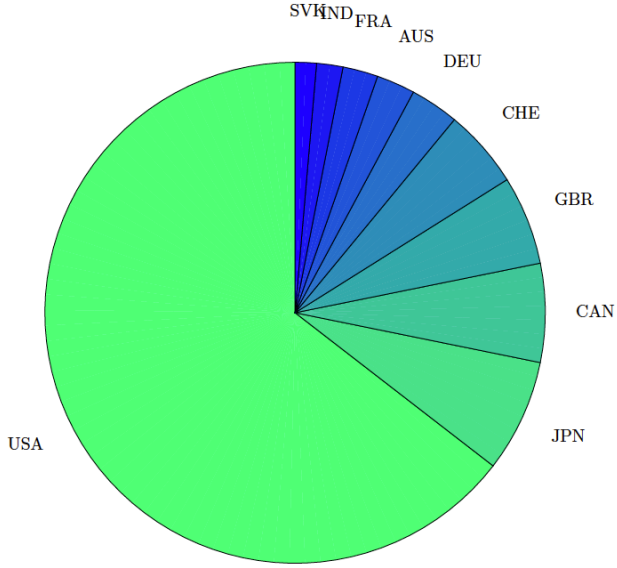


Figure K.1: The number of firms in each country.

Table K.1: The 25 countries with the largest numbers of firms in the SDC dataset.

Name	Code	# firms	% of tot.	Rank
United States	USA	11801	55.18	1
Japan	JPN	1331	6.22	2
Canada	CAN	1164	5.44	3
United Kingdom	GBR	1050	4.91	4
Switzerland	CHE	932	4.36	5
Germany	DEU	575	2.69	6
Australia	AUS	456	2.13	7
France	FRA	416	1.95	8
India	IND	311	1.45	9
Slovakia	SVK	251	1.17	10
Hong Kong	HKG	202	0.94	11
Netherlands	NLD	202	0.94	12
Taiwan	TWN	185	0.87	13
Sweden	SWE	173	0.81	14
Iceland	ISL	168	0.79	15
Italy	ITA	168	0.79	16
Singapore	SGP	164	0.77	17
Russia	RUS	143	0.67	18
Belgium	BEL	103	0.48	19
Denmark	DNK	95	0.44	20
Thailand	THA	91	0.43	21
Spain	ESP	87	0.41	22
Iran	IRN	80	0.37	23
Norway	NOR	74	0.35	24
Brazil	BRA	59	0.28	25

Table K.2: The 20 largest sectors at the 3-digit SIC level.

Sector	3-dig SIC	# firms	% of tot.	Rank
Computer and Data Processing Services	737	4996	23.36	1
Drugs	283	2399	11.22	2
Research and Testing Services	873	1088	5.09	3
Miscellaneous Investing	679	902	4.22	4
Electronic Components and Accessories	367	821	3.84	5
Computer and Office Equipment	357	679	3.17	6
Medical Instruments and Supplies	384	548	2.56	7
Telephone Communication	481	541	2.53	8
Communications Equipment	366	530	2.48	9
Measuring and Controlling Devices	382	397	1.86	10
Management and Public Relations	874	293	1.37	11
Motor Vehicles and Equipment	371	254	1.19	12
Engineering Architectural Services	871	189	0.88	13
Crude Petroleum and Natural Gas	131	180	0.84	14
Professional and Commercial Equipment	504	173	0.81	15
Misc. Business Services	738	169	0.79	16
Industrial Inorganic Chemicals	281	163	0.76	17
Electric Services	491	145	0.68	18
Plastics Materials and Synthetic	282	142	0.66	19
Misc. Chemical Products	289	139	0.65	20

K.3. Alliance Durations and R&D Network Construction

In the following sections we discuss alternative ways to construct the R&D collaboration network. In Section K.3.1 we make different assumptions about the (fixed) duration of an alliance. Moreover, in Section K.3.2 we construct the network by assuming that the duration of an alliance is random (based on the subset of observed alliances with known start and end date).

K.3.1. Variable Alliance Durations from 3 to 7 Years

Figure K.2 shows the average degree $\bar{d}$ over the years 1985 to 2012 with different durations of an alliance ranging from 3 to 7 years. We observe that the second peak becomes more pronounced relative to the first peak for shorter durations of an alliance.

K.3.2. Random Alliance Durations Following Empirical Distribution

For some of the R&D alliances in the Thomson SDC database their duration is reported. The number of such alliances for which both the starting and end date are known is 311. The average alliance duration is 2.88 years with a standard deviation of 2.33. The left panel in Figure K.3 shows the histogram of the empirical alliance durations with a fitted exponential distribution overlaid. The right panel in Figure K.3 shows the average degree, $\bar{d}$, when the duration of an alliance is drawn from the empirical distribution of alliance durations. The plot of the average degree, $\bar{d}$, is reminiscent of the top left panel in Figure 2 where a fixed alliance duration of 5 years has been assumed.

K.4. Sectoral Composition

Figure K.4, and Table K.2 show the 10, respectively 20, largest sectors at the 3-digit SIC level. The largest sector at the SIC-737 level is Computer and Data Processing Services, with 4,996 firms (23.36 % of the total), followed by the Drugs sector, with 2,399 firms (11.22 % of the total).

Figures K.5 and K.6, respectively, show the empirical average degree $\bar{d}$ over the years 1985 to 2012 for the firms in the 10 largest SIC 3-digit sectors (cf. Table K.2 and Figure K.4). In all sectors

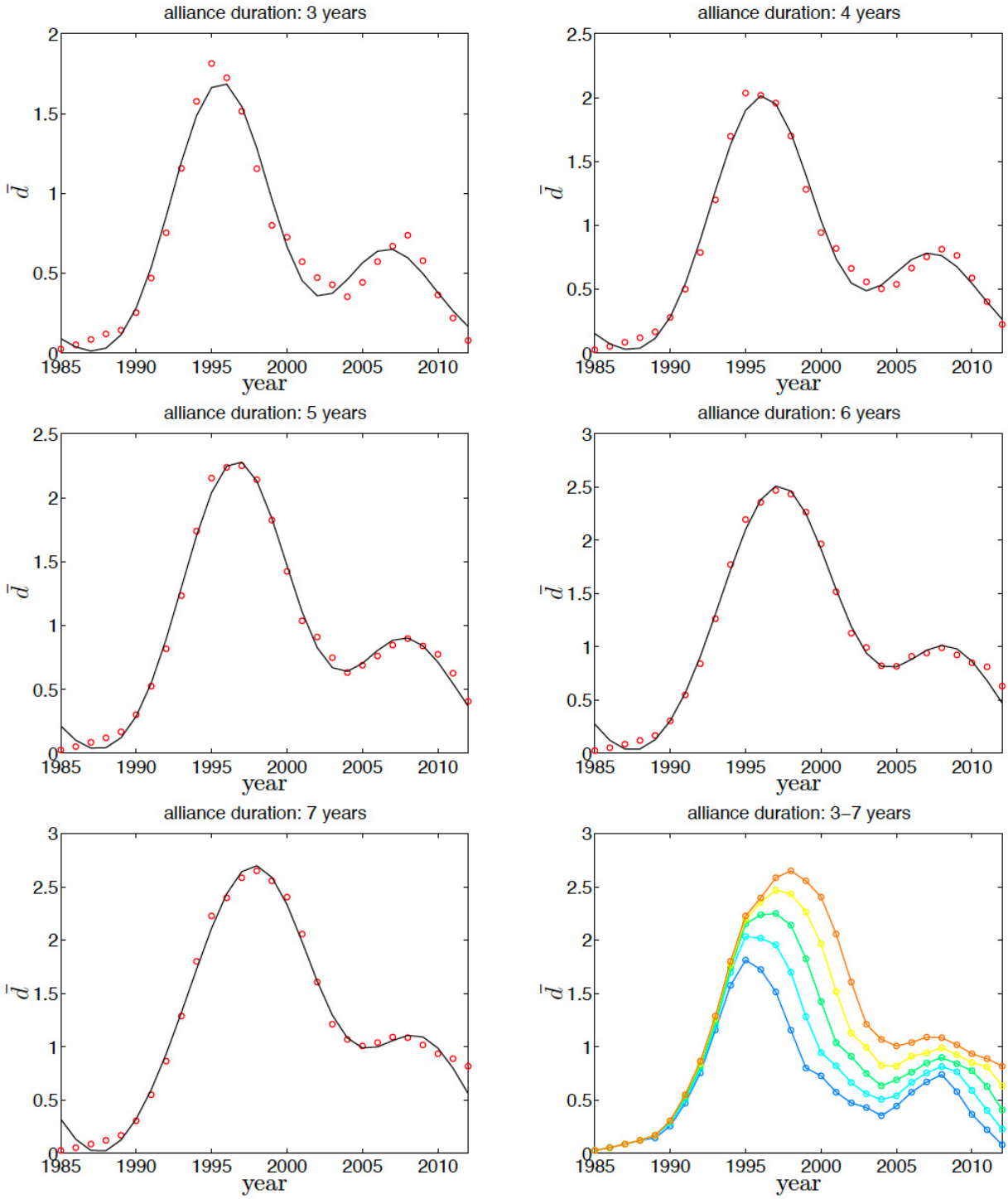


Figure K.2: The average degree $\bar{d}$ over the years 1985 to 2012. The circles indicate the empirical observations while the curve indicates a Fourier regression fit as in Figure 2.

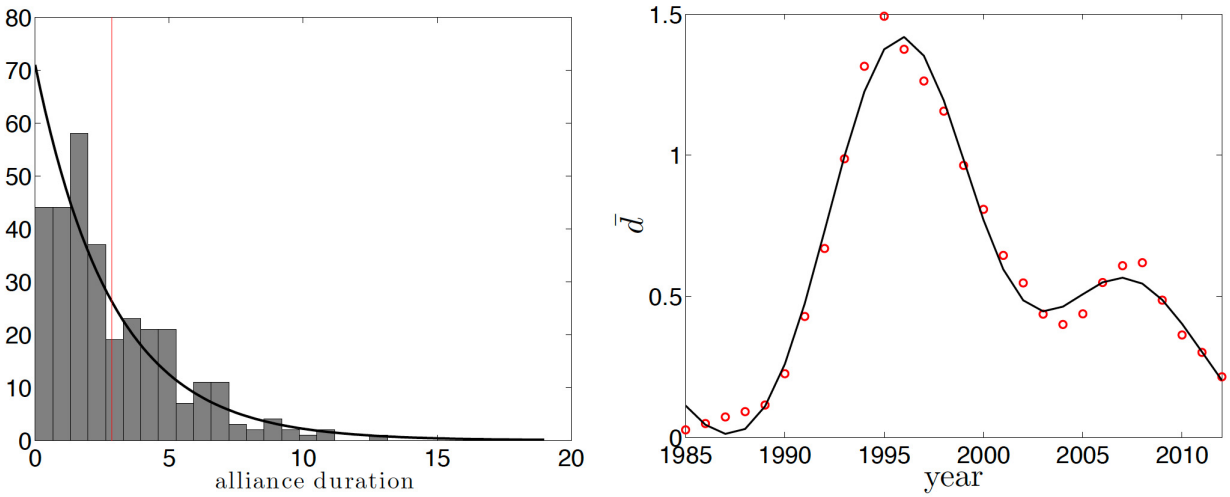


Figure K.3: (Left panel) The histogram of the empirical alliance durations with a fitted exponential distribution overlaid. The vertical line indicates the mean of 2.88 years. (Right panel) The average degree, $\bar{d}$, when the duration of an alliance is drawn from the empirical distribution of the observed alliance durations. The circles indicate the empirical observations while the curve indicates a Fourier regression fit as in Figure 2.

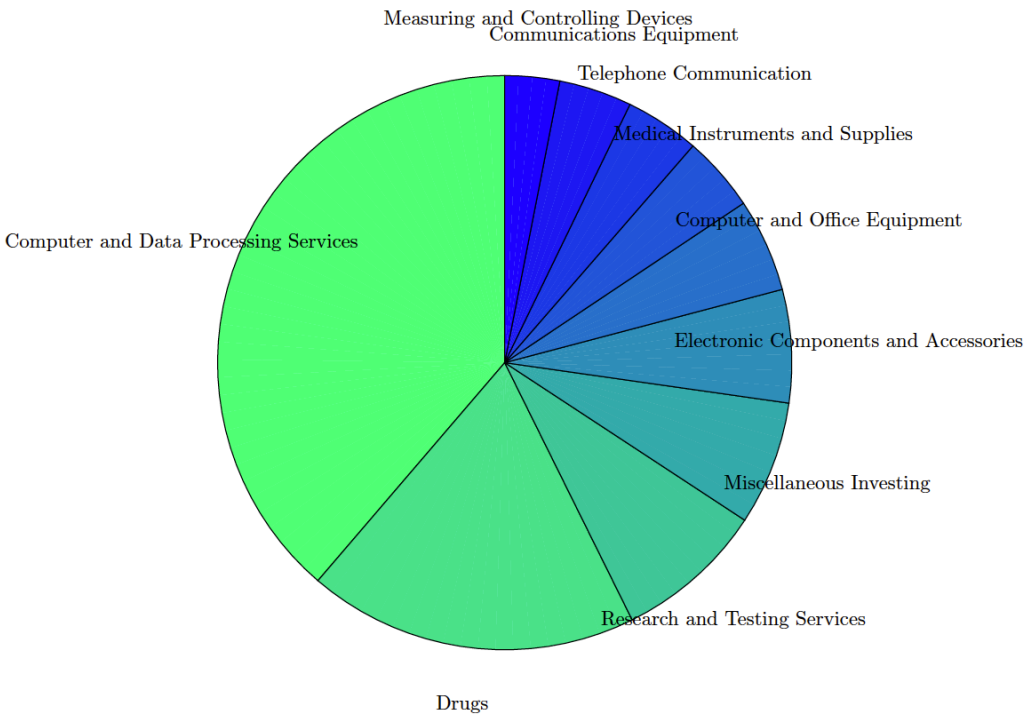


Figure K.4: The shares of the ten largest sectors at the 3-digit (right panel) SIC levels. See also Table K.2.

Table K.3: Estimated coefficients with their standard deviations, t-statistics and p-values for the regression model $\bar{d}_t = \frac{1}{2}a_0 + \sum_{j=1}^{3}(a_j \cos(\omega_j t) + b_j \sin(\omega_j t))$, with $t \in [0, T]$, the parameters a_0, a_j, b_j, $j = 1, \ldots, 4$ (Fourier coefficients), where the angular frequency is given by $\omega_j \equiv \frac{2\pi}{T} j$, $\bar{d}_t$ is periodic with period T, and the average degree $\bar{d}_t$ stemming from the CATI database. See e.g. Hamilton [1994, Chapter 6.2] for further details.

θ	$\hat{\theta}$	$\hat{\sigma}_\theta$	$t_{\hat{\theta}}$	$p_{\hat{\theta}}$
a_0	0.8999^{***}	0.0182	49.4020	0.0000
a_1	-0.2405^{***}	0.0257	−9.3380	0.0000
b_1	0.0822^{***}	0.0257	3.1927	0.0053
a_2	-0.0616^{**}	0.0257	−2.3939	0.0284
b_2	-0.2107^{***}	0.0257	−8.1794	0.0000
a_3	0.0999^{***}	0.0257	3.8793	0.0012
b_3	−0.0251	0.0257	−0.9767	0.3424
a_4	0.0232	0.0257	0.9035	0.3788
b_4	-0.0588^{**}	0.0257	−2.2854	0.0354

The number of observations is $T = 26$, the error degrees of freedom is 17. The root mean squared error is given by 0.0929. R^2 is 0.919, adjusted R^2 is 0.88. F-statistic vs. constant model is 24, and the p-value is approximately zero.
*** Statistically significant at 1% level.
** Statistically significant at 5% level.
* Statistically significant at 10% level.

we observe a non-monotonic pattern for the average degree, reminiscent of the oscillations that we osberve in the aggregate in Figure 2.

K.5. Alternative Data on R&D Collaborations

To check the robustness of our findings in Section 2 we use an alternative datasource on interfirm R&D collaborations that has been widely used in the literature stemming from the Cooperative Agreements and Technology Indicators (CATI) database [cf. Hagedoorn, 2002]. The database only records agreements for which a combined innovative activity or an exchange of technology is at least part of the agreement. Moreover, only agreements that have at least two industrial partners are included in the database, thus agreements involving only universities or government labs, or one company with a university or lab, are disregarded. The data collection of the database ends in the year 2006, and it has a more extensive coverage of firms in the biotech sector [cf. Schilling, 2009].

Figure K.7 shows the average degree over time with a Fourier fit superimposed. Similar to the average degree shown in Figure 2 in Section 2, which was based on collaborations in the Thomson SDC database, we observe an oscillation in the average degree over time with a decreasing amplitude, but with a different phase potentially due to the different sectoral composition of the CATI data. The oscillatory pattern is also supported by the significance of the Fourier coefficient estimates shown in Table K.3.

K.6. U.S. Business Cycle

Since 55% of the firms in our data are from the U.S. (cf. Figure K.1 and Table K.1) one might think that the cyclical pattern in the R&D alliance data is driven by movements of the U.S. business cycle. Figure K.8 shows the U.S. business cycle as estimated in Magrini et al. [2013] together with

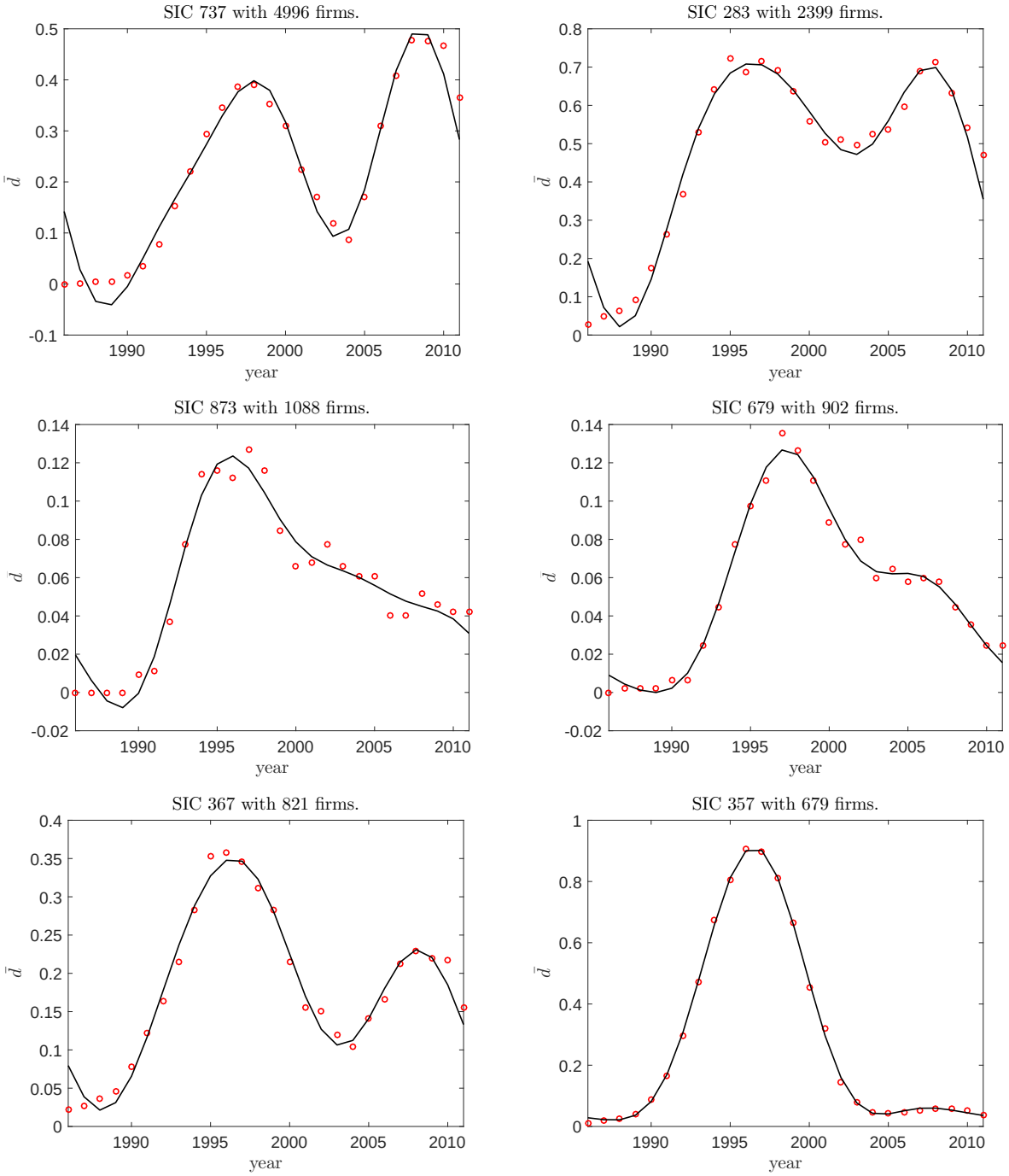


Figure K.5: The average degree $\bar{d}$ over the years 1985 to 2012 for the firms in the 6 largest SIC 3-digit sectors: 737, 283, 873, 679, 367 and 357. The circles indicate the empirical observations while the curve indicates a Fourier regression fit as in Figure 2.

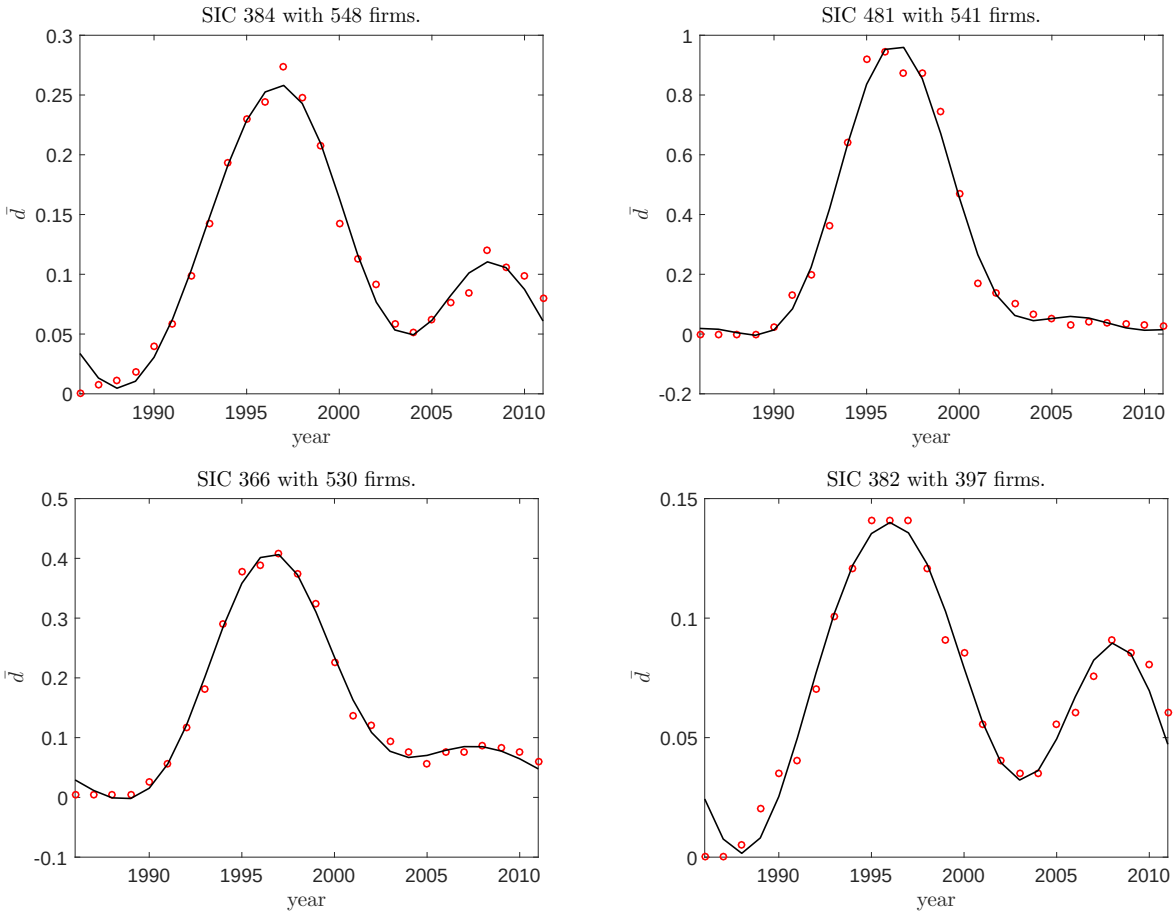


Figure K.6: The average degree $\bar{d}$ over the years 1985 to 2012 for the firms in the 7th to the 10th largest SIC 3-digit sectors: 384, 481, 366 and 382. The circles indicate the empirical observations while the curve indicates a Fourier regression fit as in Figure 2.

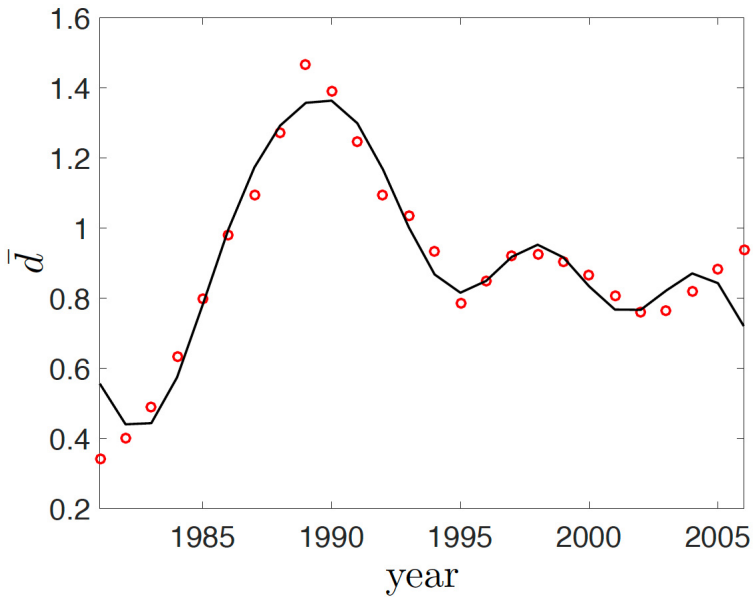


Figure K.7: The average degree $\bar{d}$ over the years 1980 to 2006 from the CATI database assuming that a collaboration lasts for 5 years. The circles indicate the empirical observations while the curve indicates a Fourier regression fit with the parameter estimates from Table K.3.

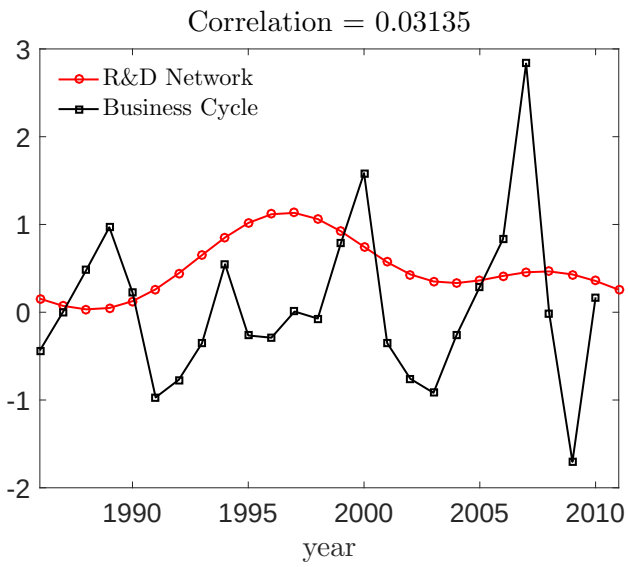


Figure K.8: The R&D collaboration intensity and the U.S. business cycle taken from Magrini et al. [2013]. Both are weakly correlated with a correlation coefficient of 0.031.

the R&D collaboration intensity.[70] The figure indicates that both are only weakly correlated with a correlation coefficient of 0.031.

L. Estimation Algorithm

The purpose of the LF-MCMC algorithm is to estimate the parameter vector $\boldsymbol{\delta} \equiv (\alpha, \beta, \gamma, \rho, \eta, \lambda, \tau, b, c, \xi, \chi, \vartheta, \kappa)$, of the model on the basis of the summary statistics of the average degree $\mathbf{S}^o \equiv (\bar{d}_t^{\text{obs}})_{t=1985}^{2011}$. The algorithm generates a Markov chain which is a sequence of parameters $(\boldsymbol{\delta}_s)_{s=1}^S$ with a stationary distribution that approximates the distribution of each parameter value $\delta \in \boldsymbol{\delta}$ conditional on the observed statistic $\mathbf{S}^o$.

Definition 1. *Consider the statistics* $\mathbf{S}$ *and denote by* $\mathbf{S}^o$ *the observed statistics. Further, let* $\Delta(\mathbf{S}^o, \mathbf{S})$ *be a measure of distance between the realized statistic* $\mathbf{S}$ *of the model with parameter vector* $\boldsymbol{\delta}$ *and the observed statistic* $\mathbf{S}^o$*. Then we consider the Markov chain* $(\boldsymbol{\delta}_s)_{s=1}^S$ *induced by the following algorithm:*

(i) *Given* $\boldsymbol{\delta}$*, propose* $\boldsymbol{\delta}'$ *according to the proposal density* $q_s(\boldsymbol{\delta} \to \boldsymbol{\delta}')$*.*

(ii) *Generate a network according to* $\boldsymbol{\delta}'$ *and calculate the summary statistics* $\mathbf{S}'$*.*

(iii) *Calculate*

$$h(\boldsymbol{\delta}, \boldsymbol{\delta}') = \min\left(1, \frac{q_s(\boldsymbol{\delta}' \to \boldsymbol{\delta})}{q_s(\boldsymbol{\delta} \to \boldsymbol{\delta}')} \mathbb{1}_{\{\Delta(\mathbf{S}', \mathbf{S}^o) < \epsilon_s\}}\right),$$

where $\epsilon_s \geq 0$ *is a monotonic decreasing sequence of threshold values,* $\epsilon_s \downarrow \epsilon^{\min}$*, and* $\Delta : \mathbb{R}_+^T \times \mathbb{R}_+^T \to \mathbb{R}_+$ *is a distance metric in* $\mathbb{R}_+^T$*.*

(iv) *Accept* $\boldsymbol{\delta}'$ *with probability* $h(\boldsymbol{\delta}, \boldsymbol{\delta}')$*, otherwise stay at* $\boldsymbol{\delta}$ *and go to (i).*

Marjoram et al. [2003] have shown that the distribution generated by the above algorithm converges to the true conditional distribution of the parameter vector $\boldsymbol{\delta}$, given the observations $\mathbf{S}^o$ and the threshold values.

[70] We would like to thank Stefano Magrini and Hasan Engin Duran for making their data available to us.

Table L.1: Estimation of the model parameters $\delta \in \boldsymbol{\delta} \equiv (\alpha, \beta, \gamma, \rho, \eta, \lambda, \tau, b, c, \xi, \chi, \vartheta, \kappa)$ for the competitive case when $\theta = 0$. The table shows simulated averages of the parameters and their standard deviations,[a] after the chain has converged.[b]

δ	μ_δ	$\bar{\sigma}_\delta$	σ_δ	ι_δ	$p_\delta(S)$
α	0.8698	0.3168	0.0685	105.5937	0.0453
β	0.2202	0.0255	0.0057	122.7108	0.6719
γ	0.0001	0.0000	0.0000	47.5886	0.2926
ρ	75.5182	6.9830	1.5636	121.2832	0.7003
η	1.4279	0.1088	0.0241	123.7539	0.8121
λ	2.1479	0.1925	0.0432	114.0198	0.6897
τ	0.0083	0.0039	0.0009	117.1475	0.1080
b	1279.8068	8.8795	1.9927	111.8273	0.9750
c	1.2616	0.0300	0.0065	58.5831	0.9305
ξ	0.7468	0.0680	0.0144	82.2030	0.7841
χ	0.5308	0.0766	0.0164	117.0896	0.6909
ϑ	16.2861	2.9225	0.6548	119.3006	0.5073
κ	0.7047	0.1438	0.0323	115.9874	0.9763
n			500		
S			1000		

[a] μ_δ is the average and $\bar{\sigma}_\delta$ is the simulation standard deviation of the respective parameter, while σ_δ is the standard deviation calculated from batch means (of length 10) for each parameter $\delta \in \boldsymbol{\delta}$ [Chib, 2001]. ι_δ is the integrated autocorrelation time which should be much smaller than the number S of iterations of the Markov chain [Sokal, 1996].

[b] $p_\delta(S)$ is the p-value of Geweke's spectral density diagnostic (converging in distribution to a standard normal random variable as $S \to \infty$) indicating the convergence of the chain [Brooks and Roberts, 1998; Geweke, 1992]. The maximum number of iterations, S, has been chosen such that sufficiently high values of $p_\delta(S)$ were obtained.

The proposal distribution $q_s(\boldsymbol{\delta} \to \boldsymbol{\delta}')$ is a truncated normal distribution $\boldsymbol{\delta}' \sim \mathcal{N}(\boldsymbol{\delta}, \boldsymbol{\Sigma}_s)$ $\mathbb{1}_{[\boldsymbol{\delta}^{\min}, \boldsymbol{\delta}^{\max}]}(\boldsymbol{\delta})$ for each parameter $\delta \in \boldsymbol{\delta}$ with a diagonal variance-covariance matrix $\boldsymbol{\Sigma}_s = \mathrm{diag}\{\sigma^2_{1,s}, \ldots, \sigma^2_{L,s}\}$. More precisely, for each parameter $\theta_i \in \mathbb{R}_+$ we choose a proposal distribution given by

$$q_s(\delta \to \delta') = \frac{\phi(\delta'|\delta, \sigma_s^2)}{\tilde{\phi}(\delta^{\max}|\delta, \sigma_s^2) - \tilde{\phi}(\delta^{\min}|\delta, \sigma_s^2)} \mathbb{1}_{[\delta^{\min}, \delta^{\max}]}(\delta'),$$

where $\phi(\delta|\mu, \sigma^2)$ and $\tilde{\phi}(\delta|\mu, \sigma^2)$ are the pdf and cdf, respectively, of a normally distributed random variable with mean μ and variance σ^2. During the "burn-in" phase [Chib, 2001], we consider a monotonic decreasing sequence of thresholds given by $\epsilon_s \geq \epsilon_{s+1} \geq \ldots \geq \epsilon^{\min}$ with $\epsilon_{s+1} = \max\left\{(1-\gamma)\epsilon_s, \epsilon^{\min}\right\}$ and $\gamma = 0.05$. Similarly, we assume a decreasing sequence of variances $\sigma_s^2 \geq \sigma_{s+1}^2 \geq \ldots \geq (\sigma^{\min})^2$ with $\sigma_{s+1}^2 = \max\left\{(1-\gamma)\sigma_s^2, (\sigma^{\min})^2\right\}$ for the proposal distribution $q_s(\delta \to \delta')$. The maximum number of iterations, S, has been chosen such that reasonably high values of $p_\delta(S)$ were obtained. As a measure of distance we choose the Euclidean distance $\Delta(\mathbf{S}, \mathbf{S}^o) = \sqrt{\sum_{t=1985}^{2012} \left(\bar{d}_t - \bar{d}_t^{\text{obs}}\right)^2}$. The parameter ranges are $\alpha \in [0, 5]$, $\beta \in [0, 5]$, $\gamma \in [0, 0.01]$, $b \in [0, 5000]$, $c \in [0, 5]$, $\tau \in [0, 0.1]$, $\rho \in [10, 1000]$, $\eta \in= [0, 5]$, $\lambda \in [0, 5]$, $\xi \in [0, 2]$, $\chi \in [0, 2]$, $\vartheta \in [0, 30]$ and $\kappa \in [0, 2]$. The parameters $\epsilon^{\min}$ are chosen sufficiently small after long experimentation with different starting values and burn-in periods. The parameter estimates from this procedure are shown in Table L.1.

References

Aghion, P., Akcigit, U., Bergeaud, A., Blundell, R., and Hémous, D. (2015). Innovation and top income inequality. *National Bureau of Economic Research, Working Paper No. w21247*.

Aghion, P., Bloom, N., Blundell, R., Griffith, R., and Howitt, P. (2005). Competition and Innovation: An Inverted-U Relationship. *Quarterly Journal of Economics*, 120(2):701–728.

Anderson, D. F. (2012). An efficient finite difference method for parameter sensitivities of continuous time markov chains. *SIAM Journal on Numerical Analysis*, 50(5):2237–2258.

Atalay, E., Hortacsu, A., Roberts, J., and Syverson, C. (2011). Network structure of production. *Proceedings of the National Academy of Sciences*, 108(13):5199.

Ballester, C., Calvó-Armengol, A., and Zenou, Y. (2006). Who's who in networks. wanted: The key player. *Econometrica*, 74(5):1403–1417.

Belhaj, M., Bervoets, S., and Deroian, F. (2016). Optimal networks in games with local complementarities. *Theoretical Economics*, 11.

Bonacich, P. (1987). Power and centrality: A family of measures. *American Journal of Sociology*, 92(5):1170–1182.

Brooks, S. P. and Roberts, G. O. (1998). Assessing convergence of Markov chain Monte Carlo algorithms. *Statistics and Computing*, 8(4):319–335.

Brualdi, R. A. and Solheid, Ernie, S. (1986). On the spectral radius of connected graphs. *Publications de l' Institute Mathématique*, 53:45–54.

Cameron, A. and Trivedi, P. (2005). *Microeconometrics: methods and applications*. Cambridge University Press.

Cator, E. and Van Mieghem, P. (2013). Susceptible-infected-susceptible epidemics on the complete graph and the star graph: Exact analysis. *Physical Review E*, 87(1):012811.

Chib, S. (2001). Markov chain Monte Carlo methods: computation and inference. *Handbook of econometrics*, 5:3569–3649.

Cvetkovic, D. and Rowlinson, P. (1990). The largest eigenvalue of a graph: A survey. *Linear and Multinilear Algebra*, 28(1):3–33.

Eeckhout, J. and Jovanovic, B. (2002). Knowledge spillovers and inequality. *American Economic Review*, 92(5):1290–1307.

Geweke, J. (1992). *Bayesian Statistics 4*, chapter Evaluating the Accuracy of Sampling-Based Approaches to the Calculation of Posterior Moments, pages 169–193. Oxford University Press.

Gibson, M. A. and Bruck, J. (2000). Efficient exact stochastic simulation of chemical systems with many species and many channels. *The Journal of Physical Chemistry A*, 104(9):1876–1889.

Gross, T., D'Lima, C. J. D., and Blasius, B. (2006). Epidemic dynamics on an adaptive network. *Physical review letters*, 96(20):208701.

Hagedoorn, J. (2002). Inter-firm R&D partnerships: an overview of major trends and patterns since 1960. *Research Policy*, 31(4):477–492.

Hall, B. H., Jaffe, A. B., and Trajtenberg, M. (2001). The NBER Patent Citation Data File: Lessons, Insights and Methodological Tools. *NBER Working Paper No. 8498*.

Hamilton, J. (1994). *Time series analysis*, volume 2. Cambridge University Press.

Hanaki, N., Nakajima, R., and Ogura, Y. (2010). The dynamics of R&D network in the IT industry. *Research Policy*, 39(3):386–399.

Hojman, D. and Szeidl, A. (2008). Core and periphery in networks. *Journal of Economic Theory*, 139(1):295–309.

Jackson, M. and Rogers, B. (2007). Relating network structure to diffusion properties through stochastic dominance. *The BE Journal of Theoretical Economics*, 7(1):1–13.

Jovanovic, B. (2009). The technology cycle and inequality. *The Review of Economic Studies*, 76(2):707–729.

Jovanovic, B. and MacDonald, G. M. (1994). Competitive diffusion. *Journal of Political Economy*, 102(1):24–52.

Jovanovic, B. and Rob, R. (1989). The growth and diffusion of knowledge. *The Review of Economic Studies*, 56(4):569–582.

Keeling, M. J. and Eames, K. T. (2005). Networks and epidemic models. *Journal of the Royal Society Interface*, 2(4):295–307.

Khalil, H. K. (2002). *Nonlinear Systems*. Prentice Hall.

König, M., Tessone, C., and Zenou, Y. (2014). Nestedness in networks: A theoretical model and some applications. *Theoretical Economics*, 9:695–752.

König, M. D., Battiston, S., Napoletano, M., and Schweitzer, F. (2011). The efficiency and stability of R&D networks. *Games and Economic Behavior*, 75:694 –713.

Magrini, S., Gerolimetto, M., and Duran, H. E. (2013). Business cycle dynamics across the US states. *B.E. Journal of Macroeconomics*, 13:795–822.

Mahadev, N. and Peled, U. (1995). *Threshold Graphs and Related Topics*. North Holland.

Marjoram, P., Molitor, J., Plagnol, V., and Tavaré, S. (2003). Markov chain Monte Carlo without likelihoods. *Proceedings of the National Academy of Sciences*, 100(26):15324.

Miller, J. C. and Kiss, I. Z. (2014). Epidemic spread in networks: Existing methods and current challenges. *Mathematical modelling of natural phenomena*, 9(2):4.

Newman, M. (2010). *Networks: An Introduction*. Oxford University Press.

Peled, U. N., Petreschi, R., and Sterbini, A. (1999). (n, e)-graphs with maximum sum of squares of degrees. *Journal of Graph Theory*, 31(4):283–295.

Rogers, T., Clifford-Brown, W., Mills, C., and Galla, T. (2012). Stochastic oscillations of adaptive networks: application to epidemic modelling. *Journal of Statistical Mechanics: Theory and Experiment*, 2012(08):P08018.

Schilling, M. (2009). Understanding the alliance data. *Strategic Management Journal*, 30(3):233–260.

Sokal, A. D. (1996). *Monte Carlo Methods in Statistical Mechanics: Foundations and New Algorithms*.

Trajtenberg, M., Shiff, G., and Melamed, R. (2009). The "names game": Harnessing inventors, patent data for economic research. *Annals of Economics and Statistics*, (93/94):79–108.

Van Mieghem, P., Omic, J., and Kooij, R. (2009). Virus spread in networks. *IEEE/ACM Transactions on Networking*, 17(1):1–14.

Westbrock, B. (2010). Natural concentration in industrial research collaboration. *The RAND Journal of Economics*, 41(2):351–371.