Technology Cycles in Dynamic R&D Networks☆

Michael D. König[a]

[a]*Department of Economics, University of Zurich, Schönberggasse 1, CH 8001 Zurich, Switzerland.*

Abstract

In this paper we study the coevolutionary dynamics of knowledge creation, diffusion and the formation of R&D collaboration networks. Differently to previous works, knowledge is not treated as an abstract scalar variable but represented by a portfolio of ideas that changes over time through 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 analyze the behavior of this dynamic process and its convergence to a stationary state, in relation to the rates at which innovations and costly R&D collaboration opportunities arrive, and the rate of creative destruction leading to the obsolescence of existing technologies. We quantify the innovation gains from collaborations, and show that there exists a critical level for the technology learning success probability in collaborations below which an economy with weak in-house R&D capabilities does not innovate even in the presence of R&D collaborations. Moreover, we show that the interplay between knowledge diffusion and network formation can give rise to a cyclical pattern in the collaboration intensity, which can be described as a damped oscillation. We confirm this novel observation using an empirical sample of a large R&D collaboration network over the years 1985 to 2011. We then study the efficient network structure, compare it to the decentralized equilibrium structures generated, and design an optimal network policy to maximize welfare in the economy. Our efficiency analysis further allows us to study the effect of competition on innovation in R&D intensive industries where R&D collaborations between firms are commonly observed.

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

1. Introduction

R&D collaborations play an important role in the creation and diffusion of new technologies. Conversely, new technological opportunities impact the formation of R&D collaborations. In this paper we study the two-way influence of innovation, technology diffusion and R&D network formation.

This paper develops the first tractable model to study endogenously the 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. The technology portfolios change over time through innovation and knowledge spillovers from imitation and learning across

☆I would like to thank Andrea Montanari for numerous comments, guidance and support. Moreover, I would like to thank Kiminori Matsuyama, Chad Jones, Philippe Aghion, Matt Jackson, Francis Bloch, Armin Schmutzler, Frank Page, Tim Hellmann, Douglas White, Woody Powell, John Hagedoorn, Michelle Sovinsky, Yves Zenou, Fabrizio Zilibotti, Hans Gersbach, Maik Schneider and seminar participants at the University of Zurich, and the audiences of the 13th SAET Conference on Current Trends in Economics at MINES ParisTech in 2013 for helpful comments. Financial support from Swiss National Science Foundation through research grants PBEZP1–131169 and 100018_140266 is greatfully acknowledged. Finally I would like to thank SIEPR and the Department of Economics at Stanford University for their kind hospitality during 2010–2012 when parts of this paper have been written.

Email address: michael.koenig@econ.uzh.ch (Michael D. König)

collaborating firms. The growth of knowledge is thus an interactive process between innovation and imitation [cf. Jovanovic and Rob, 1989; König et al., 2012]. Moreover, some technologies can become obsolete [cf. Adams, 1990; Klette and Kortum, 2004], while R&D collaborations have a finite lifetime, are costly and their profitability is plagued with uncertainty [Harrigan, 1988; Kelly et al., 2002].

A key finding of the model is the existence of a threshold for the learning and imitation 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.[1] This 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]. We further analyze changes of the threshold with respect to various parameters of the model, and, in particular, find that competition lowers the threshold. 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. Moreover, the change in the threshold with the uncertainty in the profitability of R&D collaborations is non-monotonic. We further study the knowledge gains from R&D collaborations. Our results show that these are higher in the presence of competition. Moreover, we find that the gains are increasing with increasing in-house R&D capabilities, but only if these are below a threshold that depends on the knowledge obsolescence rate, and decreasing otherwise.

We then test some of the implications of the model using a large firm-level panel dataset on R&D collaborations over several decades and various sectors matched (partially) with patent data. 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 technology cycles, in a tractable framework that is also amenable to policy analysis. Our theoretical results further indicate that the cyclicality 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 form collaborations which allow them to get access to the technology. However, once the technology has sufficiently diffused trough the network, the market size shrinks, and so do the incentives to collaborate. As a result, the economy experiences periods of high collaborative activity followed by periods of low collaborative activity [cf. Matsuyama, 1999]. This has important policy implications. If policy makers want to increase welfare 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. Jovanovic and MacDonald, 1994; Jovanovic and Rob, 1990; Schumpeter, 1934].

We then investigate the efficient network structure, and compare it to the decentralized equilibrium [cf. e.g. König et al., 2014; Westbrock, 2010]. Our analysis indicates that equilibrium networks tend to be less centralized than the efficient structure, and in the empirical application of our model we quantify the welfare loss (by providing a lower bound) incurred by a suboptimal network structure. Moreover, we find that in the absence of competition the efficient network has a core periphery structure and can be characterized as a "nested split graphs" [cf. König et al., 2011; König et al.,

[1] Similarly, Kelly [2001, 2009] finds a threshold in a static spatial environment.

[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).

2013].[3] Our analysis further allows us to investigate the impact of competition on social welfare. We find that competition is welfare increasing. 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. Aghion et al., 2013; Kretschmer et al., 2012; Schmutzler, 2010], by focussing on R&D intensive industries where R&D collaborations between firms are commonly observed.

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]. However, only recently it has been recognized that the R&D network structure is highly unstable. 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 [Gay and Dousset, 2005; Hagedoorn and van Kranenburg, 2003]. Gulati et al. [2010] find a rise and fall of "small worlds" in the R&D alliance network over time.[4] A possible explanation for this phenomenon might be that in the pursue of complementary knowledge firms form small worlds initially but the excessive formation of these ties makes the information they gather redundant and ultimately leads to the decline of the small world [Hagedoorn and Frankort, 2008; Powell et al., 2005]. As Jovanovic and Rob [1989] put 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." Various empirical studies have also documented the convergence of firms' knowledge bases in sectors like electronics and the biotechnology industries [see e.g. Fai and Von Tunzelmann, 2001; Gambardella and Torrisi, 1998; Nesta and Dibiaggio, 2003; Patel and Pavitt, 1997]. In order to explain this phenomenon, in this paper we develop a tractable model in which firms experience decreasing returns from collaboration 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 the competition of firms across different sectors [cf. Matsuyama, 1999].[5]

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 [e.g. Aghion et al., 2013; Anderson and Tushman, 1990; Geroski and Walters, 1995; Goodwin, 1946; Jovanovic and Rob, 1990; Kuznets, 1940; Schumpeter, 1934].[6] For example, Andersen [1999] investigates the growth of different technological classes and identifies an S-shaped (Sigmoid curved) pattern over time [see also Griliches, 1957]. Technology cycles have been found in the income of patents as a function of age (see e.g. Giummo [2010] and Jovanovic [2009]) and the diffusion of chip technologies (Jovanovic and MacDonald [1994]). Cycles have also been found in firm R&D expenditures [Barlevy, 2007]. Franke [2001] studies oscillations in the growth rates of average productivity. Klepper [1996] shows that the

[3] 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 Mahadev and Peled, 1995].

[4] A small world network is characterized by high clustering and a short average path length between the nodes in the network [Watts and Strogatz, 1998]. It has been argued that such small worlds are advantageous in generating and diffusing innovations in networks [cf. Cowan and Jonard, 2004].

[5] In particular, a purely technology based explanation as e.g. conjectured in Hagedoorn and Frankort [2008] for these cycles ignoring market and competition effects does not seem to be sufficient.

[6] Business cycles are another prominent instance for the unstable and periodic patterns that can be observed in economic activity [Desai and Ormerod, 1998; Goodwin, 1951; Kaldor, 1940], and technological change has been one of the explanations for their occurrence [Galí, 1999; Holly and Petrella, 2012]. Here we find, both empirically and theoretically, that technological development in R&D networks can follow a cyclical pattern. These cycles could be one of the mechanisms that trigger business cycles.

number of firms and the rate and diversity of product innovation eventually decline along a product life cycle. Cyclical patterns have further been observed empirically in mergers and acquisitions [Golbe and White, 1993], and joint ventures [Gomes-Casseres, 2002]. 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 [e.g. Amir et al., 2003; Amir and Wooders, 2000; Bloch, 1995, 1997; D'Aspremont and Jacquemin, 1988]. For example, Dawid and Hellmann [2014]; Goyal and Joshi [2003] have investigated the formation of networks of R&D collaborating firms in which firms can share knowledge of a cost reducing technology. König et al. [2011]; 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. 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. A key consequence is that in these models larger firms 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, where the small firm possess knowledge of a key technology that is particularly valuable to the larger firm [cf. Powell et al., 2005]. Here we propose a model in which even large firms have incentives to collaborate with smaller ones when these hold some technologies of interest.

Another strand of literature has studied the process of knowledge diffusion in an exogenously given communication or social network. In the mathematics, epidemiology, computer science and physics literature the spread of epidemics on networks has been extensively studied [see e.g. Acemoglu et al., 2011; Anderson et al., 1992; Berger et al., 2005; Chatterjee and Durrett, 2009; Pastor-Satorras and Vespignani, 2001; Van Mieghem et al., 2009; Wang et al., 2003].[7,8] 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. Meagher and Rogers [2004] and Andergassen et al. [2006] study the process of innovation and knowledge diffusion on an exogenous network structure. Further examples include Lopez-Pintado [2008] and Montanari and Saberi [2010]. In particular, Montanari and Saberi [2010] investigate the speed of diffusion of innovations in a network in relation to certain topological characteristics of the network. These works, however, do not explain the network structure but take it as exogenously given. We improve on them, by analyzing the endogenous formation of networks in which innovation and knowledge diffusion takes place. While the above mentioned literature finds that it is typically the largest eigenvalue of the (adjacency matrix associated with the) network which determines a threshold below which epidemics do not spread [see e.g. Newman, 2010], 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] and Blume et al. [2011]. However, these papers do not take into account the incentives of agents to form links. For example, Gross et al. [2006] 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

[7]See also Chapter 17 in Newman [2010] for an overview and introduction.

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

the network is formed according to a specific random process. Similarly, 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. As in Blume et al. [2011] and Gross et al. [2006], the link creation and removal process is mechanistic, and does not depend on the marginal payoffs agents receive from forming or severing links. In our model link formation is based on a standard 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.[9] Moreover, none of these papers is applied to the current context of R&D collaborations, nor has an empirical application.

Only few studies analyze the interplay between knowledge creation, diffusion and network evolution. Most notably, Baum et al. [2010]; Berliant and Fujita [2008, 2009, 2011]; Cowan and Jonard [2004, 2008] 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 either abstract away from the network structure of collaborations, or they are based on numerical simulation studies 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.

Outline of the paper. The paper is organized as follows. In Section 2 we provide an empirical motivation for our analysis by investigating the interaction between the patent portfolios of firms and the formation of R&D collaborations, and the average number of collaborations in a panel of a real world R&D network. In Section 3 we introduce the model, Section 4 describes the innovation process, and Section 5 the formation of the network. Section 6 describes the coevolution of the technology portfolios and the network, while Section 7 investigates the coevolution of the knowledge stocks and the network. The equilibrium analysis is given in Section 8, while Section 9 analyzes efficiency. An empirical application and a calibration of the model's parameters is provided in Section 10. Section 11 concludes. All proofs are relegated to Appendix B.

2. Empirical R&D Networks

Data. To motivate our model we consider a sample of R&D alliances ranging over the years 1985 to 2011. The data stems from the Thomson SDC alliance database.[10] 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 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].[11]

---

[9] In the language of statistical mechanics our generalization of e.g. Gross et al. [2006] is similar to the generalization of the n-vector model over the classic Ising model [cf. e.g. Grimmett, 2010; Stanley, 1968].

[10] For an overview and comparison of different types of R&D alliance data sets see Schilling [2009].

[11] 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] finding that the normal life span of most alliances is no more than five years. Moreover, Harrigan [1988] studies 895 alliances from 1924 to 1985 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

From our sample of 21,478 firms we could further 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.

Micro-level evidence. 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. Let $h_{ik,t} \in \{0,1\}$, $i = 1,\ldots,n$, $k = 1,\ldots,N$ be the indicator variable of firm i indicating whether it possesses the technology k at time t, and let $\mathbf{h}_{it}$ denote the vector of technologies of firm i describing the patent portfolio of firm i. Let the support of $\mathbf{h}_{it}$ be given by $\mathsf{S}(\mathbf{h}_{it})$ and its cardinality given by $|\mathsf{S}(\mathbf{h}_{it})| = \langle \mathbf{h}_{it}, \mathbf{u}\rangle$, where $\mathbf{u}$ is a vector of ones and $\langle\cdot,\cdot\rangle$ is the usual scalar product in $\mathbb{R}^N$. In other words, $|\mathsf{S}(\mathbf{h}_{it})|$ counts the number of technologies known to i. Moreover, the technologies j possesses but i does not, can be written as $\langle \mathbf{h}_{it}^c, \mathbf{h}_{jt}\rangle$, while the technologies i possesses but j does not, is $\langle \mathbf{h}_{jt}^c, \mathbf{h}_{it}\rangle$. Further, let $a_{ij,t} \in \{0,1\}$ be the indicator for whether firms i and j have an R&D collaboration at time t.

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 non-linear regression model[12]

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

to 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 spillover coefficient, $\hat{\alpha}_1$, is highly significant. This illustrates the importance of R&D collaborations for the diffusion of ideas across firms [cf. Jovanovic and MacDonald, 1994].

In order to illustrate the second effect, i.e. the impact of the technology portfolios on R&D collaborations, we estimate the following the nonlinear regression model

$$\mathbb{P}\left(a_{ij,t} = 1\right) = \frac{e^{\beta_0 + \beta_1 \langle \mathbf{h}_{it}^c, \mathbf{h}_{jt}\rangle}}{1 + e^{\beta_0 + \beta_1 \langle \mathbf{h}_{it}^c, \mathbf{h}_{jt}\rangle}} \frac{e^{\beta_0 + \beta_1 \langle \mathbf{h}_{jt}^c, \mathbf{h}_{it}\rangle}}{1 + e^{\beta_0 + \beta_1 \langle \mathbf{h}_{jt}^c, \mathbf{h}_{it}\rangle}}, \tag{2}$$

to 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, $\langle \mathbf{h}_{it}^c, \mathbf{h}_{jt}\rangle$, and i possesses but j does not, $\langle \mathbf{h}_{jt}^c, \mathbf{h}_{it}\rangle$, on their propensity to form a collaboration. This indicates the importance of complementarity in the technology portfolios for the creation of R&D collaborations [cf. Jovanovic and Rob, 1989]. The above exploratory empirical results will serve as a motivation for our general model introduced in Sections 4 and 5.

last less than 10 years (2/3 of the total), the average lifetime turns out to be 3.9 years.

[12] Let $Y_{ik,t} \in \mathbb{N}$ denote the count variable for the number of patents of firm i in technology class k at time t. Moreover, let $h_{ik,t} = \mathbb{1}_{\{Y_{ik,t} > 0\}}$ be the indicator variable whether firm i has a patent in technology class k at time t. If $Y_{ik,t} \sim \text{Pois}(\lambda)$ with rate λ then $\mathbb{P}(h_{ik,t} = 1) = 1 - e^{-\lambda}$. Conversely, it follows that $\lambda = -\ln\left(1 - \mathbb{P}(h_{ik,t} = 1)\right)$.

Table 1: Estimated coefficients with their standard deviations, t-statistics and p-values for the regression model of Equation (3).

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.

Macro-level evidence. Exemplary networks for the years 1985, 1990, 1995, 2000, 2005 and 2010 can be seen in Figure 1. The figure demonstrates that the evolution of the network is highly non-stationary. The average number of collaborations, $\bar{d}_t$, per year t is shown in Figure 2. The figure demonstrates 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.

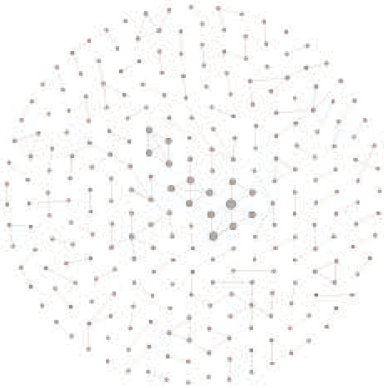
In order to investigate the non-stationary (oscillatory) pattern in the network data at the aggregate level, we perform an estimation procedure similar to Golbe and White [1993] to test whether the R&D collaboration intensity, as measured by the average number of collaborations $\bar{d}_t$ in a given year t, shows a cyclical trend. For this purpose we estimate a regression model of the form[13]

$$y_t = \frac{1}{2}a_0 + \sum_{j=1}^{3} \left(a_j \cos\left(\omega_j t\right) + b_j \sin\left(\omega_j t\right)\right), \quad t \in [0, T], \tag{3}$$

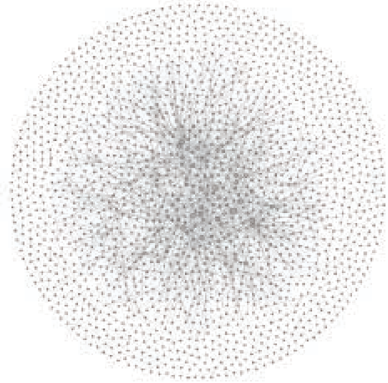
with the parameters $a_0, a_j, b_j,\ j = 1, \ldots, 3$ (Fourier coefficients), where the angular frequency is given by $\omega_j \equiv \frac{2\pi}{T} j$, and y_t is periodic with period T. The results of this regression can be seen in Table 2. The table shows that all coefficients are statistically significant. Following Golbe and White [1993] we take this as an indicator for the presence of a cyclical pattern in the data.[14] From the fact that the amplitude of the cycle is decreasing over time we conclude that the average degree follows a "damped oscillation". In the next section we develop a model that can generate such a cyclical pattern in the R&D collaboration activities of the firms.

[13]Equation (3) is a Fourier series representation of the time series $\{y_t\}_{t\in[0,T]}$. See e.g. Hamilton [1994, Chapter 6.2] for further details.

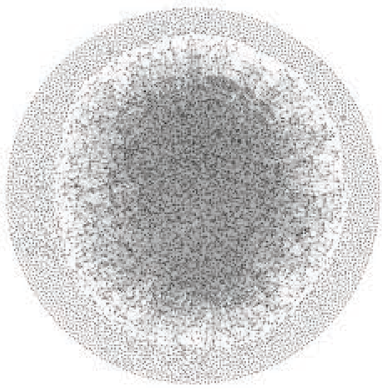
[14]Note that Hagedoorn and van Kranenburg [2003] investigate the presence of cycles in a different, smaller data set, over the years 1960–1998. They find evidence for the non-stationary nature of the collaboration intensity but no conclusive evidence for an oscillatory pattern. Despite the different way in which their data was collected we suspect that this is due to the fact that they do not explicitly look at the average degree of firms in the R&D network, their data set covers a shorter period of observation than ours, is much smaller and therefore potentially more noisy.



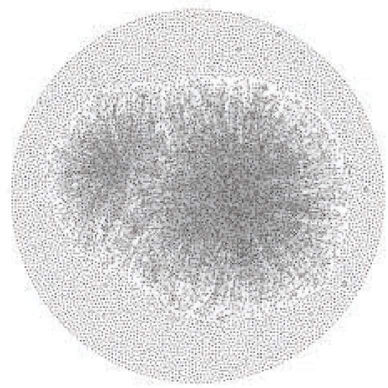
(a) 1985: $n = 263$, $m = 171$.



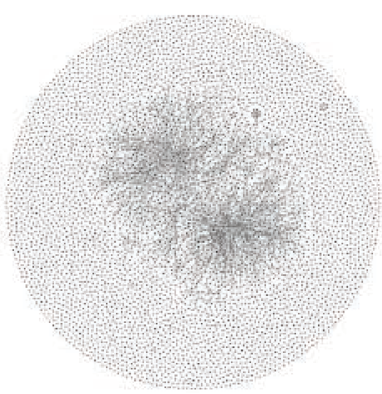
(b) 1990: $n = 1815$, $m = 1970$.



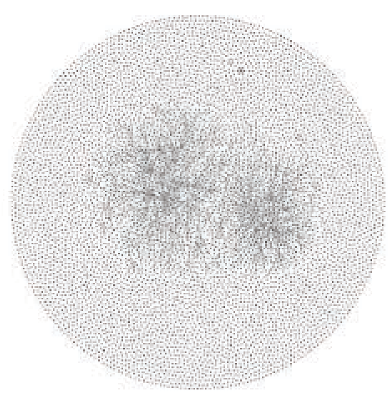
(c) 1995: $n = 9572$, $m = 13977$.



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



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



(f) 2010: $n = 6683$, $m = 5032$.

Figure 1: Network snapshots for the years (a) 1985, (b) 1990, (c) 1995, (d) 2000, (e) 2005 and (f) 2010. A node's size indicates its eigenvector centrality.

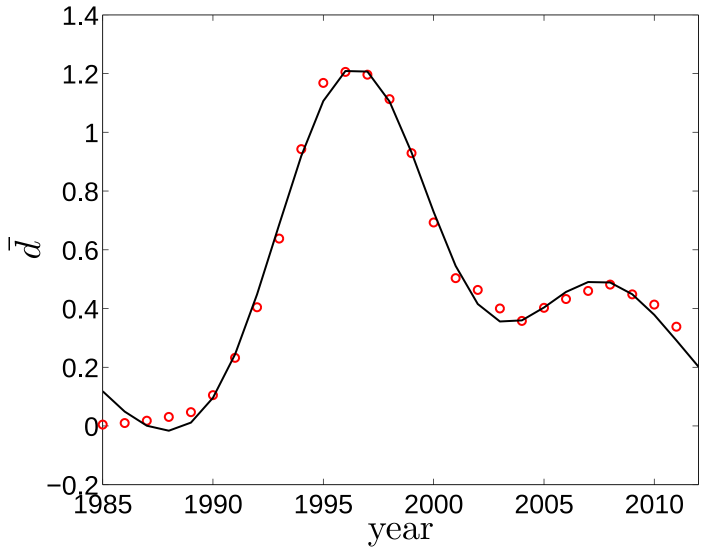


Figure 2: The average degree $\bar{d}$ over the years 1985 to 2012. The circles indicate the empirical observations while the curve indicates the function in Equation (3) with the parameter estimates from Table 2.

3. The Model

We consider a Schumpeterian model of monopolistic competition as in e.g. Acemoglu et al. [2006]; Aghion and Howitt [2009]; König et al. [2012]. 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

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

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, as an input to R&D and also as an input to the production of intermediate goods. The profit maximization program yields the following inverse demand function for intermediate goods,

$$p_i = \left(\frac{A_i}{x_i}\right)^{1-\alpha},$$

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, where $1 < \phi \leq 1/\alpha$. The firm sets the price equal to the unit cost, $p_i = \phi$, and sells at that price the equilibrium quantity $x_i = \phi^{-\frac{1}{1-\alpha}} A_i$. The gross profit earned by a firm i in an intermediate sector i, not taking into account any R&D collaboration costs, will then be a linear function of its productivity

$$\tilde{\pi}_i = (p_i - 1)\, x_i = \psi A_i, \tag{4}$$

where $\psi = \frac{\phi-1}{\alpha}\phi^{-\frac{1}{1-\alpha}}$ which is monotonically increasing in α and decreasing in ϕ. In equilibrium, output is proportional to aggregate productivity as follows $Y = \frac{1}{\alpha}\phi^{-\frac{\alpha}{1-\alpha}} \sum_{i=1}^{n} A_i = \frac{1}{\alpha}\phi^{-\frac{\alpha}{1-\alpha}} A$, where aggregate productivity is $A = \sum_{i=1}^{n} A_i$,

Next, we consider the same economic environment as above, but now assume that a firm can produce more than one intermediate good by introducing multiproduct firms [cf. Bernard et al., 2011]. In this setup firms can produce new varieties in different sectors by applying their technological knowledge in all sectors, similar to variety expanding models [cf. Jones, 1995, 2005]. More precisely, 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 the contest success function [Corchón, 2007; Fullerton and McAfee, 1999; Meland and Straume, 2007; Tullock, 1980]:[15]

$$\mathbb{P}(\text{firm } i \text{ produces in sector } j) = \frac{A_i}{\sum_{k=1}^{n} A_k}. \tag{5}$$

If firm i becomes the producer, it earns a gross profit of $\tilde{\pi}_i = \psi A_i$. The contests in each sector are assumed to be independent. Firm i's expected gross profit from all n sectors is then given by[16]

$$\tilde{\pi}_i = \psi A_i n \frac{A_i}{\sum_{j=1}^{n} A_j} = \psi \frac{A_i^2}{\frac{1}{n}\sum_{j=1}^{n} A_j}. \tag{6}$$

$$\pi_i = \tilde{\pi} - cd_i = \theta A_i + (1-\theta)\frac{A_i^2}{\frac{1}{n}\sum_{j=1}^{n} A_j} - cd_i, \tag{7}$$

where $\theta \in \{0, 1\}$ is a (zero/one) competition parameter,[17] we have normalized $\psi = 1$, $c \in \mathbb{R}_+$ 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 of size n.[18]

We assume that the productivity A_i of firm i is a linear function of the number of technologies (size of the technology portfolio) owned by the firm [cf. Klette and Kortum, 2004]. 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. Then we assume that

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

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 $\mathbb{R}^N$.

[15] Alternatively, we could assume that a firm can win a patent in a patent race that allows it to produce the intermediate good variety [cf. Futia, 1980; Reinganum, 1985], with a duration of one unit of time, and that the good becomes obsolete after the expiration of the patent.

[16] 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 $\sum_{j=1}^{n} \binom{n}{j} p_i^j (1-p_i)^{n-j} = np_i$.

[17] Note that θ is a measure for monopoly power: $\theta = 1$ indicates local monopolists in each sector, while $\theta = 0$ indicates competition across different sectors. An alternative interpretation – which we do not emphasize here – is that θ measures the scope and generality of the technologies used by the firms, i.e. the extent to which firms are using general purpose technologies (GPT) [cf. Aghion et al., 2013; Jovanovic and Rousseau, 2005].

[18] As in Berliant and Fujita [2008]; Dawid and Hellmann [2014]; König et al. [2011]; König et al. [2011]; Roketskiy [2011]; Westbrock [2010] we do not explicitly incorporate the firm's R&D expenditure decision in the productivity, and consequently firms' profits in Equation (8). Instead we focus on the strategic choice of a firm's collaboration partners. We can thus view Equation (8) as a reduced form that allows us to study the coevolution of the network and knowledge portfolios emanating from the firms' strategic linking decisions. Moreover, as in the models analyzed in König et al. [2012]; Lucas and Moll [2011]; Perla and Tonetti [2012] firms can choose between in-house R&D and copying another firm's technology by forming (or not) an R&D collaboration. In particular, in our model firms which do not participate in R&D collaborations innovate through in-house R&D instead.

Normalizing also $a = 1$, the profit function from Equation (7) then becomes

$$\pi_i(\mathbf{h}) = \theta(1 + b|\mathsf{S}(\mathbf{h}_i)|) + (1-\theta)\frac{1 + 2b|\mathsf{S}(\mathbf{h}_i)| + b^2|\mathsf{S}(\mathbf{h}_i)|^2}{1 + b\frac{1}{n}\sum_{j=1}^n |\mathsf{S}(\mathbf{h}_j)|} - cd_i, \quad \theta \in \{0,1\}, \tag{8}$$

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 $i \in \mathcal{N}$.

In the following sections we consider a dynamic environment, where in every period an existing final good is replaced with a new one. In the non-competitive case ($\theta = 1$) the final good uses inputs from the same intermediate goods producing firms at every period. In contrast, in the competitive case ($\theta = 0$), the firms supplying the intermediate goods are redrawn every period according to the contest success function in Equation (5).[19] The probability with which a firm becomes the producer of a variety depends on its productivity relative to the aggregate productivity of all other firms in the economy, and these productivities change over time through innovation or learning other firms's technologies in R&D collaborations.

4. Innovation, Spillovers and Marginal Profits from Collaboration

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 knowledge arrives as a Poisson process with an innovation rate that depends on the stock of knowledge of the firm [cf. Dasgupta and Stiglitz, 1980, 1981; Klette and Kortum, 2004; Loury, 1979]. We also allow for spillovers between collaborating firms such that the rate with which a firm i makes an innovation in the knowledge category k 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, if it does not know it already, at a rate[20]

$$\nu_{ik,t} = \underbrace{\gamma + \alpha\sum_{l=1}^N h_{il,t}}_{\text{innovation}} + \underbrace{\beta\sum_{j=1}^n a_{ij,t}h_{jk,t}}_{\text{imitation}}. \tag{9}$$

With rate λ, each knowledge category can also become obsolete [cf. e.g. Adams, 1990; Andergassen et al., 2006; Caballero and Jaffe, 1993; Klette and Kortum, 2004].[21]

In the following we assume that collaborative R&D agreements between firms have only a finite lifetime. The fact that collaborations do not last forever is a quite natural feature of real-world networks. Ehrhardt et al. [2008] put forward inter-firm alliances and scientific collaborations as examples of networks in a volatile environment. For inter-firm alliances, Hagedoorn [2002] for research partnerships, Kogut et al. [2007] for joint ventures, Harrigan [1988] for alliances and Park and Russo [1996] for (equity-based) joint ventures provide empirical evidence on this phenomenon. For example, Harrigan [1988] studies 895 alliances from 1924 to 1985 and concludes that the average life-span of

---

[19] The competitive case builds on 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]). In this model the probability of a firm i to succeed with innovating in period t is given by $\mathbb{P}(\text{firm } i \text{ succeeds at time } t) = \frac{A_{it}}{\sum_{j=1}^n A_{jt}}$, where A_{it} is the productivity of firm i at time t. This functional form is also known as a Tullock contest success function [cf. Baye and Hoppe, 2003; Tullock, 1980]. For a stochastic derivation see Jia [2008].

[20] In models such as the one by Klette and Kortum [2004], where the firm's R&D expenditure choice is explicitly considered, one finds that a firm scales up its R&D expenditure (and its innovation probability) in proportion to its knowledge capital. Hence, we can view Equation (9) as a reduced form where the innovation rate increases with the stock of knowledge of the firm and spillovers from collaborating firms.

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

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.

Let $\tau > 0$ denote the expected duration of a collaborative R&D agreement. When evaluating a potential collaboration, a firm i computes its discounted profit at time $t+\tau$,[22] taking the current network G_t as given [cf. Jackson and Watts, 2002],[23] while discounting future profits at a rate $\delta \equiv \frac{1}{1+r} > 0$. The firm i's present discounted profit at time t is then given by

$$V_i(\mathbf{h}_t, G_t) = \pi_i(\mathbf{h}_t) + \frac{1}{1+r}\mathbb{E}_t\left(\pi_i(\mathbf{h}_{t+\tau})|\mathbf{h}_t, G_t\right).$$

In the following we 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.[24] Then we can derive the change in the present discounted profits of a firm from forming a collaboration as follows.

Proposition 1. *The change in the present discounted profit of the firm i from forming the link ij for $\theta \in \{0,1\}$ can be written as*

$$V_i(\mathbf{h}_t, G_t + ij) - V_i(\mathbf{h}_t, G_t) = \beta\tau\delta\left(\theta b + (1-\theta)\frac{2b(1+b|\mathsf{S}(\mathbf{h}_{i,t})|)}{1+b\bar{h}_t(\mathbf{h}_t, G_t)}\right)\langle \mathbf{h}^c_{it}, \mathbf{h}_{jt}\rangle - \delta c + O\left(\frac{\tau}{n}\right),$$

where $\bar{h}_t \equiv \frac{1}{n}\sum_{i=1}^n |\mathsf{S}(\mathbf{h}_{i,t})|$ denotes the average stock of knowledge at time t.

In the following we denoted by

$$g_{\theta,\tau}(\bar{h}_t) \equiv \tau\delta b\left(\theta + 2\frac{1-\theta}{1+b\bar{h}_t}\right). \tag{10}$$

Note that $g_{\theta,\tau}(\bar{h}_t)$ is decreasing with the average knowledge stock $\bar{h}_t$. Dropping the remainder term $O\left(\frac{\tau}{n}\right)$ in Proposition 1 we then get

$$\frac{V_i(\mathbf{h}_t, G_t + ij) - V_i(\mathbf{h}_t, G_t)}{\delta} \approx \beta g_{\theta,\tau}(\bar{h}_t)(1+b|\mathsf{S}(\mathbf{h}_{i,t})|)^{1-\theta}\langle \mathbf{h}^c_{it}, \mathbf{h}_{jt}\rangle - c. \tag{11}$$

From the above expression we find that marginal profits for firm i from forming a link ij are increasing

[22] Related to our setup, Sannikov and Skrzypacz [2010] consider a dynamic game of strategic interaction where players learn information continuously over time, but take actions only at discrete points in time.

[23] This assumption is reminiscent of myopic behavior and is common in the complex strategic environment that networks represent. 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 cooperations we think that this assumption is not too restrictive.

[24] This assumption also guarantees that firms need only limited information about the knowledge portfolios of other firms. In particular a firm needs only to know the knowledge portfolios of its alliance partners (but not of any other firm), and the total average portfolio size (e.g. from some statistic of the aggregate innovativeness of the economy), in order to compute its present discounted profit. Any higher order corrections would require information about the knowledge possessed by the neighbors' neighbors, their neighbors, and so on. Since this information is hard to obtain (firms typically do not make their R&D programs public), this would be a strong assumption. Moreover, there exists empirical evidence that agents in a network use only information limited to their immediate neighborhood when deciding about their actions [e.g. Friedkin, 1983]. Related theoretical models of networked agents with limited information sets are DeMarzo et al. [2003]; Golub and Jackson [2012]; Jackson and Golub [2010], Montanari and Saberi [2010] and König [2011].

in the number of ideas that firm j has but firm i does not have, $\langle \mathbf{h}_{it}^c, \mathbf{h}_{jt}\rangle$, and - in the competitive case - the total stock of knowledge possessed by firm i, $|\,\mathsf{S}(\mathbf{h}_{i,t})|$, relative to the average stock of knowledge $\bar{h}_t$. In the first case (without competition, $\theta = 1$), 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 competition, $\theta = 0$) shows that relatively more competitive firms are better in reaping the gains from getting access to complementary knowledge than less competitive 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 the present discounted profits of Equation (11).

5. Innovation and R&D Network Formation

In the following we introduce more formally the knowledge creation, diffusion and R&D network formation process.

Definition 1. *Consider a population of firms $\mathcal{N} = \{1, \ldots, n\}$. Each firm $i \in \mathcal{N}$ is equipped with a knowledge vector $\mathbf{h}_{it} \in \mathcal{H}^N = \{0,1\}^N$, in a network $G_t = (\mathcal{N}, \mathcal{E}_t) \in \mathcal{G}_n$, $\mathcal{E}_t \in \mathcal{N} \times \mathcal{N}$, at time $t \in \mathbb{R}_+$. Denote by $\mathbf{h}_t$ the $n \times N$ matrix with rows equal to $\mathbf{h}_{it}$ for each $i = 1, \ldots, n$. We consider the continuous time Markov process $(\mathbf{h}_t, G_t)_{t \in \mathbb{R}_+}$, in which the following events happen in a small time interval $[t, t + \Delta t)$, $\Delta t \geq 0$:*

Innovation: *Each firm $i \in \mathcal{N}$ discovers knowledge category $k = 1, \ldots, N$ at a rate*

$$\nu_{ik,t} = \gamma + \alpha \sum_{l=1}^{N} h_{il,t} + \beta \sum_{j=1}^{n} a_{ij,t} h_{jk,t},$$

and the probability that firm i discovers idea j in the time interval $[t, t + \Delta t)$ is given by

$$\mathbb{P}(h_{ik,t+\Delta t} = 1 | h_{ik,t} = 0, \mathbf{h}_t, G_t) = 1 - e^{-\nu_{ik,t}\Delta t} = \nu_{ik,t}\Delta t + o(\Delta t). \tag{12}$$

Knowledge Obsolescence: *At rate $\lambda \geq 0$, each idea $k = 1 \ldots, N$ becomes obsolescent. That is, the probability that idea j becomes obsolescent in the time interval $[t, t + \Delta t)$ is given by*

$$\mathbb{P}(h_{ik,t+\Delta t} = 0 | h_{ik,t} = 1, \mathbf{h}_t, G_t) = 1 - e^{-\lambda\Delta t} = \lambda\Delta t + o(\Delta t). \tag{13}$$

Link Creation: *Each (unordered) pair of firms $i, j \in \mathcal{N} \times \mathcal{N}$ receives an opportunity to create the link ij with rate $\rho \geq 0$. If the pair i, j receives such an opportunity, then the link ij is created, if it is not present, with probability*[25]

$$\begin{aligned}
&\mathbb{P}(G_{t+\Delta t} = G_t + ij | \mathbf{h}_t, G_t) \\
&\quad = \rho\,\mathbb{P}(\{V_i(\mathbf{h}_t, G_t + ij) + \varepsilon_{it} > V_i(\mathbf{h}_t, G_t)\} \cap \{V_j(\mathbf{h}_t, G_t + ij) + \varepsilon_{jt} > V_j(\mathbf{h}_t, G_t)\}) \\
&= \rho \frac{e^{\eta(\beta g_{\theta,\tau}(\bar{h}_t)(1+b|\,\mathsf{S}(\mathbf{h}_{i,t})|)^{1-\theta}\langle \mathbf{h}_{it}^c, \mathbf{h}_{jt}\rangle - c)}}{1 + e^{\eta(\beta g_{\theta,\tau}(\bar{h}_t)(1+b|\,\mathsf{S}(\mathbf{h}_{i,t})|)^{1-\theta}\langle \mathbf{h}_{it}^c, \mathbf{h}_{jt}\rangle - c)}} \frac{e^{\eta(\beta g_{\theta,\tau}(\bar{h}_t)(1+b|\,\mathsf{S}(\mathbf{h}_{j,t})|)^{1-\theta}\langle \mathbf{h}_{jt}^c, \mathbf{h}_{it}\rangle - c)}}{1 + e^{\eta(\beta g_{\theta,\tau}(\bar{h}_t)(1+b|\,\mathsf{S}(\mathbf{h}_{j,t})|)^{1-\theta}\langle \mathbf{h}_{jt}^c, \mathbf{h}_{it}\rangle - c)}} \Delta t + o(\Delta t),
\end{aligned} \tag{14}$$

[25] The probability that firms i and j form an R&D collaboration depends on both finding a collaboration profitable, that is, it must hold that both $V_i(\mathbf{h}_t, G_t + ij) + \varepsilon_{it} > V_i(\mathbf{h}_t, G_t)$ and $V_j(\mathbf{h}_t, G_t + ij) + \varepsilon_{jt} > V_j(\mathbf{h}_t, G_t)$.

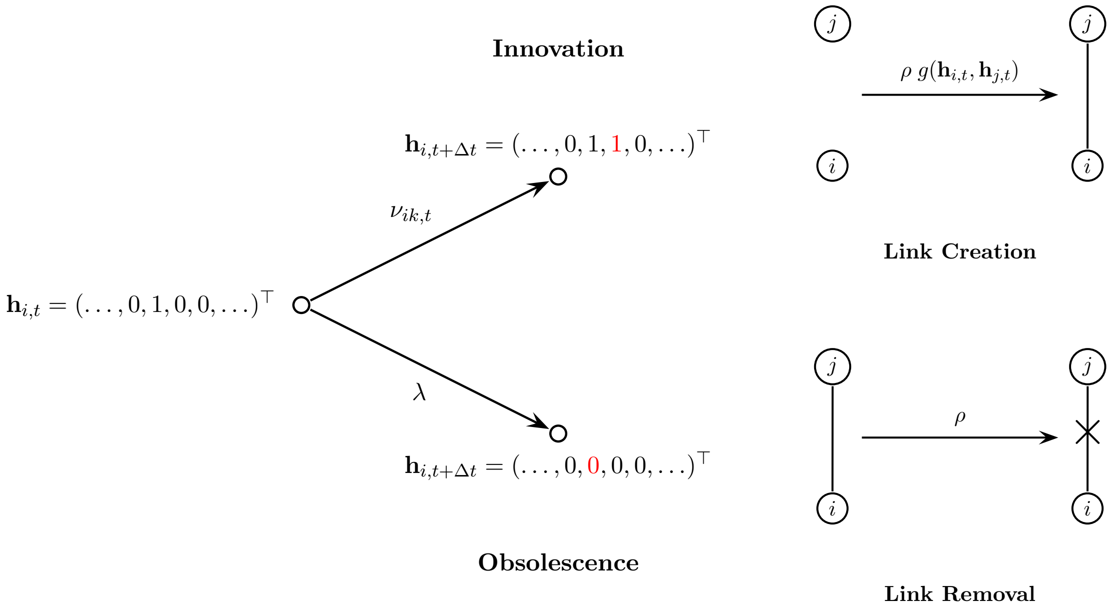


Figure 3: Illustration of the network formation and innovation process outlined in Definition 1.

where we have assumed that profits from forming a link are perturbed by identically and independently logistically distributed error terms $\varepsilon_{it}, \varepsilon_{jt}$ with parameter η/δ.

Link Removal: *An existing link ij is removed when the collaboration between i and j expires. This happens at a rate $\rho = \rho_0 + 1/\tau$, so that*

$$\mathbb{P}(G_{t+\Delta t} = G_t - ij | \mathbf{h}_t, G_t) = \rho \Delta t + o(\Delta t). \tag{15}$$

In the following we assume that the link creation and link removal rates are identical and given by ρ. This entails no loss of generality but helps us to simplify our notation. An illustration of the stochastic process introduced in Definition 1 is shown in Figure 3. Further note that the introduction of noise in marginal profits from collaboration leading to Equation (14) is quite natural, as the establishment of an R&D collaboration is fraught with ambiguity and uncertainty [cf. Kelly et al., 2002]. Moreover, Podolny and Page [1998] document that many collaborations fail and are terminated early. We allow for this possibility by including the term ρ_0 in the rate ρ adding to the inverse of the expected duration τ of a collaboration.

Note that without knowledge obsolescence (i.e. when we set $\lambda = 0$) the firms' technology portfolios eventually become complete, and there would be no incentives to form collaborations any more so that the network would be empty. Moreover, without link removal (i.e. when we set $\rho = 0$) the network would eventually become complete.[26] Both of these extreme cases are at odds with the R&D network structures that we observe in the data.

In the next section we analyze the evolution of the number of firms with different technology portfolios and the number of links between them.

[26] 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 SIS model for epidemics spreading on a static network.

6. Coevolution of Knowledge Portfolios and the Network

From Definition 1 describing the dynamic process of network formation and knowledge diffusion it is possible to obtain a coupled system of ordinary differential equations that completely describes the evolution of the average number of firms with a certain knowledge portfolio and the probability of a link between any pair of firms with given knowledge 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* $\xi_t(\mathbf{h},\mathbf{h}') \equiv \mathbb{P}(a_{ij,t}=1|\mathbf{h}_{it}=\mathbf{h},\mathbf{h}_{jt}=\mathbf{h}')$*, and let the fraction of firms with knowledge vector* $\mathbf{h}$ *be* $x_t(\mathbf{h}) \equiv \mathbb{P}(\mathbf{h}_{it}=\mathbf{h})$*. Introduce the rescaled parameters* $\beta \to \beta/n$*,* $c \to c/n$*,* $\eta \to \eta n$*,*[27] *and denote by*

$$g(\mathbf{h},\mathbf{h}') \equiv \frac{e^{\eta(\beta g_{\theta,\tau}(\bar{h}_t)(1+b|\mathsf{S}(\mathbf{h})|)^{1-\theta}\langle \mathbf{h}^c,\mathbf{h}'\rangle - c)}}{1+e^{\eta(\beta g_{\theta,\tau}(x_t)(1+b|\mathsf{S}(\mathbf{h})|)^{1-\theta}\langle \mathbf{h}^c,\mathbf{h}'\rangle - c)}} \frac{e^{\eta(\beta g_{\theta,\tau}(\bar{h}_t)(1+b|\mathsf{S}(\mathbf{h}')|)^{1-\theta}\langle \mathbf{h}'^c,\mathbf{h}\rangle - c)}}{1+e^{\eta(\beta g_{\theta,\tau}(x_t)(1+b|\mathsf{S}(\mathbf{h}')|)^{1-\theta}\langle \mathbf{h}'^c,\mathbf{h}\rangle - c)}}. \tag{16}$$

Then, in the limit of large n*,* $x_t(\mathbf{h})$ *converges in probability to the solution of the system of ODEs*

$$\begin{aligned}\frac{dx_t(\mathbf{h})}{dt} &= (\gamma+\alpha(|\mathsf{S}(\mathbf{h})|-1))\sum_{k\in\mathsf{S}(\mathbf{h})} x_t(\mathbf{h}-\mathbf{e}_k) + \lambda \sum_{k\in\mathsf{S}(\mathbf{h}^c)} x_t(\mathbf{h}+\mathbf{e}_k)\\ &\quad - (\lambda|\mathsf{S}(\mathbf{h})| + \gamma|\mathsf{S}(\mathbf{h}^c)| + \alpha|\mathsf{S}(\mathbf{h})||\mathsf{S}(\mathbf{h}^c)|)\, x_t(\mathbf{h})\\ &\quad + \beta \sum_{k\in\mathsf{S}(\mathbf{h})} \sum_{\mathbf{h}'\in\mathcal{H}^N: h'_k=1} \xi_t(\mathbf{h}-\mathbf{e}_k,\mathbf{h}')x_t(\mathbf{h}-\mathbf{e}_k)x_t(\mathbf{h}')\\ &\quad - \beta \sum_{k\in\mathsf{S}(\mathbf{h}^c)} \sum_{\mathbf{h}'\in\mathcal{H}^N: h'_k=1} \xi_t(\mathbf{h},\mathbf{h}')x_t(\mathbf{h})x_t(\mathbf{h}'),\end{aligned} \tag{17}$$

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* $\mathbf{h}$*,*[28] *and* $\xi_t(\mathbf{h},\mathbf{h}')$ *converges in probability to the solution of the system of ODEs*

$$\frac{d\xi_t(\mathbf{h},\mathbf{h}')}{dt} = \rho\, g(\mathbf{h},\mathbf{h}') - \rho\, \left(1+g(\mathbf{h},\mathbf{h}')\right)\xi_t(\mathbf{h},\mathbf{h}') + o(\rho). \tag{18}$$

With the dynamics of $x_t(\mathbf{h})$ and $\xi_t(\mathbf{h},\mathbf{h}')$ known from Theorem 1, the current period 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 current period average degree is $\bar{d}_t = n\sum_{\mathbf{h},\mathbf{h}'\in\mathcal{H}^N} \xi_t(\mathbf{h},\mathbf{h}')x_t(\mathbf{h})x_t(\mathbf{h}')$. 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 (18). The differential equations governing the dynamics of the expected number of links becomes considerably more involved, and can only be derived by making a *pair approximation*:[29] 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 $\tau_t(\mathbf{h},\mathbf{h}',\mathbf{h}'')$ denote the expected number of triplets with a firm with technology $\mathbf{h}$ being

[27]The assumption of the technology spillover parameter being given by β/n ensures that the contribution to the total spillovers from a single firm in a dense network is $O(1)$ [see e.g. Levin et al., 2009, Chap. 15.2].

[28]The support of $\mathbf{h}$ is $\mathsf{S}(\mathbf{h})$ and its cardinality is $|\mathsf{S}(\mathbf{h})| = \langle \mathbf{h},\mathbf{u}\rangle$, counting the number of nonzero entries in $\mathbf{h}$, with $\mathbf{u}$ being a vector of ones.

[29]See e.g. Newman [2010, Chap. 17].

connected to a firm with technology $\mathbf{h}'$ and this firm being connected to a firm with technology $\mathbf{h}''$. Then we make the following pair approximation[30]

$$\tau_t(\mathbf{h}, \mathbf{h}', \mathbf{h}'') \approx \frac{m_t(\mathbf{h}, \mathbf{h}')m_t(\mathbf{h}', \mathbf{h}'')}{n_t(\mathbf{h}')}. \tag{19}$$

Appendix A provides a complete derivation of the dynamics for an arbitrary number N of technology categories using the approximation in Equation (19). 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.

7. Coevolution of Knowledge Stocks and the Network

The description of the dynamics of our system reduces drastically if we consider the dynamics of the stocks of knowledge and the probability of a link between firms with given knowledge stocks. More formally, let the fraction of firms with a stock of knowledge of s, $0 \leq s \leq N$, be given by $\binom{N}{s}\bar{x}_t(s)$, where we have denoted by

$$\bar{x}_t(s) \equiv \frac{1}{\binom{N}{s}} \sum_{\mathbf{h} \in \mathcal{H}^N : |\mathsf{S}(\mathbf{h})| = s} x_t(\mathbf{h}), \tag{20}$$

$\bar{x}_t(s)$ being the solution of Equation (17) and 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 $\binom{N}{s}\binom{N}{s'}\bar{\xi}_t(s, s')$, where we have introduced

$$\bar{\xi}_t(s, s') \equiv \frac{1}{\binom{N}{s}} \frac{1}{\binom{N}{s'}} \sum_{\substack{\mathbf{h} \in \mathcal{H}^N : |\mathsf{S}(\mathbf{h})| = s \\ \mathbf{h}' \in \mathcal{H}^N : |\mathsf{S}(\mathbf{h})| = s'}} \xi_t(\mathbf{h}, \mathbf{h}'). \tag{21}$$

and $\xi_t(\mathbf{h}, \mathbf{h}')$ being the solution of Equation (18). Further, define the symmetric matrix, $\bar{g}(s, s') = \bar{g}(s', s)$ for all $0 \leq s, s' \leq N$, given by

$$\bar{g}(s, s') \equiv \frac{1}{\binom{N}{s}} \frac{1}{\binom{N}{s'}} \sum_{\substack{\mathbf{h} \in \mathcal{H}^N : |\mathsf{S}(\mathbf{h})| = s \\ \mathbf{h}' \in \mathcal{H}^N : |\mathsf{S}(\mathbf{h}')| = s'}} g(\mathbf{h}, \mathbf{h}'), \tag{22}$$

with $g(\mathbf{h}, \mathbf{h}')$ as in Equation (16) in Theorem 1. Then the dynamics for the fraction $\bar{x}_t(s)$ of firms with knowledge stock s, and the probability $\bar{\xi}_t(s, s')$ of a link between firms with knowledge stocks s and s', respectively, are given by the following proposition.

Proposition 2. *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 $\bar{\xi}_t(s, s')$ for any $0 \leq s, s' \leq N$ defined as in Equations (20) and (21). Then $\bar{x}_t(s)$ is the solution of the system of ODEs*

$$\begin{aligned} \frac{d\bar{x}_t(s)}{dt} &= (\gamma s + \alpha(s-1)s)\bar{x}_t(s-1) + \lambda(N-s)\bar{x}_t(s+1) - (\lambda s + \gamma(N-s) + \alpha s(N-s))\bar{x}_t(s) \\ &\quad + \beta \sum_{s'=1}^{N} \binom{N-1}{s'-1} \left(s\bar{\xi}_t(s-1, s')\bar{x}_t(s-1)\bar{x}_t(s') - (N-s)\bar{\xi}_t(s, s')\bar{x}_t(s)\bar{x}_t(s') \right), \end{aligned} \tag{23}$$

[30] The rationale for Equation (19) is that the expected number of links between firms with technology $\mathbf{h}$ and technology $\mathbf{h}'$ is given by $m_t(\mathbf{h}, \mathbf{h}')$, and the expected number of links to firms with technology $\mathbf{h}''$ involving a firm $\mathbf{h}'$ is given by $\frac{m_t(\mathbf{h}', \mathbf{h}'')}{n_t(\mathbf{h}')}$. A more detailed discussion can be found e.g. in Do and Gross [2009]; Gross et al. [2006]; Keeling and Eames [2005]. The pair approximation becomes exact for "locally tree-like" networks [cf. Dembo and Montanari, 2010].

and $\bar{\xi}_t(s,s')$ *is the solution of the system of ODEs*

$$\frac{d\bar{\xi}_t(s,s')}{dt} = \rho\, \bar{g}(s,s') - \rho\,(1+\bar{g}(s,s'))\bar{\xi}_t(s,s') + o(\rho). \tag{24}$$

The stationary solution of the above dynamics for the special case of $\beta = 0$ is particularly simple and given in the next corollary.

Corollary 1. *Let the expected number of firms with knowledge stock* s *be denoted by*

$$\tilde{x}_t(s) \equiv \sum_{\mathbf{h}\in\mathcal{H}^N:|\mathsf{S}(\mathbf{h})|=s} x_t(\mathbf{h}), \quad 0 \leq s \leq N,$$

then in the case of $\beta = 0$ *the stationary distribution of* $\tilde{x}(s) = \lim_{t\to\infty}\tilde{x}_t(s)$ *is given by*

$$\tilde{x}(s) = \left(\sum_{k=0}^{N} \frac{\prod_{j=1}^{k-1}(N-j)(\gamma+\alpha j)}{\prod_{j=1}^{k}\lambda j}\right)^{-1} \frac{\prod_{k=1}^{s-1}(N-k)(\gamma+\alpha k)}{\prod_{k=1}^{s}\lambda k}. \tag{25}$$

If we also set $\alpha = 0$ *then we obtain a binomial distribution with success probability* $\frac{\gamma}{\lambda+\gamma}$ *so that*

$$\tilde{x}(s) = \binom{N}{s}\left(\frac{\gamma}{\lambda+\gamma}\right)^{s}\left(\frac{\lambda}{\lambda+\gamma}\right)^{N-s}. \tag{26}$$

Equation (26) corresponds to a simple birth-death process with birth rate γ and death rate λ [Grimmett and Stirzaker, 2001]. Note that we can further simplify Equation (22) to

$$\begin{aligned}\bar{g}(s,s') = \frac{1}{\binom{N}{s}}\frac{1}{\binom{N}{s'}} \sum_{k=\max\{0,s'-s\}}^{\min\{N-s,s'\}} \binom{N-s}{k}\binom{s}{s'-k}\binom{N}{N-s} \\ \times \frac{e^{\eta(\beta g_{\theta,\tau}(x_t)(1+bs)^{\theta}k-c)}}{1+e^{\eta(\beta g_{\theta,\tau}(x_t)(1+bs)^{\theta}k-c)}} \frac{e^{\eta(\beta g_{\theta,\tau}(x_t)(1+bs')^{\theta}(s-s'+k)-c)}}{1+e^{\eta(\beta g_{\theta,\tau}(x_t)(1+bs')^{\theta}(s-s'+k)-c)}}.\end{aligned} \tag{27}$$

Moreover, note that the average stock of knowledge is given by $\bar{h}_t = \sum_{s=1}^{N} s\binom{N}{s}\bar{x}_t(s)$, while the average degree can be computed from $\bar{d}_t = \frac{2m_t}{n} = n\sum_{\mathbf{h},\mathbf{h}'\in\mathcal{H}^N} z_t(\mathbf{h},\mathbf{h}') = n\sum_{s,s'=0}^{N}\binom{N}{s}\binom{N}{s'}\bar{z}_t(s,s') = n\sum_{s,s'=0}^{N}\binom{N}{s}\binom{N}{s'}\bar{x}_t(s)\bar{x}_t(s')\bar{\xi}_t(s,s')$, where m_t denotes the number of links at time t. In the following section we study the stationary states of the dynamics introduced in Proposition 2 and their stability properties.

8. Equilibrium Characterization

In this section we first identify a threshold β^c such that the economy does not innovate if the technology spillover parameter β is below β^c. We then preform a comparative statics analysis of the threshold β^c, and, in particular, show that the threshold is increasing with the knowledge obsolescence rate λ and the linking cost c, while it is decreasing with the productivity function parameter b and the alliance duration τ. Moreover, the change with the uncertainty of R&D collaborations η is non-monotonic. For small spillover effects β (or alternatively, for a small number of technology categories) we further characterize the innovation gains (in terms of the average stock of knowledge) due to the presence of R&D collaborations. These gains are decreasing with the in-house R&D success rate, the linking cost and the collaboration uncertainty, while the the effect of the knowledge obsolescence rate is non-monotonic, and there exists a critical value λ^c as a function of the in-house R&D success rate γ such that the gains from collaboration are increasing with λ if $\lambda < \lambda^c$, and decreasing, otherwise.

The 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 $\bar{g}(s.s')$ does not depend on $\bar{x}(s)$, and letting $\rho \to \infty$. The following lemma 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.

Lemma 1. *Consider the limit $\rho \to \infty$ and independent markets with $\theta = 1$ in Proposition 2. Then the fixed points for $\bar{\xi}(s, s')$ of Equation (24) are given by*

$$\bar{\xi}(s, s') = \frac{\bar{g}(s, s')}{1 + \bar{g}(s, s')},$$

while the fixed points of Equation (23) satisfy

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

Lemma 1 shows that $\bar{x}(s)$ is higher for all $s > 0$ when $\beta > 0$. Hence, the presence of R&D collaborations leads to a higher average knowledge stock in the economy. Moreover, from Lemma 1 we find that for $\gamma, \alpha \to 0$ a stationary solution is always given by $\bar{x}(s) = \delta_{s,0}$, where firms have empty technology portfolios and thus vanishing stock of knowledge. However, this trivial stationary state is not the only stationary state if β exceeds a threshold. Moreover, the trivial stationary state becomes unstable if β is higher than this threshold.

Proposition 3. *Consider the limit of $\rho \to \infty$, $\theta = 1$ and $\gamma, \alpha \to 0$ in Proposition 2. Then the unique, asymptotically stable stationary state is $\bar{x}(s) = \delta_{s,0}$ if $\beta < \beta^c$, with*

$$\beta^c = \lambda b \eta \tau \left(e^{c\eta} + 2\right) + \frac{W\left(\lambda b \eta \tau \left(e^{c\eta} + 1\right) e^{\eta(c - b\tau(e^{c\eta}+2)\lambda)}\right)}{b \eta \tau}, \tag{28}$$

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

Proposition 3 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 non-vanishing innovation activities (i.e. a significant fraction of firms has non-empty technology portfolios, $\bar{x}(0) < 1$) if the spillover parameter β exceeds a threshold β^c. Proposition 3 further states that when $\beta > \beta^c$ then the trivial stationary state becomes unstable and there exists non-trivial stable stationary state. Moreover, one can show that (see the proof of Proposition 3 in Appendix B)

$$\frac{\partial \beta^c}{\partial \lambda} > 0, \quad \frac{\partial \beta^c}{\partial c} > 0, \quad \frac{\partial \beta^c}{\partial b} < 0, \quad \frac{\partial \beta^c}{\partial \tau} < 0, \tag{29}$$

while β^c is a convex function of η as indicated in Figure 5 for different values of the linking cost c. The existence of a threshold 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 incorporated other firms' technologies [cf. Cohen and Levinthal, 1990; Griffith et al., 2003].

In contrast, in the absence of technology spillovers, when $\beta = 0$, the following corollary follows immediately from Lemma 1.

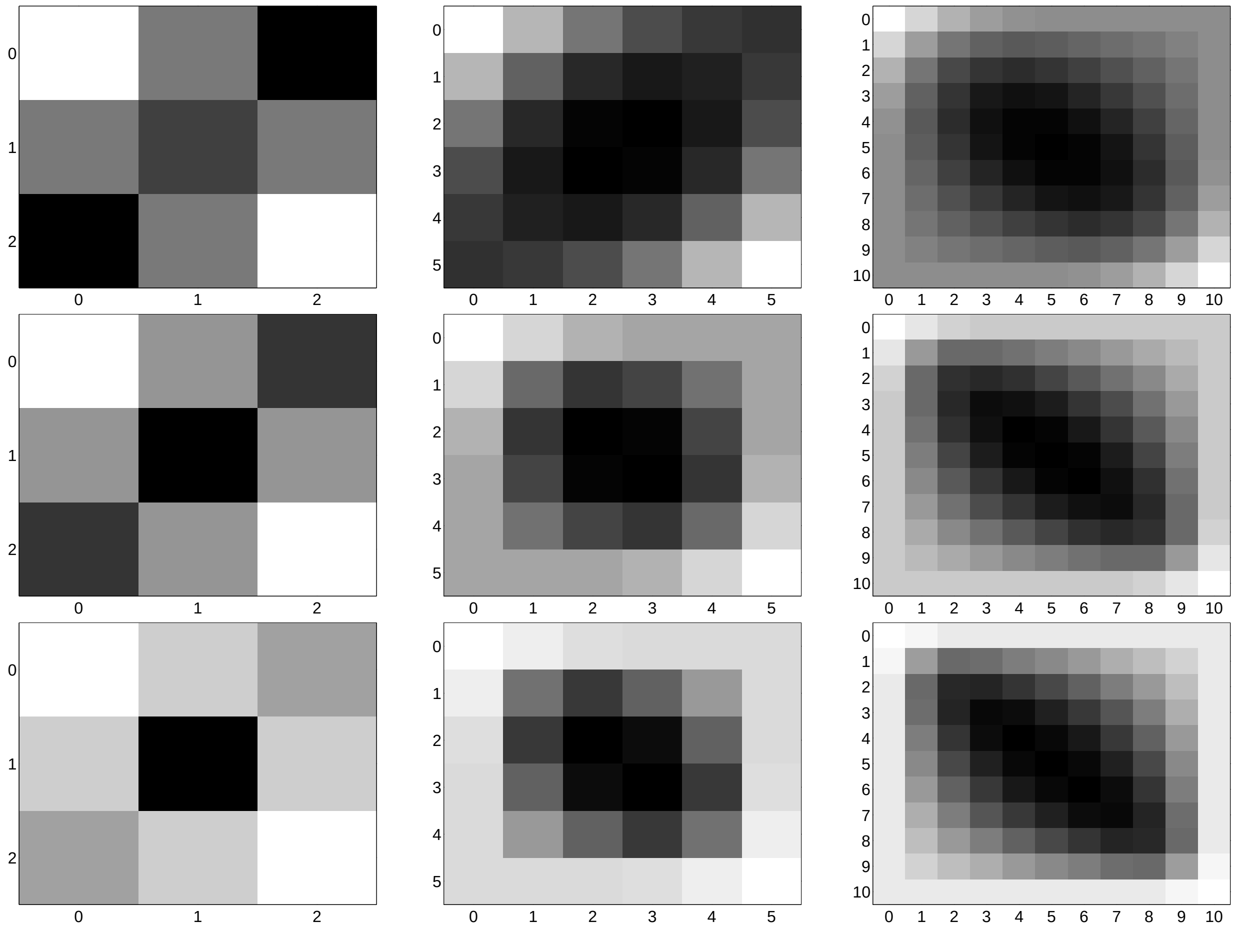


Figure 4: Examples of the stationary expected fraction of links $\bar{\xi}(s, s')$, $0 \leq s, s' \leq N$ with $c = 1$, $\theta = 1$, $\beta\tau b = 1$, for $\eta = 1, 2, 3$ (rows) and $N = 2, 5, 10$ (columns) (where higher values are black and lower values are white). We observe that with increasing 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. Assortativity has been observed in empirical studies of R&D collaboration networks.

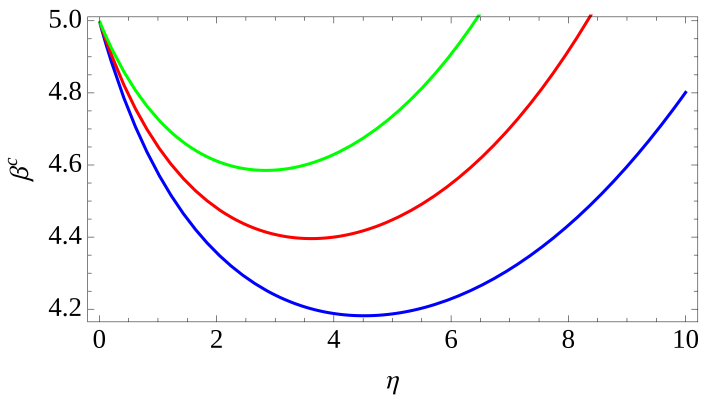


Figure 5: The threshold level β as a function of η for $b = 10$, $\tau = 0.01$, $\lambda = 1$, and $c \in \{1, 1.25, 1.5\}$.

Corollary 2. *Consider $\rho \to \infty$ and let $\beta = 0$, $\theta = 1$ in Proposition 2, then the stationary stocks of knowledge are given by*

$$\bar{x}(s) = \left(\sum_{k=0}^{N} \binom{N}{k} \prod_{l=0}^{k-1} \frac{\gamma + \alpha l}{\lambda} \right)^{-1} \prod_{k=0}^{s-1} \frac{\gamma + \alpha k}{\lambda}. \tag{30}$$

The next proposition characterizes the stationary state in the limit small β/γ, that is, when the in-house R&D capabilities are much higher than the spillovers from collaboration.

Proposition 4. *Consider the limit $\rho \to \infty$ and independent markets with $\theta = 1$ in Proposition 2. Then for $\beta/\gamma \to 0$, the stationary stocks of knowledge are given by*

$$\bar{x}(s) = \bar{x}_0(s) + \frac{\beta}{\gamma} \frac{b}{a^2} \left(\frac{b_s}{b} - \bar{x}_0(s) \right) + O\left(\frac{\beta}{\gamma} \right)^2, \tag{31}$$

where $\bar{x}_0(s) \equiv \bar{x}(s)|_{\beta=0}$ is given in Equation (30),

$$b_s = \sum_{k=0}^{s-1} \prod_{l \neq k}^{s-1} \frac{\gamma + \alpha l}{\lambda} \sum_{s'=1}^{N} \binom{N-1}{s'-1} \frac{\bar{g}(k, s')}{1 + \bar{g}(k, s')} \prod_{k''=0}^{s'-1} \frac{\gamma + \alpha k''}{\lambda},$$

$b = \sum_{s'=0}^{N} \binom{N}{s'} b_{s'}$ and $a = \sum_{k=0}^{N} \binom{N}{k} \prod_{l=0}^{k-1} \frac{\gamma+\alpha l}{\lambda}$ for all $s = 0, \dots, N$. Moreover, the eigenvalues of the Jacobian $\mathbf{J}$ corresponding to the dynamical system in Equations (23) and (24) are all real, and consequently their solution trajectories do not show oscillatory behavior.

The average stock of knowledge is given by

$$\bar{h} = \sum_{s=1}^{N} s \binom{N}{s} \bar{x}(s) = \bar{h}_0 + \frac{\beta}{\gamma} \frac{b}{a^2} \left(\frac{\sum_{s=1}^{N} s \binom{N}{s} b_s}{b} - \bar{h}_0 \right) + O(\beta^2),$$

where $\bar{h}_0 = \frac{1}{a} \sum_{s=1}^{N} s \binom{N}{s} a_s$. The stationary average stock of knowledge $\bar{h}$ can be seen in Figure 6 for varying values of γ and λ. The gains from R&D collaborations are given by

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

The relative gains from R&D collaborations, $\Delta \bar{h}/\bar{h}_0$, are illustrated in the right panels in Figure 6. 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. However, the effect of the knowledge obsolescence rate λ is non-monotonic. We will identify an explicit critical value λ^c for λ as a function of γ in the next section focussing on the case of $N = 1$ for which $\Delta \bar{h}/\bar{h}_0$ is increasing with λ if $\lambda < \lambda^c$, or decreasing, otherwise.

Proposition 4 also shows that at least when spillover effects are not too strong, a non-competitive economy cannot generate the oscillatory time evolution of the average degree that we have documented in Section 2. However, as we are going to demonstrate in the following sections, such oscillations can be obtained in a competitive environment. In order to demonstrate this, and to obtain more general results beyond the case of small spillover effects, in the following two sections we will confine our analysis to the case of at most two competing technologies, with $N \in \{1, 2\}$,[31]

[31] This simplifying restriction is shared with various other works in a similar context such as Lazear [2004]; Montanari

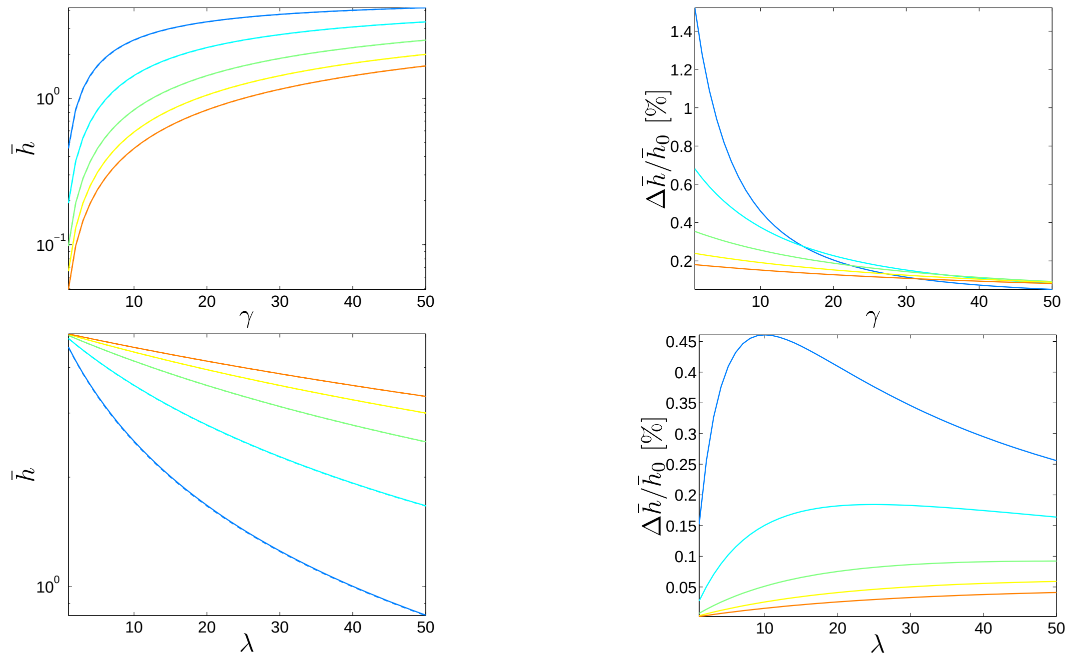


Figure 6: (Top left panel) The stationary average stock of knowledge $\bar{h}$ over different values of $\gamma \in [1, 50]$ for varying values of $\lambda \in \{10, 25, 50, 75, 100\}$. (Top right panel) The relative percentage gain in the average stock of knowledge, $\Delta\bar{h}/\bar{h}_0$, for the same parameters. (Bottom left panel) The stationary average stock of knowledge $\bar{h}$ over different values of $\lambda \in [1, 50]$ for varying values of $\gamma \in \{10, 25, 50, 75, 100\}$. (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 $\beta = 1$.

considering both, absence of competition, $\theta = 1$, and a competitive environment setting $\theta = 0$.

8.1. The Case of $N = 1$

In the case of $N = 1$ where $s \in \{0, 1\}$ we obtain from Equation (23)

$$\begin{aligned}\frac{d\bar{x}_t(1)}{dt} &= \gamma\bar{x}_t(0) - \lambda\bar{x}_t(1) + \beta\bar{\xi}_t(0,1)\bar{x}_t(0)\bar{x}_t(1), \\ \frac{d\bar{x}_t(0)}{dt} &= \lambda\bar{x}_t(1) - \gamma\bar{x}_t(0) - \beta\bar{\xi}_t(0,1)\bar{x}_t(0)\bar{x}_t(1),\end{aligned} \tag{32}$$

and from Equation (24) we get

$$\frac{d\bar{\xi}_t(0,1)}{dt} = \rho\bar{g}(0,1) - \rho(1 + \bar{g}(0,1))\bar{\xi}_t(0,1) + o(\rho), \tag{33}$$

with

$$\bar{g}(0,1) = \begin{cases} \frac{e^{\eta(\beta g_{1,\tau} - c)}}{1+e^{\eta(\beta g_{1,\tau} - c)}} \frac{e^{-\eta c}}{1+e^{-\eta c}} = \frac{e^{\eta(\beta b\tau - 2c)}}{(1+e^{\eta(\beta b\tau - c)})(1+e^{-\eta c})} & \text{if } \theta = 1, \\ \frac{e^{\eta(\beta g_{0,\tau}(x) - c)}}{1+e^{\eta(\beta g_{0,\tau}(x) - c)}} \frac{e^{-\eta c}}{1+e^{-\eta c}} = \frac{e^{\eta(\frac{2\beta\tau b}{1+bh_t(x)} - c)}}{1+e^{\eta(\frac{2\beta\tau b}{1+bh_t(x)} - c)}} \frac{e^{-\eta c}}{1+e^{-\eta c}} & \text{if } \theta = 0, \end{cases} \tag{34}$$

---

and Saberi [2010]; Young [2002], and it is also at the center of the analysis of two competing technologies in Jovanovic and MacDonald [1994].

where the average stock of knowledge is given by $\bar{h}_t(x) = \bar{x}_t(1)$. 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 $\lim_{\eta\to 0} \bar{g}(0,1) = \frac{1}{4}$. We then can state the following proposition, characterizing the stationary states and their stability properties.

Proposition 5. *Consider the limit $\rho \to \infty$, $N = 1$ in Proposition 2 and denote by $x = \lim_{t\to\infty} \bar{x}_t(1)$ and $z = \lim_{t\to\infty} \bar{\xi}_t(0,1)$.*

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

$$\beta^c = \lambda\left(e^{c\eta}+2\right) + \frac{W\left((2-\theta)\lambda b\eta\tau\left(e^{c\eta}+1\right)e^{\eta(c-(2-\theta)b\tau(e^{c\eta}+2)\lambda)}\right)}{(2-\theta)b\eta\tau}, \quad \theta \in \{0,1\}, \tag{35}$$

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

(ii) ***No competition:*** *Let $\theta = 1$, then the stationary state of the dynamic system in Equations (32) and (33) is asymptotically stable and given by*

$$\begin{aligned} x &= \frac{g\beta - \gamma - g\gamma - \lambda - g\lambda + \sqrt{4g(1+g)\beta\gamma + (\gamma+\lambda+g(\gamma+\lambda-\beta))^2}}{2g\beta}, \\ z &= \frac{g}{1+g}, \end{aligned} \tag{36}$$

where $g \equiv \bar{g}(0,1)$ is given in Equation (34) and $\bar{x}_t(0) = 1 - \bar{x}_t(1)$. The Jacobian has only negative, real eigenvalues, so that the solution trajectories for $\bar{x}_t(0), \bar{x}_t(1)$ and $\bar{\xi}_t(0,1)$ do not exhibit oscillatory behavior. Moreover, in the limit of $\gamma \to 0$ the non-trivial solution is given by

$$x = 1 - \frac{\lambda(2 + e^{\eta(c-b\beta\tau)}\left(e^{b\beta\eta\tau} + e^{c\eta} + 1\right))}{\beta}. \tag{37}$$

(iii) ***Competition:*** *Let $\theta = 0$, $\gamma = 0$ and consider small τ such that terms of the order $O(\tau^2)$ can be neglected. Then the non-trivial stationary state of the dynamic system in Equations (32) and (33) is given*

$$\begin{aligned} x = \frac{1}{2Ab\beta n}&\left(\beta n\left(e^{2c\eta}(-2\beta b\eta\tau + b - 1) - 2e^{c\eta}(b(\beta\eta\tau - 1) + 1) + 2(b-1)\right) - A^2 b\lambda\right. \\ &+ \left(A^4 b^2\lambda^2 + 2A^2 b\beta\lambda n\left(2e^{c\eta}(\beta b\eta\tau - b - 1) + e^{2c\eta}(2\beta b\eta\tau - b - 1) - 2(b+1)\right)\right. \\ &\left.\left.+\beta^2 n^2\left(2e^{c\eta}(\beta b\eta\tau + b + 1) + e^{2c\eta}(2\beta b\eta\tau + b + 1) + 2(b+1)\right)^2\right)^{\frac{1}{2}}\right), \end{aligned} \tag{38}$$

and

$$z = \frac{1}{A}\left(1 + \frac{2b\beta\eta\tau e^{c\eta}\left(e^{c\eta}+1\right)}{A(bx+1)}\right), \tag{39}$$

where we have denoted by $A = 2 + 2e^{c\eta} + e^{2c\eta}$.

(iv) ***Large shocks:*** *In the case of $\eta \to 0$ and $\theta \in \{0,1\}$ the asymptotically stable stationary state is given by*

$$x = \frac{-5(\gamma+\lambda) + \sqrt{(\beta n - 5(\gamma+\lambda))^2 + 20\beta\gamma n} + \beta n}{2\beta n}, \tag{40}$$

and

$$z = \frac{1}{5},$$

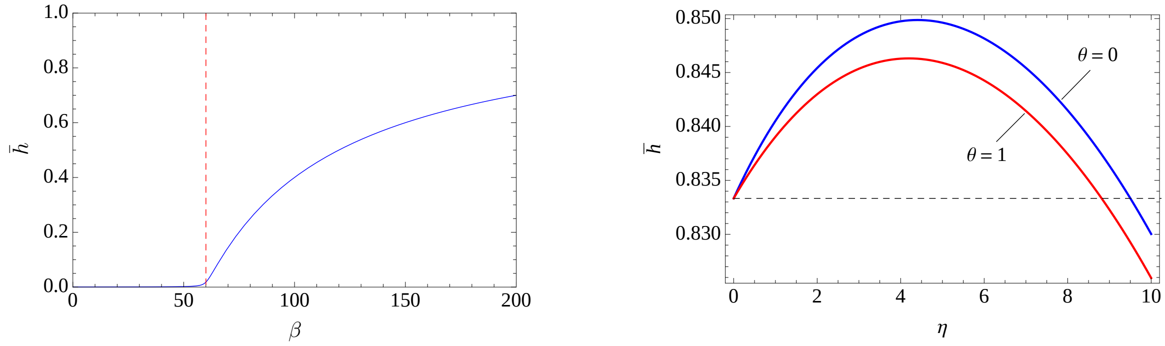


Figure 7: (Left panel) Stationary average knowledge stock $\bar{h}$ as a function of β. The dashed line indicates the threshold β^c from Equation (35). (Right panel) The stationary average knowledge stock $\bar{h}$ as a function of η for $\theta \in \{0,1\}$, $\lambda = 1$, $\beta = 30$, $b = 1$, $\tau = 0.01$, $c = 0.1$ and $\gamma = 0$. The dashed line indicates the solution from Equation (40). The average knowledge stock in the competitive case ($\theta = 0$) is always higher than in the non-competitive case ($\theta = 1$).

and the Jacobian has only real eigenvalues, so that the solution trajectories for $\bar{x}_t(0), \bar{x}_t(1)$ and $\bar{\xi}_t(0,1)$ do not exhibit oscillatory behavior.

(vi) ***Small shocks:*** *When $\eta \to \infty$, $\theta \in \{0,1\}$ and $c > 0$ then the asymptotically stable stationary state is given by $\lim_{t\to\infty} \bar{x}_t(1) = \frac{\gamma}{\lambda+\gamma}$ and $\lim_{t\to\infty} \bar{\xi}_t(0,1) = 0$ and the Jacobian has only real eigenvalues.*

The left panel in Figure 7 also illustrates the stationary fraction $\bar{h}$ as a function of β together with the threshold β^c from Equation (35). A significant fraction of firms has on average knowledge of the technology once the spillover parameter β exceeds the critical value β^c. The same comparative statics as in Equation (29) hold for the critical level β^c. Further, from Equation (35) we find that the threshold β^c is lower in the competitive case ($\theta = 0$) than in the non-competitive case ($\theta = 1$). Hence, introducing competition lowers the threshold above which innovation can take off in the economy. Moreover, from Equation (37) in Proposition 5 we find that the asymptotic stocks of knowledge $\bar{h} \equiv \lim_{t\to\infty} \bar{x}_t(1)$ satisfy (see also the proof of Proposition 5 in Appendix B)

$$\frac{\partial \bar{h}}{\partial \lambda} < 0, \quad \frac{\partial \bar{h}}{\partial c} < 0, \quad \frac{\partial \bar{h}}{\partial \beta} > 0, \quad \frac{\partial \bar{h}}{\partial b} > 0, \quad \frac{\partial \bar{h}}{\partial \tau} > 0.$$

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 $c\left(e^{b\beta\eta\tau} + 2e^{c\eta} + 1\right) < b\beta\tau\left(e^{c\eta} + 1\right)$ and $\frac{\partial \bar{h}}{\partial \eta} < 0$ otherwise. This is shown in the right panel of Figure 7. As Figure 7 illustrates, we find that the average knowledge stock in the competitive case ($\theta = 0$) is higher than in the non-competitive case ($\theta = 1$). From Equation (40) we also find that $\frac{\partial \bar{h}}{\partial \gamma} > 0$ (see also the proof of Proposition 5 in Appendix B). Further, from Equation (36) for the case of $\theta = 1$ in the limit of small β we can write

$$\bar{h} = \frac{\gamma}{\gamma + \lambda} + \beta \frac{\gamma\lambda}{(2e^{c\eta} + e^{2c\eta} + 2)(\gamma + \lambda)^3} + O\left(\beta^2\right),$$

which is what we would get from Equation (31) in the case of $N = 1$. We then find for $\Delta\bar{h}$ that (see also the proof of Proposition 5 in Appendix B)

$$\frac{\partial \Delta\bar{h}}{\partial c} < 0, \quad \frac{\partial \Delta\bar{h}}{\partial \eta} < 0,$$

while

$$\frac{\partial \Delta \bar{h}}{\partial \gamma}\begin{cases} > 0 & \text{if } \gamma < \frac{\lambda}{2}, \\ < 0 & \text{otherwise.} \end{cases}$$

Similarly, we find that

$$\frac{\partial \Delta \bar{h}}{\partial \lambda}\begin{cases} > 0 & \text{if } \lambda < \frac{\gamma}{2}, \\ < 0 & \text{otherwise.} \end{cases}$$

Increasing linking costs c or a reduction in the noise η unanimously reduce the innovation gains $\Delta \bar{h}$ from collaboration, while the effect of the in-house R&D innovation rate γ and the knowledge obsolescence rate λ are ambiguous, and increase $\Delta \bar{h}$ only if they are below a threshold level.

When we do not make the assumption that ρ is large, then we need to take into account the remainder term of the order of $o(\rho)$ in Equation (33). Using the pair approximation of Equation (19) we can state the following proposition:[32]

Proposition 6. *Let $N = 1$, $\theta = 1$ in Proposition 2 and denote by $x = \lim_{t\to\infty} \bar{x}_t(1)$, $z_1 = \lim_{t\to\infty} \bar{\xi}_t(0,1)$, $z_2 = \lim_{t\to\infty} \bar{\xi}_t(0,0)$ and $z_3 = \lim_{t\to\infty} \bar{\xi}_t(1,1)$. Assume that the pair approximation in Equation (19) holds.*

(i) The stationary state is given by

$$\begin{aligned} x &= \frac{z_1\beta - \gamma - \lambda + A(z_1)}{2z_1\beta} \\ z_2 &= \frac{2xz_1\lambda + g_2\rho - g_2 x\rho}{2x\lambda - (1+g_2)(-1+x)\rho} \\ z_3 &= \frac{2(-1+x)z_1(xz_1\beta+\gamma) - g_3 x\rho}{2(-1+x)(xz_1\beta+\gamma) - (1+g_3)x\rho}, \end{aligned} \tag{41}$$

where z_1 is the root of

$$\begin{aligned} 2g_1 - 2(1+g_1)z_1 &= \frac{2(g_3(-1+z_1)+z_1)\lambda\,(z_1\beta - \gamma - \lambda + A(z_1))}{(z_1\beta+\gamma+\lambda - A(z_1))\,(2\lambda+\rho+g_3\rho)} \\ &+ \frac{(g_2(-1+z_1)+z_1)\,(z_1\beta+\gamma+\lambda-A(z_1))^2\,(z_1\beta+\gamma-\lambda+A(z_1))}{(z_1\beta-\gamma-\lambda+A(z_1))\,(z_1\beta(2\lambda+\rho+g_2\rho)+(-\gamma-\lambda+A(z_1))\,(2\lambda-(1+g_2)\rho))}, \end{aligned}$$

and $A(z_1) \equiv \sqrt{4z_1\beta\gamma + (-z_1\beta+\gamma+\lambda)^2}$.

(ii) In the case of $\gamma = 0$ the Jacobian has only real eigenvalues, trajectories do not oscillate, in first order of large ρ the non-trivial asymptotically stable solution is characterized by

$$\begin{aligned} z_1 &= \frac{1}{2\beta(g_2+1)(g_3+1)} \\ &\times \Big(\sqrt{(g_2+1)\,(4\beta(g_3+1)(g_1(g_2+1)(g_3+1)\rho+\lambda(g_2-g_3)) + (g_2+1)(\beta g_3 - (g_1+1)(g_3+1)\rho)^2)} \\ &-(g_1+1)g_2(g_3+1)\rho - g_1 g_3\rho - g_1\rho + \beta g_2 g_3 + \beta g_3 - \rho(1+g_3)\Big), \end{aligned} \tag{42}$$

and

$$x = 1 - \frac{\lambda}{\beta z_1},$$

[32] Appendix A provides a complete derivation of the dynamics for an arbitrary number N of technology categories in terms of a system of ODEs.

while the threshold level for β is given by

$$\beta^c = \frac{\lambda\left(2e^{c\eta}+e^{2c\eta}+2\right)\left((\lambda+2\rho)e^{b\beta\eta\tau}+(\lambda+\rho)e^{\eta(b\beta\tau+c)}+e^{c\eta}(\lambda+\rho)+e^{2c\eta}(\lambda+\rho)\right)}{\lambda\left(e^{c\eta}+1\right)\left(e^{b\beta\eta\tau}+e^{c\eta}\right)+\rho\left(2e^{c\eta}+e^{2c\eta}+2\right)e^{b\beta\eta\tau}}.$$

In the case of $\gamma = 0$ we find from Proposition 6 that (see the proof of Proposition 6 in Appendix B)

$$\frac{\partial \beta^c}{\partial \rho} < 0, \quad \frac{\partial \bar{h}}{\partial \rho} > 0.$$

Hence, a more adaptive network (higher ρ) implies a lower threshold β^c and a higher average stock of knowledge $\bar{h}$. Moreover, we note that in the non-competitive case of $\theta = 1$ for all the cases analyzed, the trajectories did not exhibit any oscillatory behavior (as indicated by the Jacobian having only real eigenvalues). However, this does not hold for the competitive case with $\theta = 0$, where such oscillations could be observed.

8.2. The Case of $N = 2$

When $N = 2$ with $s \in \{0, 1, 2\}$ we obtain from Equation (23)

$$\begin{aligned}
\frac{d\bar{x}_t(0)}{dt} &= 2\left(\lambda\bar{x}_t(1) - \gamma\bar{x}_t(0) - \beta(\bar{\xi}_t(0,1)\bar{x}_t(0)\bar{x}_t(1) + \bar{\xi}_t(0,2)\bar{x}_t(0)\bar{x}_t(2))\right) \\
\frac{d\bar{x}_t(1)}{dt} &= \gamma\bar{x}_t(0) + \lambda\bar{x}_t(2) - (\lambda+\gamma+\alpha)\bar{x}_t(1) + \beta\left(\bar{\xi}_t(0,1)\bar{x}_t(0)\bar{x}_t(1) - \bar{\xi}_t(1,1)\bar{x}_t(1)^2\right. \\
&\quad \left.+\bar{\xi}_t(0,2)\bar{x}_t(0)\bar{x}_t(2) - \bar{\xi}_t(1,2)\bar{x}_t(1)\bar{x}_t(2)\right) \\
\frac{d\bar{x}_t(2)}{dt} &= 2\left((\gamma+\alpha)\bar{x}_t(1) - \lambda\bar{x}_t(2) + \beta\left(\bar{\xi}_t(1,1)\bar{x}_t(1)^2 + \bar{\xi}_t(1,2)\bar{x}_t(1)\bar{x}_t(2)\right)\right),
\end{aligned} \tag{43}$$

and from Equation (24) we obtain

$$\begin{aligned}
\frac{d\bar{\xi}_t(0,1)}{dt} &= \frac{1}{2}\rho\tilde{g}(0,1) - \rho\left(1 + \frac{1}{2}\tilde{g}(0,1)\right)\bar{\xi}_t(0,1) + o(\rho) \\
\frac{d\bar{\xi}_t(0,2)}{dt} &= \rho\tilde{g}(0,2) - \rho\left(1 + \tilde{g}(0,2)\right)\bar{\xi}_t(0,2) + o(\rho) \\
\frac{d\bar{\xi}_t(1,1)}{dt} &= \frac{1}{4}\rho\tilde{g}(1,1) - \rho\left(1 + \frac{1}{4}\tilde{g}(1,1)\right)\bar{\xi}_t(1,1) + o(\rho) \\
\frac{d\bar{z}_t(1,2)}{dt} &= \frac{1}{2}\rho\tilde{g}(1,2) - \rho\left(1 + \frac{1}{2}\tilde{g}(1,2)\right)\bar{\xi}_t(1,2) + o(\rho).
\end{aligned} \tag{44}$$

We next identify the stationary states of the stochastic process and their stability properties.

Proposition 7. *Consider the limit $\rho \to \infty$, $N = 2$ in Proposition 2 and denote by $x_1 \equiv \lim_{t\to\infty}\bar{x}_t(1)$, $x_2 \equiv \lim_{t\to\infty}\bar{x}_t(2)$ and $z_1 \equiv \lim_{t\to\infty}\bar{\xi}_t(0,1)$, $z_2 \equiv \lim_{t\to\infty}\bar{\xi}_t(0,2)$, $z_3 \equiv \lim_{t\to\infty}\bar{\xi}_t(1,1)$, $z_4 \equiv \lim_{t\to\infty}\bar{\xi}_t(1,2)$.*

(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 Equation (35).*

(ii) ***No competition:*** *Assume that $\alpha = \gamma = 0$ and consider $\theta = 1$, then in the limit of $\tau \to 0$ the non-trivial stationary state of Equation (43) is given by*

$$x_2 = \frac{\beta x_1^2 z_1}{\lambda - \beta x_1 z_1},$$

and

$$
\begin{aligned}
x_1 = {} & \frac{1}{3\sqrt[3]{A}\beta^2 z_1^2(4\beta(3\beta z_1(z_1-z_2)+2z_1-z_2)+1)} \\
& \times(-A^{2/3}+2\sqrt[3]{A}\beta z_1(\lambda+\beta(\beta z_1^2(6\beta(z_1-z_2)+1)+\lambda(12\beta z_1(z_1-z_2)+9z_1-2z_2))) \\
& -4\beta^6 z_1^6(6\beta(z_1-z_2)+1)^2+4\beta^4\lambda z_1^4(\beta(12(3\beta^2 z_1(z_1-z_2)(2z_1+z_2) \\
& +\beta(4z_1^2-2z_1z_2+z_2^2)+z_1)-5z_2)+1)-\beta^2\lambda^2 z_1^2(4\beta(144\beta^3 z_1^2(z_1-z_2)^2 \\
& +48\beta^2 z_1(3z_1^2-4z_1z_2+z_2^2)+\beta(48z_1^2-27z_1z_2+4z_2^2)+6z_1-z_2)+1))),
\end{aligned}
$$

where A is a function of β,η and c (provided in the proof of the proposition in the appendix), while the stationary state of Equation (44) is given by

$$
\begin{aligned}
z_1 = z_3 = z_4 &= \frac{1}{2e^{c\eta}+e^{2c\eta}+2} + \frac{b\beta\eta\tau e^{c\eta}\left(e^{c\eta}+1\right)}{\left(2e^{c\eta}+e^{2c\eta}+2\right)^2} + O\left(\tau^2\right), \\
z_2 &= \frac{1}{2e^{c\eta}+e^{2c\eta}+2} + \frac{2b\beta\eta\tau e^{c\eta}\left(e^{c\eta}+1\right)}{\left(2e^{c\eta}+e^{2c\eta}+2\right)^2} + O\left(\tau^2\right).
\end{aligned}
$$

If we also set $c=0$ then the nontrivial stationary state of Equation (43) simplifies to $x_1 = \frac{5\lambda(\beta-5\lambda)}{\beta^2}$ and $x_2 = \frac{(\beta-5\lambda)^2}{\beta^2}$ with $z_1=z_2=z_3=z_4=\frac{1}{5}$. If also $\alpha < \frac{\lambda}{2}$ then the Jacobian has only real eigenvalues, so that the solution trajectories for $\bar{x}_t(s)$ and $\bar{\xi}_t(s,s')$, $s,s' \in \{0,1,2\}$, do not exhibit oscillatory behavior.

(iii) ***Large shocks:*** *Assume that $\gamma = 0$ and that $\bar{x}_0(0) < 1$. In the limit of $\eta \to 0$ the stationary state of the dynamic system in Equation (43) is given by $x_1 = x_2 = 0$, or*

$$
x_2 = \frac{2\alpha(A-5\alpha)+\lambda(-25\alpha+A-\beta)}{2\alpha\beta},
$$

and

$$
\begin{aligned}
x_1 = {} & \frac{1}{6\alpha\beta}\Big[15\alpha^2+45\alpha\lambda-2A(2\alpha+\lambda) \\
& +\alpha\left(100\alpha\lambda^3+\lambda^2\left(-175\alpha^2+2\beta(\beta-A)+10\alpha\beta\right)+\alpha^2\left(25\alpha^2+\beta(4A+5\beta)+10\alpha\beta\right)\right. \\
& \left.+2\alpha\lambda\left((\beta-5\alpha)^2-A\beta\right)\right)^{\frac{1}{2}}+\alpha^2\beta+2\alpha\beta\lambda\Big],
\end{aligned}
$$

where we have denoted by $A \equiv \sqrt{100\alpha\lambda+(5\alpha+\beta)^2}$, while the stationary state of Equation (44) is given by $z_1=z_2=z_3=z_4=\frac{1}{5}$. If also $\alpha < \frac{\lambda}{2}$, then the Jacobian has only real eigenvalues, so that the solution trajectories for $\bar{x}_t(s)$ and $\bar{\xi}_t(s,s')$, $s,s' \in \{0,1,2\}$, do not exhibit oscillatory behavior.

(iv) ***Small shocks:*** *Assume that $\beta b\tau > c$ in the case of $\theta = 1$ and $\frac{2b(b+1)\beta\tau}{4b+1} > c$ in the case of $\theta = 0$. In the limit of $\eta \to \infty$ starting from an empty graph $\bar{K}_n$ the stationary state of the dynamic system in Equation (43) is given by*

$$
x_1 = \frac{-3\alpha\gamma-3(\gamma+\lambda)^2+A}{2\beta\gamma},
$$

and

$$
x_2 = \frac{1}{12\beta\gamma^2\lambda}\left(3\alpha\gamma+3\gamma^2-6\gamma\lambda-3\lambda^2+A\right)\left(-3\alpha\gamma-3\left(\gamma+\lambda\right)^2+A\right),
$$

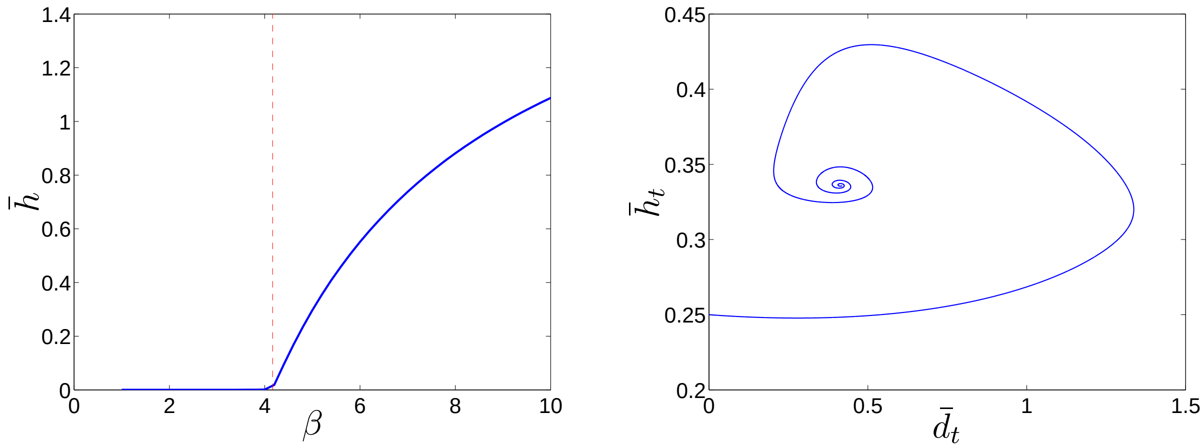


Figure 8: (Left panel) The asymptotic average stock of knowledge $\lim_{t\to\infty} \bar{h}_t = \lim_{t\to\infty} 2(\bar{x}_t(1) + \bar{x}_t(2))$ for $\eta = 0$, $\gamma = 0.01$, $b = 1$, $\rho = 10$, $n = 500$, $\lambda = 500$ as a function of β with the threshold β^c indicated with a dashed line from Equation (35). (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 $\alpha = 0$, $\beta = 700$, $b = 69$, $c = 2900$, $\gamma = 1$, $\eta = 0.01$, $\lambda = 400$, $n = 500$, $\rho = 1/\tau + 300$ and $\tau = 0.01$. The initial condition is $x_0(0) = 0.75$, $\bar{x}_0(1) = 0.125$, $\bar{x}_0(2) = 0$ and an empty network.

where we have denoted by $A = \sqrt{12\beta\gamma^2\lambda + 9\left(\alpha\gamma + (\gamma+\lambda)^2\right)^2}$, while the stationary state of Equation (44) is given by $z_1 = z_2 = z_4 = 0$ and $z_3 = \frac{1}{3}$. If also $\gamma = 0$ then the unique stationary state is $\lim_{t\to\infty} \bar{x}_t(0) = 1$. If $\beta b\tau < c$ in the case of $\theta = 1$ and $\frac{2b(b+1)\beta\tau}{4b+1} < c$ in the case of $\theta = 0$ then the stationary state is given by $z_1 = z_2 = z_3 = z_4 = 0$ and x_1, x_2 are determined by Equation (30). In the case of $\theta = 1$ the Jacobian has only real eigenvalues, so that the solution trajectories for $\bar{x}_t(s)$ and $\bar{\xi}_t(s, s')$, $s, s' \in \{0, 1, 2\}$, do not exhibit oscillatory behavior. In the case of $\theta = 0$ this holds if $\frac{2b(b+1)\beta\tau}{4b+1} < c$ and $\lambda > \frac{\gamma}{3}$.

This threshold β^c in part (i) of Proposition 7 is indicated with a dashed line in the right panel in Figure 8. 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 contrast to the case of $N = 1$ where the parameter α did not play any role, here we find in both cases, small and large shocks, 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 small shocks 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.

The left panel in Figure 8 shows the asymptotic average stock of knowledge $\lim_{t\to\infty} \bar{h}_t = \lim_{t\to\infty} 2(\bar{x}_t(1)+\bar{x}_t(2))$ for $\eta = 0$, $\gamma = 0.01$, $b = 1$, $\rho = 10$, $n = 500$, $\lambda = 500$ as a function of β with the threshold β^c indicated with a dashed line from Equation (35). The right panel in Figure 8 shows solution trajectories for general levels of ρ (instead of taking the limit of large ρ) using a pair approximation of Equation (19). The figure illustrates that in the competitive case ($\theta = 0$) the average stock of knowledge and the average degree can oscillate and spiral towards their stationary state. We will discuss this more in relation to empirically observed networks in Section 10.

We find that for all cases for which we could obtain an analytic characterization, cyclical behavior did not occur when we assumed that there is no competition ($\theta = 1$). In contrast, we observed cyclical patterns in our numerical integration of the governing differential equations when competition was allowed ($\theta = 1$), as for example in Figure 8. This is a first indication that competition increases the variably of the network density over time, and is a necessary ingredient to explain the pattern with have observed in Section 2.

9. Efficiency

In this section we define welfare from a social planners perspective and show that it 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 provide explicit welfare characterizations for a given network structure, and show that more centralized structures increase welfare. We then derive the welfare gains from competition, and compare different levels of uncertainty in R&D collaborations.

The social welfare function is given by aggregate profits

$$W_\theta(\mathbf{h}, G) = \sum_{i=1}^n \pi_i(\mathbf{h}, G) = \sum_{i=1}^n \left(\theta + (1-\theta) \frac{A_i(\mathbf{h}_i)}{\frac{1}{n}\sum_{j=1}^n A_j(\mathbf{h}_j)} \right) A_i(\mathbf{h}_i) - 2mc, \tag{45}$$

where m is the number of links in the network. In the case of independent markets ($\theta = 1$) we obtain

$$W_1(\mathbf{h}, G) = \sum_{i=1}^n A_i(\mathbf{h}_i) - 2mc,$$

In this case social welfare is increasing with the total productivity $A(\mathbf{h}) = \sum_{i=1}^n A_i(\mathbf{h}_i)$ in the economy and decreasing with the number of links m in the network. Using the fact that $A_i(\mathbf{h}_i) = a + b|\mathsf{S}(\mathbf{h}_i)|$ and setting $a = 1$ we obtain

$$\begin{aligned} W_1(\mathbf{h}, G) &= n + b \sum_{i=1}^n |\mathsf{S}(\mathbf{h}_i)| - 2mc \\ &= n \left(1 + b \frac{1}{n} \sum_{i=1}^n |\mathsf{S}(\mathbf{h}_i)| - \frac{2m}{n} c \right) \\ &= n(1 + b\bar{h} - c\bar{d}), \end{aligned} \tag{46}$$

where $\bar{h}$ is the average stock of knowledge and $\bar{d}$ the average degree. In contrast, in the case of competitive markets ($\theta = 0$) social welfare can be written as

$$W_0(\mathbf{h}, G) = \frac{\sum_{i=1}^n A_i(\mathbf{h}_i)^2}{\frac{1}{n}\sum_{i=1}^n A_i(\mathbf{h}_i)} - 2mc.$$

Gross social welfare is proportional to the ratio of the second to the first sample moment of the productivity distribution, and net welfare decreasing with the number of links m in the network.

Similarly, for $A_i(\mathbf{h}_i) = a + b|\mathsf{S}(\mathbf{h}_i)|$ with $a = 1$ we get

$$
\begin{aligned}
W_0(\mathbf{h}, G) &= \frac{n + 2b\sum_{i=1}^n |\mathsf{S}(\mathbf{h}_i)| + b^2\sum_{i=1}^n |\mathsf{S}(\mathbf{h}_i)|^2}{1 + \frac{1}{n}\sum_{i=1}^n |\mathsf{S}(\mathbf{h}_i)|} - 2mc \\
&= n\left(\frac{1 + 2b\frac{1}{n}\sum_{i=1}^n |\mathsf{S}(\mathbf{h}_i)| + b^2\frac{1}{n}\sum_{i=1}^n |\mathsf{S}(\mathbf{h}_i)|^2}{1 + b\frac{1}{n}\sum_{i=1}^n |\mathsf{S}(\mathbf{h}_i)|} - \frac{2m}{n}c\right) \\
&= n\left(1 + \frac{b\bar{h} + b^2\sigma_h^2 + b^2\bar{h}^2}{1 + b\bar{h}} - c\bar{d}\right) \\
&= n\left(1 + b\bar{h} + \frac{b^2\sigma_h^2}{1 + b\bar{h}} - c\bar{d}\right),
\end{aligned}
\tag{47}
$$

where σ_h^2 is the variance in the stocks of knowledge. With welfare in the case of independent markets from Equation (46) we then can write

$$
W_0(\mathbf{h}, G) - W_1(\mathbf{h}, G) = \frac{nb^2\sigma_h^2}{1 + b\bar{h}}. \tag{48}
$$

It follows that, when the stocks of knowledge are exogenously given, welfare in the competitive case 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 $c_v = \sigma_h/\bar{h}$).

We next take into account that the knowledge stocks are endogenous. The social planner's goal is to maximize welfare $W_\theta(\mathbf{h}, G)$ by choosing the network $G \in \mathcal{G}^n$ and knowing that the dynamics of knowledge depend on the network structure G of collaborating firms. Hence, we can write the social planner's problem as follows

$$
\max\nolimits_{G\in\mathcal{G}^n} V_{\theta,r}(G) = \max\nolimits_{G\in\mathcal{G}^n} \int_{t=0}^{\infty} e^{-rt}\mathbb{E}(W_\theta(\mathbf{h}_t, G)|G)dt \tag{49}
$$

subject to Equation (74). We denote by G^* the efficient network solving the above optimization problem.

9.1. The Non-Competitive Case

When $\theta = 1$ we obtain

$$
V_{1,r}(G) = \int_0^\infty e^{-rt}\mathbb{E}(W_1(\mathbf{h}_t, G)|G)dt = \frac{n}{r} + nb\int_0^\infty e^{-rt}\mathbb{E}\left(\sum_{i=1}^n |\mathsf{S}(\mathbf{h}_{i,t})| \,\middle|\, G\right)dt - \frac{2mc}{r}.
$$

We now 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.[33] 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

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

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

$$\begin{aligned}
\mathbb{E}(X_i(t)X_j(t)|G) &= \mathbb{E}(\mathbb{1}_{\{h_{ik,t}=1\}} \mathbb{1}_{\{h_{jk,t}=1\}}|G) \\
&= \mathbb{P}(h_{ik,t} = 1, h_{jk,t} = 1|G) \\
&= \mathbb{P}(h_{jk,t} = 1|h_{ik,t} = 1, G)\mathbb{P}(h_{ik,t} = 1|G) \\
&= \mathbb{P}(h_{jk,t} = 1|h_{ik,t} = 1, G)\mathbb{E}(\mathbb{1}_{\{h_{ik,t}=1\}}|G) \\
&= \mathbb{P}(h_{jk,t} = 1|h_{ik,t} = 1, G) y_i(t).
\end{aligned}$$

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

$$\frac{dy_i(t)}{dt} = \gamma - (\lambda + \gamma) y_i(t) + \beta \sum_{j=1}^{n} a_{ij} y_j(t) - \beta \sum_{j=1}^{n} a_{ij} y_i(t) y_j(t). \tag{50}$$

The following lemma describes the evolution of the stocks of knowledge according to Equation (50) 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 $i, j = 1, \ldots, n$.

From the above lemma 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 $i = 1, \ldots, n$ and $k = 1, \ldots, N$. 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 $\mathbf{A}$.

The objective function of the social planner for $\theta = 1$ can then be written as follows

$$V_{1,r}(G) = \int_0^\infty e^{-rt}\mathbb{E}(W_1(\mathbf{h}_t, G)|G)dt = \frac{n}{r} + nbN\gamma \sum_{j=1}^n \frac{\gamma\langle \mathbf{u}, \mathbf{v}_j\rangle^2}{\gamma+\lambda-\beta\mu_j}\left(\frac{1}{r} - \frac{1}{r+\gamma+\lambda-\beta\mu_j}\right) - \frac{2mc}{r}.$$

A discussion for the derivation without the assumption that $\text{Cov}(h_{ik,t}, h_{jk,t}|G) = 0$ 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].

In the limit of weak discounting for small values of r we can write[34]

$$V_{\theta,r}(G) = \int_0^\infty e^{-rt}\mathbb{E}(W_\theta(\mathbf{h}_t, G)|G)dt = O(1) + O\left(\frac{1}{r}\right)\lim_{t\to\infty}\mathbb{E}(W_\theta(\mathbf{h}_t, G)|G).$$

In the case of $\theta = 1$ we then obtain

$$V_{1,r}(G) = \int_0^\infty e^{-rt}\mathbb{E}(W_1(\mathbf{h}_t, G)|G)dt = O(1) + O\left(\frac{1}{r}\right)\left(n + b\lim_{t\to\infty}\mathbb{E}\left(\sum_{i=1}^n |\mathsf{S}(\mathbf{h}_{i,t})|\,\middle|\, G\right) - 2mc\right).$$

The next proposition derives the asymptotic knowledge stocks in the case of $\alpha = 0$.

Proposition 8. *Consider a given network G, let $\alpha = 0$, assume that asymptotically pair-correlations can be neglected, i.e. $\lim_{t\to\infty}\text{Cov}(h_{ik,t}, h_{jk,t}|G) = 0$ for all $i = 1,\ldots,n$ and $k = 1,\ldots,N$. Then*

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

$$\lim_{t\to\infty}\mathbb{E}\left(\sum_{i=1}^n |\mathsf{S}(\mathbf{h}_{i,t})|\,\middle|\, G\right) = N\left(1 - \cfrac{\frac{\lambda}{\lambda+\gamma}}{1 + \frac{\beta}{\lambda+\gamma}d_i - \frac{\beta\lambda}{(\lambda+\gamma)^2}\sum_{j=1}^n \cfrac{a_{ij}}{1+\frac{\beta}{\lambda+\gamma}d_j - \ddots}}\right), \tag{51}$$

(ii) we have the bounds

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

(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 $\mathbf{b}\left(G, \frac{\beta}{\lambda+\gamma}\right)$ is the Bonacich centrality vector defined as $\mathbf{b}\left(G, \frac{\beta}{\lambda+\gamma}\right) \equiv \left(\mathbf{I}_n - \frac{\beta}{\lambda+\gamma}\mathbf{A}\right)^{-1}\mathbf{u}$.

For example, a third-order continued fraction approximation to the stationary knowledge stocks

[34] Related works such as Aghion et al. [2005, Sec. III.C ff.] assume that the discounting rate r equals zero for all of their analytic results.

is given by

$$\lim_{t\to\infty} \mathbb{E}\left(\sum_{i=1}^{n} |\mathsf{S}(\mathbf{h}_{i,t})| \middle| G\right) \approx N\left(1 - \frac{\frac{\lambda}{\gamma+\lambda}}{1+\frac{\beta}{\gamma+\lambda}\sum_{j=1}^{n} a_{ij}\left(1-\frac{\frac{\lambda}{\gamma+\lambda}}{1+\frac{\beta}{\gamma+\lambda}\sum_{k=1}^{n} a_{jk}}\left(1-\frac{\frac{\lambda}{\gamma+\lambda}}{1+\frac{\beta}{\gamma+\lambda}d_k}\right)\right)}\right) \tag{52}$$

Note that the importance of the Bonacich centrality in relation to equilibria and aggregate outcome in network games has been prominently studied in Ballester et al. [2006]. We then can write the value function as follows

$$V_{1,r}(G) = \int_0^\infty e^{-rt}\mathbb{E}(W_1(\mathbf{h}_t, G)|G)dt = O(1) + O\left(\frac{n}{r}\right)\begin{cases} 1 + bN - c\bar{d}, & \text{if } \beta \gg \lambda, \\ 1 + \frac{\gamma bN}{\lambda+\gamma}\frac{1}{n}\left\|\mathbf{b}\left(G, \frac{\beta}{\lambda+\gamma}\right)\right\|_1 - c\bar{d}, & \text{if } \beta \ll \lambda, \\ 1 + \frac{bN\gamma}{\lambda+\gamma} - c\bar{d}, & \text{if } \beta \to 0. \end{cases}$$

From the above analysis we find that when $\theta = 1$ the efficient graph G^* is characterized by having a large largest eigenvalue μ_1 while minimizing on the number of links m. We then can give the following proposition.

Proposition 9. *Let $\alpha = 0$, and $h_{ik,0} = 0$ for all $i = 1, \ldots, n$ and $k = 1, \ldots, N$. Moreover, let $\mathcal{G}(n, m)$ denote the class of graphs with n nodes and m links. Then the graph maximizing the value function $\lim_{r\to 0} V_{1,r}$ is the graph with the largest eigenvalue in $\mathcal{G}(n, m)$, and hence a nested split graph.*

Nested split graphs are also known as *threshold graphs* [cf. Diaconis et al., 2008; Mahadev and Peled, 1995]. From Proposition 9 it follows that the efficient graph in the class of graphs $\mathcal{G}(n)$ with n nodes must be a nested split graph. A candidate for such a graph is the star $K_{1,n-1}$, which has been studied in Durrett [2007, Lemma 4.8.2],[35] or the *nested star architecture* $F_{n,d}$ studied in König et al. [2011] maximizes welfare.

In the following we provide two examples for which we explicitly compute the asymptotic knowledge stocks. First, we consider a k-regular graph.[36]

Corollary 3. *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 $i = 1, \ldots, n$ and $k = 1, \ldots, N$. Then in the k-regular graph the asymptotic knowledge stocks are given by*

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

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

Corollary 4. *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 $i = 1, \ldots, n$ and $k = 1, \ldots, N$. Then in the star network $K_{1,n-1}$ the asymptotic knowledge*

[35] See also Berger et al. [2005]; Cator and Van Mieghem [2012].

[36] 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].

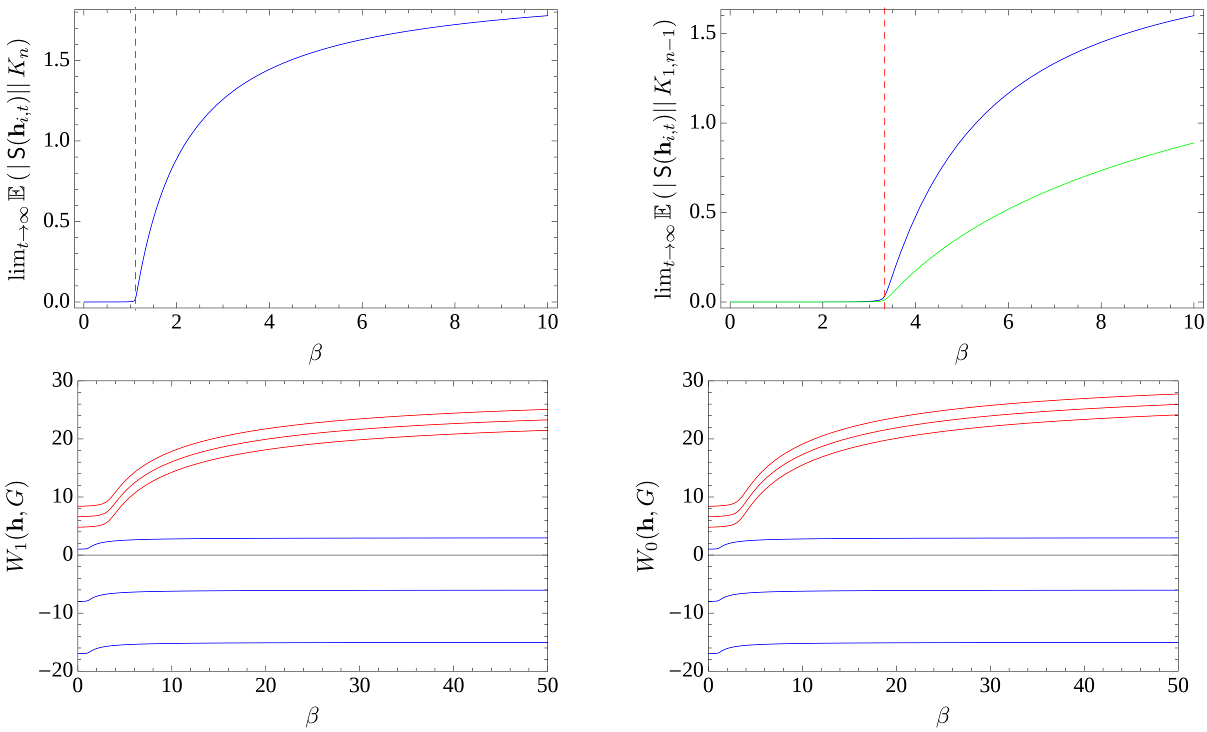


Figure 9: (Top left panel) The asymptotic knowledge stock of the firms in the complete graph K_n for varying values of β with $\gamma = 0.001$, $\lambda = 10$, $N = 2$ and $n = 10$. The critical value for the spillover parameter β is $\beta^c = \frac{\gamma+\lambda}{\mu_1(K_n)} = \frac{\gamma+\lambda}{n-1}$. (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 welfare 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 welfare 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 $\theta = 0$.

stocks are given by

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

where

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

An illustration for the asymptotic knowledge stocks is given in Figure 9 for the complete graph K_n and the star $K_{1,n-1}$.

9.2. The Competitive Case

Next, we consider the competitive case with $\theta = 0$. The value function can then be written as follows

$$V_{0,r} = \int_0^\infty e^{-rt}\mathbb{E}(W_0(\mathbf{h}_t, G)|G)dt = n\int_0^\infty e^{-rt}\mathbb{E}\left(1 + b\bar{h}_t + \frac{b^2\sigma^2_{h,t}}{1+b\bar{h}_t} - c\bar{d}_t \middle| G\right)dt.$$

In the following we consider the steady state in the limit of $t \to \infty$ and assume pairwise independence $\lim_{t\to\infty}\mathrm{Cov}(h_{ik,t}, h_{jk,t}|G) = 0$ for all $i \neq j$. Then $\bar{h}$ and σ^2_h are the sample mean and variance of the i.i.d. random variables $\{|\mathsf{S}(\mathbf{h}_i)|\}_{i=1}^n$, defined by $\bar{h} = \frac{1}{n}\sum_{i=1}^n|\mathsf{S}(\mathbf{h}_i)|$ and $\sigma^2_h = \frac{1}{n}\sum_{i=1}^n|\mathsf{S}(\mathbf{h}_i)|^2 - \bar{h}^2$. Note that, under pairwise independence and $\alpha = 0$, $|\mathsf{S}(\mathbf{h}_i)|$ follows a binomial distribution $\mathrm{Binom}(N,p)$ with success probability $p = \mathbb{E}(h_{ik})$, mean $\mu = Np$ and variance $\sigma^2 = Np(1-p)$. In the limit of large n, by the CLT we then have that

$$\sqrt{n}(\bar{h} - \mu)|G)) \sim \mathcal{N}(0, \sigma^2), \quad \text{as } n \to \infty,$$

and

$$\sqrt{n}(\sigma^2_h - \sigma^2)|G)) \sim \mathcal{N}(0, 2\sigma^4), \quad \text{as } n \to \infty.$$

Next, denote by $\boldsymbol{\delta} = (\bar{h}, \sigma^2_h)^\top$. Because the mean and variance are independent for normally distributed random variables, we can write

$$\sqrt{n}(\boldsymbol{\delta} - \mathbb{E}(\boldsymbol{\delta})) \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}), \quad \text{as } n \to \infty,$$

with

$$\boldsymbol{\Sigma} = \begin{bmatrix}\sigma^2 & 0 \\ 0 & 2\sigma^4\end{bmatrix}.$$

In the following we denote by $\Phi(\bar{h}, \sigma^2_h) = \frac{1}{n}W_0(\mathbf{h}, G)$. Since $\Phi(\bar{h}, \sigma^2_h)$ is a continuous function of $\bar{h}$ and σ^2_h we can directly apply the delta method to show that

$$\sqrt{n}(\Phi(\boldsymbol{\delta}) - \Phi(\mathbb{E}(\boldsymbol{\delta}))) \sim \mathcal{N}(\mathbf{0}, \mathbf{V}_\Phi), \quad \text{as } n \to \infty,$$

where

$$\mathbf{V}_\Phi = \frac{\partial\Phi(\mathbb{E}(\boldsymbol{\delta}))}{\partial\boldsymbol{\delta}}\boldsymbol{\Sigma}\frac{\partial\Phi(\mathbb{E}(\boldsymbol{\delta}))}{\partial\boldsymbol{\delta}^\top}.$$

In particular, this implies that

$$\lim_{n\to\infty}\mathbb{E}(\Phi(\bar{h}, \sigma^2_h)) = \Phi(\mathbb{E}(\bar{h}), \mathbb{E}(\sigma^2_h)) = \Phi(Np, Np(1-p)),$$

where $p = \mathbb{E}(h_{ik})$. It then follows that

$$\lim_{t\to\infty}\lim_{n\to\infty}\frac{1}{n}\mathbb{E}\left(1 + b\bar{h}_t + \frac{b^2\sigma^2_{h,t}}{1+b\bar{h}_t} - c\bar{d}_t \middle| G\right) = 1 + b\left(Np + \frac{1-p}{1+\frac{1}{bNp}}\right) - c\bar{d},$$

Since this is an increasing function in p,[37] asymptotic welfare attains its maximum at $p = 1$, where we get

$$\lim_{n\to\infty}\frac{1}{n}V_{0,r}(G) = o(1) + O\left(\frac{1}{r}\right)\left(1 + bN - c\bar{d}\right).$$

[37] The derivative w.r.t. p is given by $b\left(\frac{bn+1}{(bnp+1)^2} + n - 1\right) \geq 0$.

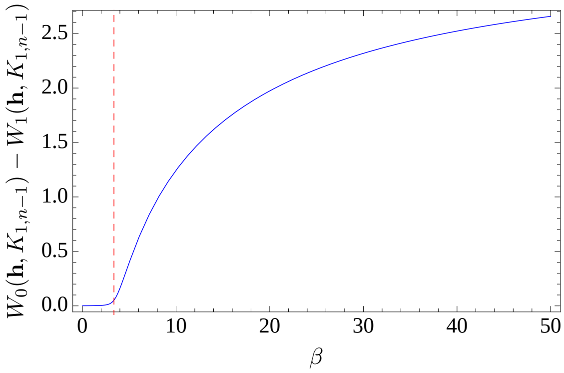


Figure 10: The welfare gain from competition in the star network $K_{1,n-1}$ with the same parameter values as in Figure 9. The threshold $\beta^c = (\lambda + \gamma)/\sqrt{n-1}$ is indicated with a dashed line.

This is identical to the case without competition (with $\theta = 1$), when $\beta \gg \lambda$ or $\beta \to 0$ and $\gamma \gg \lambda$ (cf. Proposition 8).

Figure 9 also shows asymptotic welfare 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 $\theta = 0$, welfare is increasing with decreasing cost c. Moreover, welfare for the star $K_{1,n-1}$ is always higher than in the complete graph K_n, indicating that centralization has a conducive effect on welfare. This resembles previous studies of R&D networks which abstracted away any technological dynamics [cf. Westbrock, 2010]. Moreover, observe that the variance in the knowledge stocks in the complete graph is zero, so that welfare in both cases, $\theta = 1$ and $\theta = 0$ is the same. However, welfare gains can be obtained from competition in the case of the star $K_{1,n-1}$. These gains are shown in Figure 10. We see that the gains are increasing with increasing values of β, but only after β exceeds the threshold β^c.

9.3. Welfare Gains from Competition

By considering the two polar opposite cases of independent markets ($\theta = 1$) and full competition ($\theta = 0$) we can investigate whether competition has a conducive or detrimental effect on innovation [cf. Aghion et al., 2005; Schmutzler, 2010]. In the case of $N = 2$ and $\eta \to \infty$ as well as $\eta \to 0$ we know from Propositions 7 that the stationary knowledge stocks and network density are identical for the non-competitive case ($\theta = 1$) and the competitive case $\theta = 0$, so that from Equation (48) it follows that welfare in the competitive case is higher.

Proposition 10. *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 welfare gain from competition is given by

$$W_0(\mathbf{h}, G) - W_1(\mathbf{h}, G) = \frac{b^2\left(-50\alpha^3 + \alpha\left(20\lambda(A - 5\lambda) + \beta n(A - 15\lambda) - \beta^2 n^2\right) + 5\alpha^2(2A - 40\lambda - 3\beta n) + \beta\lambda n(A - \beta n)\right)}{\alpha\beta(Ab - 5b(\alpha + 2\lambda) + (b + 1)\beta n)}, \tag{53}$$

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

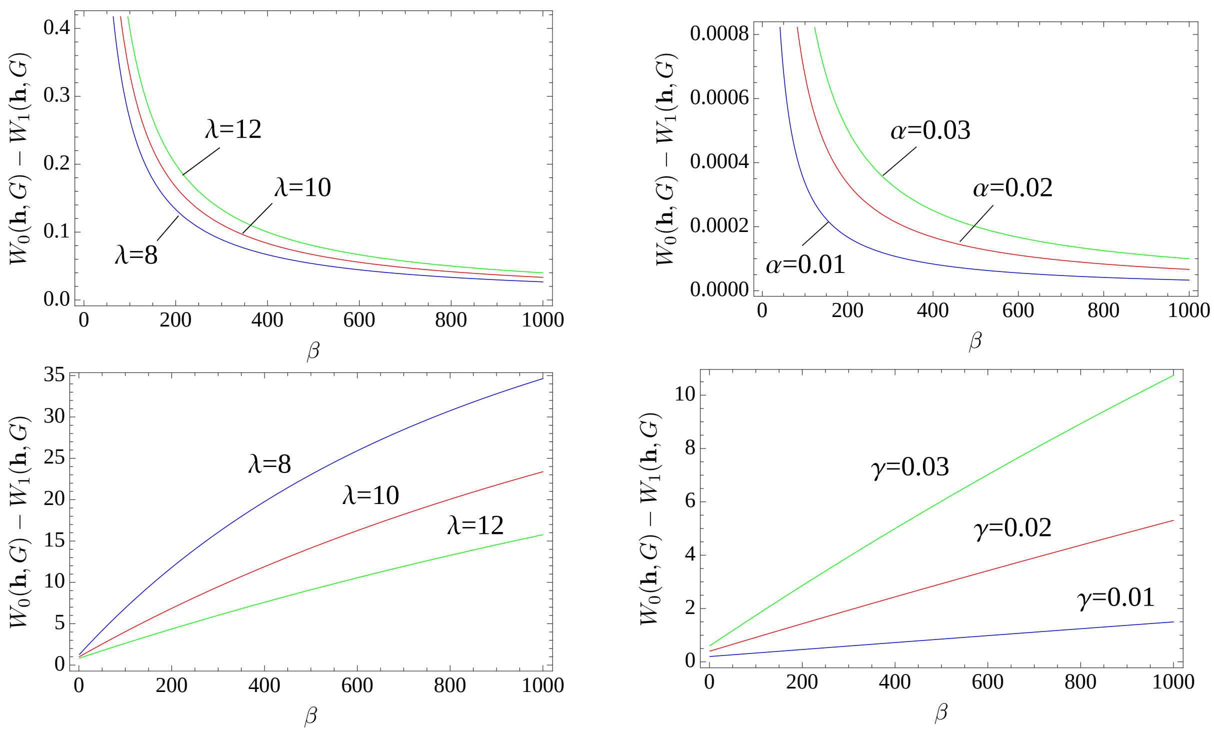


Figure 11: (Top left panel) The welfare gain from competition for different values of the knowledge obsolescence rate λ with $\alpha = 0.01$, $\gamma = 0$, $n = 100$ and $b = 1$ when $\eta \to 0$. (Top right panel) Welfare gain from competition for different values of α for the same parameter values setting $\lambda = 10$. (Bottom left panel) The welfare gain from competition for different values of the knowledge obsolescence rate λ with $\alpha = 0$, $n = 100$ and $b = 1$ when $\eta \to \infty$. (Bottom right panel) Welfare gain from competition for different values of γ for the same parameter values.

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

$$
\begin{aligned}
W_0(\mathbf{h}, G) - W_1(\mathbf{h}, G) &= \frac{1}{\beta^2\gamma^4 n + b\beta\gamma^2\left(3\alpha\gamma(\gamma+\lambda) + (\gamma+\lambda)\left(3(\gamma+\lambda)^2 - A\right) + 2\beta\gamma^2 n\right)} \\
&\times \left[b^2\left(-18\alpha^2\gamma^2(\gamma+\lambda)^2 - 6(\gamma+\lambda)^4\left(3(\gamma+\lambda)^2 - A\right) - 3\alpha\gamma\left(2(\gamma+\lambda)^2\left(6(\gamma+\lambda)^2 - A\right)\right.\right.\right. \\
&\left.\left.\left.+\beta\gamma^2 n(\gamma+2\lambda)\right) + \beta\gamma^2 n\left(A\gamma + 2A\lambda - 3(\gamma+6\lambda)(\gamma+\lambda)^2\right)\right)\right], \qquad (54)
\end{aligned}
$$

where we have denoted by

$$A = \sqrt{9\gamma^2(\alpha+\gamma)^2 + 18\gamma\lambda^2(\alpha+3\gamma) + 36\gamma\lambda^3 + 9\lambda^4 + 12\gamma^2\lambda(3\alpha+3\gamma+\beta n)}.$$

Figure 11 shows the welfare gain from competition for different values of the knowledge obsolescence rate λ, the in-house innovation rates γ and the α for both cases without profit shocks ($\eta \to \infty$) and with strong shocks ($\eta \to 0$) as a function of the spillover parameter β. The two cases can show starkly different behavior, as we find for example that the welfare gain is decreasing with strong shocks, but increasing with vanishing shocks.

10. Empirical Implications

In this section 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.[38] In order to estimate the parameters of the model we use the Likelihood-Free Markov Chain Monte Carlo (LF-MCMC) algorithm suggested by Marjoram et al. [2003].[39,40] The purpose of the LF-MCMC algorithm is to estimate the parameter vector $\boldsymbol{\delta} \equiv (\alpha, \beta, \gamma, \rho, \eta, \lambda, \tau, b, c)_{1\times L}$, $L = 9$, of the model on the basis of the summary statistics of the average degree $\mathbf{S}^{\text{o}} \equiv (\bar{d}_t^{\text{obs}})_{t=1985}^{2011}$. 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 2. *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.

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 $\boldsymbol{\Sigma}_s = \text{diag}\{\sigma_{1,s}^2, \ldots, \sigma_{L,s}^2\}$. More precisely, for each parameter $\theta_i \in \mathbb{R}_+$ we choose a proposal distribution given by

$$q_s(\delta \to \delta') = \frac{\phi(\delta' | \delta, \sigma_s^2)}{\Phi(\delta^{\max} | \delta, \sigma_s^2) - \Phi(\delta^{\min} | \delta, \sigma_s^2)} \mathbb{1}_{[\delta^{\min}, \delta^{\max}]}(\delta'),$$

where $\phi(\delta|\mu, \sigma^2)$ and $\Phi(\delta|\mu, \sigma^2)$ 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 $\alpha \in [0, 200]$, $\beta \in [0, 50]$, $\gamma \in [0, 0.025]$, $\rho \in [0, 1.5]$,

[38] In contrast, the patent data is only available for 20% of the firms, and hence any estimation relying on patents would force us to discard 80% of the data.

[39] This is essentially a simulated method of moments (SMM) estimation procedure [cf. McFadden, 1989; Pakes and Pollard, 1989].

[40] See Sisson and Fan [2011] for an introduction to LF-MCMC, Robert and Casella [2004] for a general discussion of MCMC approaches, and Chib [2001] and Chernozhukov and Hong [2003] for applications of MCMC in econometrics.

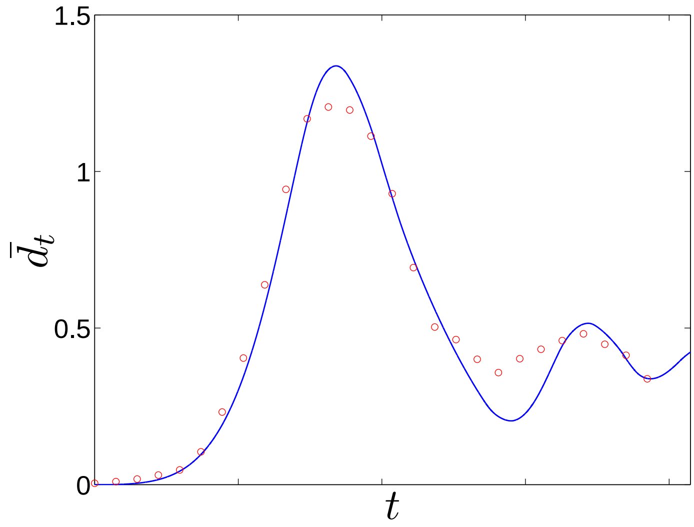


Figure 12: Comparison of the average degree $\bar{d}$ from the prediction of the theoretical model for $N = 2$ indicated with a line and the empirical observations indicated with circles. The parameters used are $\theta = 0$, $\alpha = 0$, $\beta = 700$, $b = 69$, $c = 2900$, $\gamma = 1$, $\eta = 0.01$, $\lambda = 400$, $n = 500$, $\rho = 1/\tau + 300$ and $\tau = 0.01$. The initial condition is $x_0(0) = 0.75$, $\bar{x}_0(1) = 0.125$, $\bar{x}_0(2) = 0$ and an empty network.

$\eta \in= [0, 5]$, $\lambda \in [0, 50]$, $\tau \in [0, 0.05]$, $b \in [0, 500]$ and $c \in [0, 5]$. The parameters $\epsilon^{\min}$ are choose sufficiently small after long experimentation with different starting values and burn-in periods.

The parameter estimates from the above procedure are shown in Table 2. Moreover, Figure 12 shows a comparison of the average degree from the theoretical model for $N = 2$ and the empirical observations. The parameters used are $\theta = 0$, $\alpha = 0$, $\beta = 700$, $b = 69$, $c = 2900$, $\gamma = 1$, $\eta = 0.01$, $\lambda = 400$, $n = 500$, $\rho = 1/\tau + 300$ and $\tau = 0.01$. The initial condition is $x_0(0) = 0.75$, $\bar{x}_0(1) = 0.125$, $\bar{x}_0(2) = 0$ and an empty network.

We can further infer the knowledge stocks of the firms from the observed network for every year by using Equation (50), or by using a continued fraction expansion approximation as in Equation (52) with the estimated parameter values from Table 2. This allows us to compute asymptotic welfare. We can further compute welfare for a star network with the same number of firms as in the observed network from Corollary 4. The relative welfare gains from imposing the star network structure over the different years of observation are shown in Figure 13. We find that welfare can be improved by up to 48%. This indicates the possibility for policy makers to improve the innovativeness of an economy considerably from not only fostering R&D collaborations per se but also paying attention to the overall R&D collaboration network structure.

11. Conclusion

In this paper we study the co-evolutionary dynamics of firms' technology portfolios and the formation of R&D collaborations that influence and get influenced by these technology portfolios. We investigate the stationary states of this dynamics process, and show that there exists a critical level for the technology spillover parameter below which no significant innovation takes place in the economy. Moreover, we analyze the impact of competition on innovation and R&D network formation, and find that in general competition is welfare increasing. This is due to the fact that competition leads to reallocation and the replacement of less productive firms with more centrally located ones in the network that also tend to be more productive. We further identify the efficient network structure as a nested split graph, which is characterized by a core periphery structure. The stability analysis of

Table 2: Estimation of the model parameters $\delta \in \boldsymbol{\delta} \equiv (\alpha, \beta, \gamma, \rho, \eta, \lambda, \tau, b, c)$ 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	112.6535	48.7623	9.1379	8279.6846	0.8612
β	802.4149	166.5676	34.0604	20089.8680	0.6556
γ	0.5054	0.2458	0.0171	985.8151	0.7242
ρ	366.5746	150.7862	31.8503	22590.2242	0.2957
η	0.1734	0.1147	0.0237	24211.7205	0.4745
λ	299.6674	107.8416	21.6602	13774.0216	0.9460
τ	0.0254	0.0220	0.0035	15998.2589	0.0766
b	46.5475	24.1763	4.5532	22727.7992	0.2348
c	2745.4919	177.2581	36.3354	16860.3328	0.9309
n			500		
S			200000		

[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 reasonably high values of $p_\delta(S)$ were obtained.

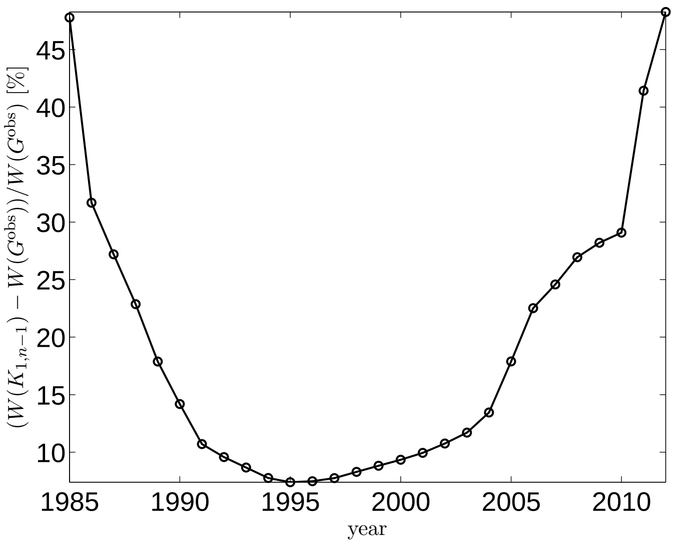


Figure 13: Relative percentage increase in welfare $(W(K_{1,n-1}) - W(G^{\text{obs}}))/W(G^{\text{obs}})$ from imposing a star network $K_{1,n-1}$ as compared to the observed network G^{obs}.

our model indicates that the R&D collaboration intensity can exhibit a cyclical pattern, which can be described as a damped oscillation. We confirm this novel observation using 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. Finally, we indicate the potential welfare loss incurred in the empirically observed networks.

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., Ozdaglar, A., and Yildiz, E. (2011). Diffusion of innovations in social networks. In *Decision and Control and European Control Conference (CDC-ECC), 2011 50th IEEE Conference on*, pages 2329–2334. IEEE.

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

Aghion, P., Akcigit, U., and Howitt, P. (2013). What do we learn from schumpeterian growth theory? *NBER Working Paper No. 18824*.

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.

Aghion, P. and Howitt, P. (2009). *The economics of growth*. MIT Press.

Amir, R., Evstigneev, I., and Wooders, J. (2003). Noncooperative versus cooperative R&D with endogenous spillover rates. *Games and Economic Behavior*, 42(2):183–207.

Amir, R. and Wooders, J. (2000). One-way spillovers, endogenous innovator/imitator roles, and research joint ventures. *Games and Economic Behavior*, 31(1):1–25.

Andergassen, R., Nardini, F., and Ricottilli, M. (2006). Innovation waves, self-organized criticality and technological convergence. *Journal of Economic Behavior & Organization*, 61(4):710–728.

Andersen, B. (1999). The hunt for s-shaped growth paths in technological innovation: a patent study. *Journal of Evolutionary Economics*, 9(4):487–526.

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.

Anderson, P. and Tushman, M. (1990). Technological discontinuities and dominant designs: A cyclical model of technological change. *Administrative science quarterly*, 35(4):604–633.

Anderson, R. M., May, R. M., and Anderson, B. (1992). *Infectious diseases of humans: dynamics and control*, volume 28. Wiley.

Ballester, C., Calvó-Armengol, A., and Zenou, Y. (2006). Who's who in networks. wanted: The key player. *Econometrica*, 74(5):1403–1417.

Barlevy, G. (2007). On the cyclicality of research and development. *The American Economic Review*, 97(4):1131–1164.

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.

Belhaj, M., Bervoets, S., and Deroian, F. (2013). Optimal networks in games with local complementarities. *Working Paper, Aix-Marseille School of Economics*.

Berger, N., Borgs, C., Chayes, J., and Saberi, A. (2005). On the spread of viruses on the internet. In *Proceedings of the sixteenth annual ACM-SIAM symposium on Discrete algorithms*, pages 301–310. Society for Industrial and Applied Mathematics.

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

Berliant, M. and Fujita, M. (2009). Dynamics of knowledge creation and transfer: The two person case. *International Journal of Economic Theory*, 5(2):155–179.

Berliant, M. and Fujita, M. (2011). The dynamics of knowledge diversity and economic growth. *Southern Economic Journal*, 77(4):856.

Bernard, A. B., Redding, S. J., and Schott, P. K. (2011). Multiproduct firms and trade liberalization. *The Quarterly Journal of Economics*, 126(3):1271–1318.

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

Bloch, F. (1997). Non-cooperative models of coalition formation in games with spillovers. *New directions in the economic theory of the environment*, 25:311.

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.

Bonacich, P. (1987). Power and centrality: A family of measures. *American Journal of Sociology*, 92(5):1170.

Brooks, S. and Roberts, G. (1998). Assessing convergence of markov chain monte carlo algorithms. *Statistics and Computing*, 8(4):319–335.

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.

Cator, E. and Van Mieghem, P. (2012). Susceptible infected susceptible epidemics on the complete graph and the star graph: Exact analysis. *Mimeo Delft University of Technology*.

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.

Chatterjee, S. and Durrett, R. (2009). Contact processes on random graphs with power law degree distributions have critical value 0. *The Annals of Probability*, 37(6):2332–2356.

Chernozhukov, V. and Hong, H. (2003). An MCMC approach to classical estimation. *Journal of Econometrics*, 115(2):293–346.

Chib, S. (2001). Markov chain monte carlo methods: computation and inference. *Handbook of econometrics*, 5:3569–3649.

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

Corchón, L. (2007). The theory of contests: a survey. *Review of Economic Design*, 11(2):69–100.

Cowan, R. and Jonard, N. (2004). Network structure and the diffusion of knowledge. *Journal of Economic Dynamics and Control*, 28(8):1557–1575.

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

Dasgupta, P. and Stiglitz, J. (1980). Uncertainty, industrial structure, and the speed of R&D. *The Bell Journal of Economics*, 11(1):1–28.

Dasgupta, P. and Stiglitz, J. (1981). Resource depletion under technological uncertainty. *Econometrica*, 49(1):85–104.

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.

Dembo, A. and Montanari, A. (2010). Ising models on locally tree-like graphs. *The Annals of Applied Probability*, 20(2):565–592.

Desai, M. and Ormerod, P. (1998). Richard goodwin: a short appreciation. *The Economic Journal*, 108(450):1431–1435.

Diaconis, P., Holmes, S., and Janson, S. (2008). Threshold graph limits and random threshold graphs. *Internet Mathematics*, 5(3):267–320.

Do, A.-L. and Gross, T. (2009). Contact processes and moment closure on adaptive networks. In *Adaptive Networks*, pages 191–208. Springer.

Durrett, R. (2007). *Random Graph Dynamics*. Cambridge University Press.

Ehrhardt, G., Marsili, M., and Vega-Redondo, F. (2008). Networks emerging in a volatile world. Economics Working Papers.

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

Fai, F. and Von Tunzelmann, N. (2001). Industry-specific competencies and converging technological systems: evidence from patents. *Structural Change and Economic Dynamics*, 12(2):141–170.

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.

Franke, R. (2001). Wave trains, innovation noise, and long waves. *Journal of Economic Behavior &*

Organization, 45(1):49–68.
Friedkin, N. (1983). Horizons of observability and limits of informal control in organizations. *Social Forces*, 62(1):54–77.
Fullerton, R. L. and McAfee, R. P. (1999). Auctionin entry into tournaments. *Journal of Political Economy*, 107(3):573–605.
Futia, C. (1980). Schumpeterian competition. *The Quarterly Journal of Economics*, 94(4):675–695.
Galí, J. (1999). Technology, employment, and the business cycle: Do technology shocks explain aggregate fluctuations? *American Economic Review*, 89(1):249–271.
Gambardella, A. and Torrisi, S. (1998). Does technological convergence imply convergence in markets? Evidence from the electronics industry. *Research Policy*, 27(5):445–463.
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*, pages 1457–1475.
Geroski, P. and Walters, C. (1995). Innovative activity over the business cycle. *The Economic Journal*, pages 916–928.
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.
Giummo, J. (2010). German employee inventors compensation records: A window into the returns to patented inventions. *Research Policy*, 39(7):969–984.
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. (2002). Joint venture cycles: The evolution of ownership strategies of US MNEs, 1945–75. *Cooperative strategies in international business: Joint ventures and technology partnerships between firms*, page 111.
Goodwin, R. (1946). Innovations and the irregularity of economic cycles. *The Review of Economics and Statistics*, 28(2):95–104.
Goodwin, R. (1951). The nonlinear accelerator and the persistence of business cycles. *Econometrica*, pages 1–17.
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.
Griliches, Z. (1957). Hybrid corn: An exploration in the economics of technological change. *Econometrica*, pages 501–522.
Grimmett, G. (2010). *Probability on graphs*. Cambridge University Press.
Grimmett, G. and Stirzaker, D. (2001). *Probability and random processes*. Oxford University Press.
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 Frankort, H. (2008). The gloomy side of embeddedness: The effects of overembeddedness on inter-firm partnership formation. *Advances in Strategic Management*, 25:503–30.
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.
Hall, B., Jaffe, A., 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.

Harrigan, K. (1988). Strategic alliances and partner asymmetries. *Management international review*, 28(4):53–72.

Holly, S. and Petrella, I. (2012). Factor demand linkages, technology shocks, and the business cycle. *Review of Economics and Statistics*, 94(4):948–963.

Horn, R. A. and Johnson, C. R. (1990). *Matrix Analysis*. Cambridge 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. (2007). Relating network structure to diffusion properties through stochastic dominance. *The BE Journal of Theoretical Economics*, 7(1):1–13.

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*, pages 24–52.

Jovanovic, B. and Rob, R. (1989). The growth and diffusion of knowledge. *The Review of Economic Studies*, 56(4):569–582.

Jovanovic, B. and Rob, R. (1990). Long waves and short waves: Growth through intensive and extensive search. *Econometrica*, pages 1391–1409.

Jovanovic, B. and Rousseau, P. L. (2005). General purpose technologies. *Handbook of economic growth*, 1:1181–1224.

Kaldor, N. (1940). A model of the trade cycle. *The Economic Journal*, pages 78–92.

Keeling, M. J. and Eames, K. T. (2005). Networks and epidemic models. *Journal of the Royal Society Interface*, 2(4):295–307.

Kelly, M. (2001). Linkages, thresholds, and development. *Journal of Economic Growth*, 6(1):39–53.

Kelly, M. (2009). Technological progress under learning by imitation. *International Economic Review*, 50(2):397–414.

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.

Klepper, S. (1996). Entry, exit, growth, and innovation over the product life cycle. *The American economic review*, pages 562–583.

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.

Kogut, B., Urso, P., and Walker, G. (2007). Emergent properties of a new financial market: American venture capital syndication, 1960–2005. *Management Science*, 53(7):1181–1198.

König, M. D. (2011). The formation of networks with local spillovers and limited observability. *SIEPR Discussion Paper 11–004, Stanford Institute for Economic Policy Research.*

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., Battiston, S., Napoletano, M., and Schweitzer, F. (2011). Recombinant knowledge and the evolution of innovation networks. *Journal of Economic Behavior & Organization*, 79(3):145–164.

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. (2012). Innovation vs. imitation and the evolution of productivity distributions. *CEPR Discussion Paper No. 8843.*

König, M. D., Tessone, C., and Zenou, Y. (2013). Nestedness in networks: A theoretical model and some applications. *Forthcoming in Theoretical Economics.*

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.

Kurtz, T. (1971). Limit theorems for sequences of jump Markov processes approximating ordinary differential processes. *Journal of Applied Probability*, 8(2):344–356.

Kuznets, S. (1940). Schumpeter's business cycles. *American Economic Review*, pages 257–271.
Lazear, E. (2004). Balanced skills and entrepreneurship. *American Economic Review*, pages 208–211.
Levin, D. A., Peres, Y., and Wilmer, E. L. (2009). *Markov chains and mixing times*. American Mathematical Society.
Lopez-Pintado, D. (2008). Diffusion in complex social networks. *Games and Economic Behavior*, 62(2):573 – 590.
Loury, G. C. (1979). Market structure and innovation. *The Quarterly Journal of Economics*, 93(3):395–410.
Lucas, R. and Moll, B. (2011). Knowledge growth and the allocation of time. University of Chicago Working Paper.
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.
McFadden, D. (1989). A method of simulated moments for estimation of discrete response models without numerical integration. *Econometrica*, pages 995–1026.
Meagher, K. and Rogers, M. (2004). Network density and R&D spillovers. *Journal of Economic Behavior & Organization*, 53(2):237–260.
Meland, F. and Straume, O. R. (2007). Outsourcing in contests. *Public Choice*, 131(3-4):315–331.
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.
Nesta, L. and Dibiaggio, L. (2003). Technology Strategy and Knowledge Dynamics: The Case of Biotech. *Industry & innovation*, 10(3):331–349.
Newman, M. (2010). *Networks: An Introduction*. Oxford University Press.
Pakes, A. and Pollard, D. (1989). Simulation and the asymptotics of optimization estimators. *Econometrica*, pages 1027–1057.
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.
Pastor-Satorras, R. and Vespignani, A. (2001). Epidemic spreading in scale-free networks. *Physical review letters*, 86(14):3200–3203.
Patel, P. and Pavitt, K. (1997). The technological competencies of the world's largest firms: Complex and path-dependent, but not much variety. *Research Policy*, 26(2):141–156.
Perla, J. and Tonetti, C. (2012). Equilibrium imitation and growth. *SSRN 1866460*.
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.
Robert, C. and Casella, G. (2004). *Monte Carlo statistical methods*. Springer Verlag.
Roketskiy, N. (2011). Competition and networks of collaboration. *NYU Working Paper*.
Rosenkopf, L. and Padula, G. (2008). Investigating the Microstructure of Network Evolution: Alliance Formation in the Mobile Communications Industry. *Organization Science*, 19(5):669.
Sandholm, W. (2010). *Population games and evolutionary dynamics*. MIT Press.
Sannikov, Y. and Skrzypacz, A. (2010). The role of information in repeated games with frequent actions. *Econometrica*, 78(3):847–882.
Schilling, M. (2009). Understanding the alliance data. *Strategic Management Journal*, 30(3):233–260.
Schmutzler, A. (2010). Is competition good for innovation? a simple approach to an unresolved question. *Foundations and Trends in Microeconomics*, 5(6):355–428.
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.
Sisson, S. and Fan, Y. (2011). *Handbook of Markov Chain Monte Carlo*, chapter Likelihood-free markov chain monte carlo. Chapman & Hall/CRC Handbooks of Modern Statistical Methods.
Sokal, A. (1996). *Monte Carlo Methods in Statistical Mechanics: Foundations and New Algorithms*.
Stanley, H. E. (1968). Dependence of critical properties on dimensionality of spins. *Phys. Rev. Lett.*,

20:589–592.
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.
Tullock, G. (1980). *Toward a theory of the rent-seeking society*, volume 97, chapter Efficient rent seeking, page 112. Texas A & M University Press.
Van Mieghem, P., Omic, J., and Kooij, R. (2009). Virus spread in networks. *IEEE/ACM Transactions on Networking*, 17(1):1–14.
Veselić, K. (1979). On real eigenvalues of real tridiagonal matrices. *Linear Algebra and its Applications*, 27:167–171.
Wang, Y., Chakrabarti, D., Wang, C., and Faloutsos, C. (2003). Epidemic spreading in real networks: An eigenvalue viewpoint. In *Proceedings. 22nd International Symposium on Reliable Distributed Systems, 2003*, pages 25–34. IEEE.
Watts, D. J. and Strogatz, S. H. (1998). Collective dynamics of small-world networks. *Nature*, 393:440–442.
Westbrock, B. (2010). Natural concentration in industrial research collaboration. *The RAND Journal of Economics*, 41(2):351–371.
Wormald, N. C. (1995). Differential equations for random processes and random graphs. *The annals of applied probability*, pages 1217–1235.
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. 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 (18) in Theorem 1.

Proposition 11. *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* $\xi_t(\mathbf{h},\mathbf{h}') \equiv \mathbb{P}(a_{ij,t}=1|\mathbf{h}_{it}=\mathbf{h},\mathbf{h}_{jt}=\mathbf{h}')$*, 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 (19),* $x_t(\mathbf{h})$ *converges in probability to the solution of Equation (17), and* $\xi_t(\mathbf{h},\mathbf{h}')$ *converges in probability to the solution of the ODE*

$$\begin{aligned}
\frac{d\xi_t(\mathbf{h},\mathbf{h}')}{dt} &= \rho g(\mathbf{h},\mathbf{h}')\left(1-\xi_t(\mathbf{h},\mathbf{h}')\right) - \rho\,\xi_t(\mathbf{h},\mathbf{h}') \\
&+ \sum_{k=1}^{N} \mathbb{1}_{\{h_k=1\}}\left(\gamma+\alpha\langle\mathbf{h}-\mathbf{e}_k,\mathbf{u}\rangle\right)\frac{x_t(\mathbf{h}-\mathbf{e}_k)}{x_t(\mathbf{h})}\left(\xi_t(\mathbf{h}-\mathbf{e}_k,\mathbf{h}')-\xi_t(\mathbf{h},\mathbf{h}')\right) \\
&+ \sum_{k=1}^{N} \mathbb{1}_{\{h'_k=1\}}\left(\gamma+\alpha\langle\mathbf{h}'-\mathbf{e}_k,\mathbf{u}\rangle\right)\frac{x_t(\mathbf{h}'-\mathbf{e}_k)}{x_t(\mathbf{h}')}\left(\xi_t(\mathbf{h},\mathbf{h}'-\mathbf{e}_k)-\xi_t(\mathbf{h},\mathbf{h}')\right) \\
&+ \lambda\sum_{k=1}^{N} \mathbb{1}_{\{h_k=0\}}\frac{x_t(\mathbf{h}+\mathbf{e}_k)}{x_t(\mathbf{h})}\left(\xi_t(\mathbf{h}+\mathbf{e}_k,\mathbf{h}')-\xi_t(\mathbf{h},\mathbf{h}')\right) \\
&+ \lambda\sum_{k=1}^{N} \mathbb{1}_{\{h'_k=0\}}\frac{x_t(\mathbf{h}'+\mathbf{e}_k)}{x_t(\mathbf{h}')}\left(\xi_t(\mathbf{h},\mathbf{h}'+\mathbf{e}_k)-\xi_t(\mathbf{h},\mathbf{h}')\right) \\
&+ \beta\sum_{k=1}^{N} \mathbb{1}_{\{h_k=1\}}\left(\xi_t(\mathbf{h}-\mathbf{e}_k,\mathbf{h}')-\xi_t(\mathbf{h},\mathbf{h}')\right)\sum_{\mathbf{h}''\in\mathcal{H}^N:h''_k=1}\xi_t(\mathbf{h}-\mathbf{e}_k,\mathbf{h}'')\frac{x_t(\mathbf{h}-\mathbf{e}_k)x_t(\mathbf{h}'')}{x_t(\mathbf{h})} \\
&+ \beta\sum_{k=1}^{N} \mathbb{1}_{\{h'_k=1\}}\left(\xi_t(\mathbf{h}'-\mathbf{e}_k,\mathbf{h})-\xi_t(\mathbf{h},\mathbf{h}')\right)\sum_{\mathbf{h}''\in\mathcal{H}^N:h''_k=1}\xi_t(\mathbf{h}'-\mathbf{e}_k,\mathbf{h}'')\frac{x_t(\mathbf{h}'-\mathbf{e}_k)x_t(\mathbf{h}'')}{x_t(\mathbf{h}')}. \qquad (55)
\end{aligned}$$

Observe that Equation (55) is of the form of Equation (18) as ρ becomes large, independently of the approximation in Equation (19).

Proposition 12. *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* $\bar{\xi}_t(s,s')$ *for any* $0 \le s,s' \le N$ *defined as in Equations (20) and (21). Then, under the pair approximation of Equation (19),* $\bar{x}_t(s)$ *converges in probability to the solution of Equation (23), and* $\bar{\xi}_t(s,s')$ *converges in probability to the solution of the ODE*

$$\begin{aligned}
\frac{d\bar{\xi}_t(s,s')}{dt} &= \rho\bar{g}(s,s') - \rho(1+\bar{g}(s,s'))\bar{\xi}_t(s,s') \\
&+ \frac{\bar{x}_t(s-1)}{\bar{x}_t(s)}s\left(\bar{\xi}_t(s-1,s')-\bar{\xi}_t(s,s')\right)\left[(\gamma+\alpha(s-1))+\beta\sum_{s''=1}^{N}\binom{N}{s''-1}\bar{\xi}_t(s-1,s'')\bar{x}_t(s'')\right] \\
&+ \frac{\bar{x}_t(s'-1)}{\bar{x}_t(s')}s'\left(\bar{\xi}_t(s'-1,s)-\bar{\xi}_t(s',s)\right)\left[(\gamma+\alpha(s'-1))+\beta\sum_{s''=1}^{N}\binom{N}{s''-1}\bar{\xi}_t(s'-1,s'')\bar{x}_t(s'')\right] \\
&+ \lambda\frac{\bar{x}_t(s+1)}{\bar{x}_t(s)}(N-s)\left(\bar{\xi}_t(s+1,s')-\bar{\xi}_t(s,s')\right)+\lambda\frac{\bar{x}_t(s'+1)}{\bar{x}_t(s')}(N-s')\left(\bar{\xi}_t(s'+1,s)-\bar{\xi}_t(s',s)\right). \qquad (56)
\end{aligned}$$

A.1. The Case of $N = 1$

In the case of $N = 1$ where $s \in \{0, 1\}$ we obtain from Equation (23)

$$\begin{aligned} \frac{d\bar{x}_t(1)}{dt} &= \gamma \bar{x}_t(0) - \lambda \bar{x}_t(1) + \beta \bar{\xi}_t(0,1)\bar{x}_t(0)\bar{x}_t(1), \\ \frac{d\bar{x}_t(0)}{dt} &= \lambda \bar{x}_t(1) - \gamma \bar{x}_t(0) - \beta \bar{\xi}_t(0,1)\bar{x}_t(0)\bar{x}_t(1), \end{aligned} \tag{57}$$

and from Equation (56) we get

$$\begin{aligned} \frac{d\bar{\xi}_t(0,1)}{dt} &= \rho \bar{g}(0,1) - \rho(1 + \bar{g}(0,1))\bar{\xi}_t(0,1) \\ &+ \frac{\bar{x}_t(0)}{\bar{x}_t(1)}(\bar{\xi}_t(0,0) - \bar{\xi}_t(0,1)) \left(\gamma + \beta \bar{\xi}_t(0,1)\bar{x}_t(1)\right) + \lambda \frac{\bar{x}_t(1)}{\bar{x}_t(0)}(\bar{\xi}_t(1,1) - \bar{\xi}_t(0,1)), \end{aligned} \tag{58}$$

and

$$\frac{d\bar{\xi}_t(0,0)}{dt} = \rho \bar{g}(0,0) - \rho(1 + \bar{g}(0,0))\bar{\xi}_t(0,0) + 2\lambda \frac{\bar{x}_t(1)}{\bar{x}_t(0)}(\bar{\xi}_t(1,0) - \bar{\xi}_t(0,0)), \tag{59}$$

and

$$\frac{d\bar{\xi}_t(1,1)}{dt} = \rho \bar{g}(1,1) - \rho(1 + \bar{g}(1,1))\bar{\xi}_t(1,1) + 2\frac{\bar{x}_t(0)}{\bar{x}_t(1)}(\bar{\xi}_t(0,1) - \bar{\xi}_t(1,1)) \left(\gamma + \beta \bar{\xi}_t(0,1)\bar{x}_t(1)\right), \tag{60}$$

with

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

where the average stock of knowledge is given by $\bar{h}_t(x) = \bar{x}_t(1)$, and

$$\bar{g}(0,0) = \bar{g}(1,1) = \frac{e^{-2\eta c}}{(1 + e^{-\eta c})^2}, \tag{61}$$

for both cases $\theta = 0$ and $\theta = 1$. 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 Definition 1 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 (57)–(60) superimposed is shown in Figure 14.

A.2. The Case of $N = 2$

When $N = 2$ with $s \in \{0, 1, 2\}$ we obtain from Equation (23)

$$\begin{aligned} \frac{d\bar{x}_t(0)}{dt} &= 2\left(\lambda \bar{x}_t(1) - \gamma \bar{x}_t(0) - \beta(\bar{\xi}_t(0,1)\bar{x}_t(0)\bar{x}_t(1) + \bar{\xi}_t(0,2)\bar{x}_t(0)\bar{x}_t(2))\right) \\ \frac{d\bar{x}_t(1)}{dt} &= \gamma \bar{x}_t(0) + \lambda \bar{x}_t(2) - (\lambda + \gamma + \alpha)\bar{x}_t(1) + \beta\left(\bar{\xi}_t(0,1)\bar{x}_t(0)\bar{x}_t(1) - \bar{\xi}_t(1,1)\bar{x}_t(1)^2\right. \\ &\left. + \bar{\xi}_t(0,2)\bar{x}_t(0)\bar{x}_t(2) - \bar{\xi}_t(1,2)\bar{x}_t(1)\bar{x}_t(2)\right) \\ \frac{d\bar{x}_t(2)}{dt} &= 2\left((\gamma + \alpha)\bar{x}_t(1) - \lambda \bar{x}_t(2) + \beta\left(\bar{\xi}_t(1,1)\bar{x}_t(1)^2 + \bar{\xi}_t(1,2)\bar{x}_t(1)\bar{x}_t(2)\right)\right), \end{aligned} \tag{62}$$

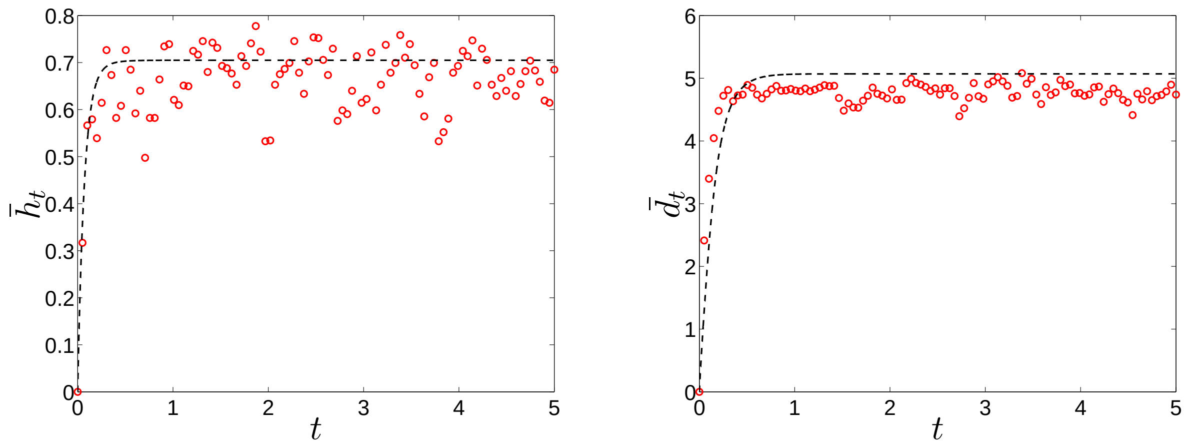


Figure 14: An example of a numerical simulation of the stochastic process introduced in Definition 1 for $N = 1$ using the "next reaction method" for simulating a continuous time Markov chain [cf. Anderson, 2012; Gibson and Bruck, 2000], and the solution of the ODEs in Equations (57)–(60) shown with a dashed line.

and from Equation (56) we obtain

$$\begin{aligned}\frac{d\bar{\xi}_t(0,1)}{dt} &= \frac{1}{2}\rho\tilde{g}(0,1) - \rho\left(1+\frac{1}{2}\tilde{g}(0,1)\right)\bar{\xi}_t(0,1) \\ &+ \frac{\bar{x}_t(0)}{\bar{x}_t(1)}\left(\bar{\xi}_t(0,0)-\bar{\xi}_t(1,0)\right)\left[\gamma+\beta\left(\bar{\xi}_t(0,1)\bar{x}_t(1)+2\bar{\xi}_t(0,2)\bar{x}_t(2)\right)\right] \\ &+ \lambda\frac{\bar{x}_t(1)}{\bar{x}_t(0)}2\left(\bar{\xi}_t(1,1)-\bar{\xi}_t(0,1)\right)+\lambda\frac{\bar{x}_t(2)}{\bar{x}_t(1)}\left(\bar{\xi}_t(2,0)-\bar{\xi}_t(1,0)\right), \end{aligned} \tag{63}$$

$$\begin{aligned}\frac{d\bar{\xi}_t(0,2)}{dt} &= \rho\tilde{g}(0,2) - \rho\left(1+\tilde{g}(0,2)\right)\bar{\xi}_t(0,2) \\ &+ 2\frac{\bar{x}_t(1)}{\bar{x}_t(2)}\left(\bar{\xi}_t(1,0)-\bar{\xi}_t(2,0)\right)\left[\gamma+\alpha+\beta\left(\bar{\xi}_t(1,1)\bar{x}_t(1)+2\bar{\xi}_t(1,2)\bar{x}_t(2)\right)\right] \\ &+ 2\lambda\frac{\bar{x}_t(1)}{\bar{x}_t(0)}\left(\bar{\xi}_t(1,2)-\bar{\xi}_t(0,2)\right), \end{aligned} \tag{64}$$

$$\begin{aligned}\frac{d\bar{\xi}_t(1,1)}{dt} &= \frac{1}{4}\rho\tilde{g}(1,1) - \rho\left(1+\frac{1}{4}\tilde{g}(1,1)\right)\bar{\xi}_t(1,1) \\ &+ 2\frac{\bar{x}_t(0)}{\bar{x}_t(1)}\left(\bar{\xi}_t(0,1)-\bar{\xi}_t(1,1)\right)\left[\gamma+\beta\left(\bar{\xi}_t(0,1)\bar{x}_t(1)+2\bar{\xi}_t(0,2)\bar{x}_t(2)\right)\right] \\ &+ 2\lambda\frac{\bar{x}_t(2)}{\bar{x}_t(1)}\left(\bar{\xi}_t(2,1)-\bar{\xi}_t(1,1)\right), \end{aligned} \tag{65}$$

$$
\begin{aligned}
&\frac{d\bar{\xi}_t(1,2)}{dt} = \frac{1}{2}\rho\tilde{g}(1,2) - \rho\left(1+\frac{1}{2}\tilde{g}(1,2)\right)\bar{\xi}_t(1,2) \\
&+ \frac{\bar{x}_t(0)}{\bar{x}_t(1)}\left(\bar{\xi}_t(0,2)-\bar{\xi}_t(1,2)\right)\left[\gamma+\beta\left(\bar{\xi}_t(0,1)\bar{x}_t(1)+2\bar{\xi}_t(0,2)\bar{x}_t(2)\right)\right] \\
&+ 2\frac{\bar{x}_t(1)}{\bar{x}_t(2)}\left(\bar{\xi}_t(1,1)-\bar{\xi}_t(2,1)\right)\left[\gamma+\alpha+\beta\left(\bar{\xi}_t(1,1)\bar{x}_t(1)+2\bar{\xi}_t(1,2)\bar{x}_t(2)\right)\right] \\
&+ \lambda\frac{\bar{x}_t(2)}{\bar{x}_t(1)}\left(\bar{\xi}_t(2,2)-\bar{\xi}_t(1,2)\right),
\end{aligned} \tag{66}
$$

$$
\begin{aligned}
&\frac{d\bar{\xi}_t(0,0)}{dt} = \rho\bar{g}(0,0') - \rho(1+\bar{g}(0,0'))\tilde{\xi}_t(0,0') \\
&+ 4\lambda\frac{\bar{x}_t(1)}{\bar{x}_t(0)}\left(\bar{\xi}_t(1,0)-\bar{\xi}_t(0,0)\right),
\end{aligned} \tag{67}
$$

$$
\begin{aligned}
&\frac{d\bar{\xi}_t(2,2)}{dt} = \rho\bar{g}(2,2) - \rho(1+\bar{g}(2,2))\tilde{\xi}_t(2,2) \\
&+ 4\frac{\bar{x}_t(1)}{\bar{x}_t(2)}\left(\bar{\xi}_t(1,2)-\bar{\xi}_t(2,2)\right)\left[\gamma+\alpha+\beta\left(\bar{\xi}_t(1,1)\bar{x}_t(1)+2\bar{\xi}_t(1,2)\bar{x}_t(2)\right)\right].
\end{aligned} \tag{68}
$$

From the definition in Equation (22) we find that

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

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

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

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

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

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

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

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

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

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

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

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

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

An example of a numerical simulation of the stochastic process introduced in Definition 1 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 (62)–(68) superimposed is shown in Figure 15.

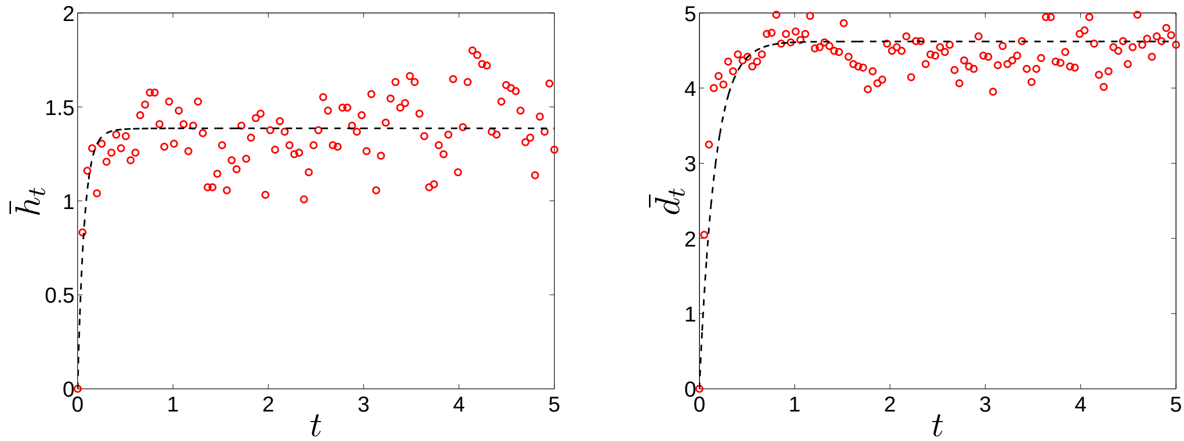


Figure 15: An example of a numerical simulation of the stochastic process introduced in Definition 1 for $N = 2$ using the "next reaction method" for simulating a continuous time Markov chain [cf. Anderson, 2012; Gibson and Bruck, 2000], and the solution of the ODEs in Equations (62)–(68) shown with a dashed line.

B. 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

$$
\begin{aligned}
\mathbb{E}_t\left(|\mathsf{S}(\mathbf{h}_{i,t+\tau})||\mathbf{h}_t, G_t\right) &= \mathbb{E}_t\left(\sum_{k=1}^{N} \mathbb{1}_{\{h_{ik,t+\tau}=1\}} \middle| \mathbf{h}_t, G_t\right) \\
&= \sum_{k=1}^{N} \mathbb{P}(h_{ik,t+\tau}=1|\mathbf{h}_t, G_t) \\
&= \sum_{k=1}^{N} \left(\nu_{ik,t}\tau \mathbb{1}_{\{h_{ik,t}=0\}} + (1-\lambda\tau)\mathbb{1}_{\{h_{ik,t}=1\}}\right) + o(\tau) \\
&= |\mathsf{S}(\mathbf{h}_{i,t})| + \sum_{k=1}^{N} \left(\nu_{ik,t}\mathbb{1}_{\{h_{ik,t}=0\}} - \lambda \mathbb{1}_{\{h_{ik,t}=1\}}\right)\tau + o(\tau) \\
&= |\mathsf{S}(\mathbf{h}_{i,t})| + \sum_{k=1}^{N} \left(\left(\gamma + \alpha|\mathsf{S}(\mathbf{h}_{i,t})| + \beta\sum_{j=1}^{n} a_{ij,t}h_{jk,t}\right)\mathbb{1}_{\{h_{ik,t}=0\}} - \lambda\mathbb{1}_{\{h_{ik,t}=1\}}\right)\tau + o(\tau) \\
&= |\mathsf{S}(\mathbf{h}_{it})| + (\gamma\tau + \alpha\tau|\mathsf{S}(\mathbf{h}_{it})|)|\mathsf{S}(\mathbf{h}_{it}^c)| - \lambda\tau|\mathsf{S}(\mathbf{h}_{it})| + \beta\tau\sum_{j=1}^{n} a_{ij,t}\langle \mathbf{h}_{it}^c, \mathbf{h}_{jt}\rangle + o(\tau).
\end{aligned} \tag{69}
$$

The product $\langle \mathbf{h}_{it}^c, \mathbf{h}_{jt}\rangle$ measures 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 the above equation as

$$\mathbb{E}_t\left(|\mathsf{S}(\mathbf{h}_{i,t+\tau})||\mathbf{h}_t, G_t\right) = f_{it} + \Delta f_{it}\tau + o(\tau).$$

Considering the network $G_t + ij$ obtained from G_t by adding the link ij we then can write

$$\mathbb{E}_t\left(|\mathsf{S}(\mathbf{h}_{i,t+\tau})||\mathbf{h}_t, G_t + ij\right) = f_{it} + \Delta f_{it}\tau + \beta f_{ij,t}\tau + o(\tau),$$

where we have also denoted by $f_{ij,t} \equiv \langle \mathbf{h}_{it}^c, \mathbf{h}_{jt}\rangle$.

From Equation (8) 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 next period's profit of firm i as

$$
\begin{aligned}
\mathbb{E}_t\left(\pi_i(\mathbf{h}_{t+\tau})|\mathbf{h}_t, G_t\right) &= 1 + \mathbb{E}_t\left(|\mathsf{S}(\mathbf{h}_{i,t+\tau t})||\mathbf{h}_t, G_t\right) - cd_{it} \\
&= 1 + f_{it} + \Delta f_{it}\tau + \beta f_{ij,t}\tau - cd_{it} + o(\tau).
\end{aligned}
$$

Then the change in the present discounted profit of a firm i from forming the link ij is given by

$$
\begin{aligned}
V_i(\mathbf{h}_t, G_t + ij) - V_i(\mathbf{h}_t, G_t) &= \delta\,\mathbb{E}_t\left(\pi_i(\mathbf{h}_{t+\tau})|\mathbf{h}_t, G_t + ij\right) - \delta\,\mathbb{E}_t\left(\pi_i(\mathbf{h}_{t+\tau})|\mathbf{h}_t, G_t\right) + o(\tau) \\
&= \delta\,\mathbb{E}_t\left(|\mathsf{S}(\mathbf{h}_{i,t+\tau})||\mathbf{h}_t, G_t + ij\right) - \delta\,\mathbb{E}_t\left(|\mathsf{S}(\mathbf{h}_{i,t+\tau})||\mathbf{h}_t, G_t\right) - \delta c + o(\tau) \\
&= \delta\left(\beta\tau f_{ij,t} - c\right) + o(\tau) \\
&= \delta\left(\beta\tau\langle \mathbf{h}_{it}^c, \mathbf{h}_{jt}\rangle - c\right) + o(\tau).
\end{aligned}
$$

To simplify our notation we have assumed in the expression above that the per period cost for an additional link needs to be paid at the end of a collaboration period.[41]

[41] Otherwise, we could introduce a cost $c' = 1 + c$ to obtain the same expression.

Next we consider the case of competitive markets when $\theta = 1$. Note that

$$\mathbb{E}_t\left(|\,\mathsf{S}(\mathbf{h}_{i,t+\tau})|^2|\mathbf{h}_t, G_t\right) = |\,\mathsf{S}(\mathbf{h}_{it})|^2 + 2|\,\mathsf{S}(\mathbf{h}_{it})|\left((\gamma+\alpha|\,\mathsf{S}(\mathbf{h}_{it})|)|\,\mathsf{S}(\mathbf{h}_{it}^c)| - \lambda|\,\mathsf{S}(\mathbf{h}_{it})| + \beta\sum_{j=1}^n a_{ij,t}\langle\mathbf{h}_{it}^c, \mathbf{h}_{jt}\rangle\right)\tau + o(\tau), \quad (70)$$

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

$$\begin{aligned}\bar{h}_{t+\tau}(\mathbf{h}_t, G_t) &\equiv \mathbb{E}_t\left(\frac{1}{n}\sum_{i=1}^n|\,\mathsf{S}(\mathbf{h}_{i,t+\tau})|\,\middle|\,\mathbf{h}_t, G_t\right)\\ &= \gamma N\tau + \frac{1-\lambda\tau-\gamma\tau}{n}\sum_{i=1}^n|\,\mathsf{S}(\mathbf{h}_{it})| + \frac{\alpha\tau}{n}\sum_{i=1}^n|\,\mathsf{S}(\mathbf{h}_{it})||\,\mathsf{S}(\mathbf{h}_{it}^c)| + \frac{\beta\tau}{n}\sum_{i=1}^n\sum_{j=1}^n a_{ij,t}\langle\mathbf{h}_{it}^c, \mathbf{h}_{jt}\rangle + o(\tau)\\ &= \bar{h}_t + \Delta\bar{h}_t\tau + o(\tau), \end{aligned} \quad (71)$$

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

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

Using a Taylor expansion around the mean (see e.g. Chap. 2.3 in Paolella [2007])

$$\mathbb{E}\left(\frac{X}{Y}\right) = \frac{\mathbb{E}(X)}{\mathbb{E}(Y)} + O\left(\frac{1}{\mathbb{E}(Y)^2}\right),$$

we then can write[42]

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

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

and

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

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

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

We then have that

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

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

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

$$= \frac{f_{it}^2+2f_{it}\Delta f_{it}\tau+2\beta f_{it}f_{ij,t}\tau+o(\tau)}{1+b(\bar{h}_t+\Delta\bar{h}_t\tau)}\frac{1}{1+\frac{1}{n}\frac{b\beta(f_{ij,t}+f_{ji,t})\tau}{1+b(\bar{h}_t+\Delta\bar{h}_t\tau)}+o\left(\frac{\tau}{n}\right)}$$

$$= \frac{f_{it}^2+2f_{it}\Delta f_{it}\tau+2\beta f_{it}f_{ij,t}\tau+o(\tau)}{1+b(\bar{h}_t+\Delta\bar{h}_t\tau)} \times\left(1-\frac{1}{n}\frac{b\beta(f_{ij,t}+f_{ji,t})\tau}{1+b(\bar{h}_t+\Delta\bar{h}_t\tau)}+o\left(\frac{\tau}{n}\right)\right)$$

$$= \frac{f_{it}^2+2f_{it}\Delta f_{it}\tau+2\beta f_{it}f_{ij,t}\tau+o(\tau)}{1+b(\bar{h}_t+\Delta\bar{h}_t\tau)}+O\left(\frac{\tau}{n}\right).$$

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

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

From the above calculations it follows that for $\theta=0$ the change in the present discounted profit can be written as

$$V_i(\mathbf{h}_t, G_t+ij)-V_i(\mathbf{h}_t, G_t) = \delta\frac{2b\beta(1+bf_{it})f_{ij,t}\tau}{1+b\bar{h}_t}-\delta c+O\left(\frac{\tau}{n}\right).$$

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

$$V_i(\mathbf{h}_t, G_t+ij)-V_i(\mathbf{h}_t, G_t) = \beta\tau\delta\left(\theta+(1-\theta)\frac{2b(1+bf_{it})}{1+b\bar{h}_t}\right)f_{ij,t}-\delta c+O\left(\frac{\tau}{n}\right). \tag{72}$$

□

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 $\mathbf{h}$. Moreover, let $\tau_t(\mathbf{h},\mathbf{h}',\mathbf{h}'')$ be 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 $\mathbf{h}''$. Further, let $x_t(\mathbf{h})=\frac{n_t(\mathbf{h})}{n}$ and $z_t(\mathbf{h},\mathbf{h}')=\frac{m_t(\mathbf{h},\mathbf{h}')}{n^2}$. Normalization requires that $\sum_{\mathbf{h}\in\mathcal{H}^N}x_t(\mathbf{h})=1$ and $\sum_{\mathbf{h},\mathbf{h}'\in\mathcal{H}^N}z_t(\mathbf{h},\mathbf{h}')=\frac{2m_t}{n^2}=\frac{\bar{d}_t}{n}$, where m_t is the expected number of links and $\bar{d}_t$ is the expected average degree. Moreover, the expected number of ideas of a 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 $\mathbf{h}'$ wand $h'_k = 1$ 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 $h'_k = 1$. Besides, $n_t(\mathbf{h})$ increases if an idea k becomes obsolete for a firm with technology $\mathbf{h}+\mathbf{e}_k$ at a rate λ, and it declines if an idea k with $h_k = 1$ becomes obsolete at the same rate λ. The expected change in the number $n_t(\mathbf{h})$ of firms with technology $\mathbf{h}$ is then given by

$$\begin{aligned} n\mathsf{F}_x(\mathbf{h}) \equiv & \sum_{k=1}^N \mathbb{1}_{\{h_k=1\}}\left((\gamma+\alpha\langle\mathbf{h}-\mathbf{e}_k,\mathbf{u}\rangle)n_t(\mathbf{h}-\mathbf{e}_k) + \beta\sum_{\mathbf{h}'\in\mathcal{H}^N:h'_k=1} m_t(\mathbf{h}-\mathbf{e}_k,\mathbf{h}')\right) \\ & - \sum_{k=1}^N \mathbb{1}_{\{h_k=0\}}\left((\gamma+\alpha\langle\mathbf{h},\mathbf{u}\rangle)n_t(\mathbf{h}) + \beta\sum_{\mathbf{h}'\in\mathcal{H}^N:h'_k=1} m_t(\mathbf{h},\mathbf{h}')\right) \\ & + \lambda\sum_{k=1}^N \mathbb{1}_{\{h_k=0\}} n_t(\mathbf{h}+\mathbf{e}_k) - \lambda\sum_{k=1}^N \mathbb{1}_{\{h_k=1\}} n_t(\mathbf{h}). \end{aligned} \tag{73}$$

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

$$\begin{aligned} \mathsf{F}_x(\mathbf{h}) = & \sum_{k=1}^N \mathbb{1}_{\{h_k=1\}}\left(\gamma+\alpha\langle\mathbf{h}-\mathbf{e}_k,\mathbf{u}\rangle + n\beta\sum_{\mathbf{h}'\in\mathcal{H}^N:h'_k=1} \frac{z_t(\mathbf{h}-\mathbf{e}_k,\mathbf{h}')}{x_t(\mathbf{h}-\mathbf{e}_k)}\mathbb{1}_{\{x_t(\mathbf{h}-\mathbf{e}_k)>0\}}\right) x_t(\mathbf{h}-\mathbf{e}_k) \\ & - \sum_{k=1}^N \mathbb{1}_{\{h_k=0\}}\left(\gamma+\alpha\langle\mathbf{h},\mathbf{u}\rangle + n\beta\sum_{\mathbf{h}'\in\mathcal{H}^N:h'_k=1} \frac{z_t(\mathbf{h},\mathbf{h}')}{x_t(\mathbf{h})}\mathbb{1}_{\{x_t(\mathbf{h})>0\}}\right) x_t(\mathbf{h}) \\ & + \lambda\sum_{k=1}^N \mathbb{1}_{\{h_k=0\}} x_t(\mathbf{h}+\mathbf{e}_k) - \lambda\sum_{k=1}^N \mathbb{1}_{\{h_k=1\}} x_t(\mathbf{h}). \end{aligned} \tag{74}$$

Observe that by introducing the rescaled variables $\beta\to\beta/n$, $c\to c/n$, $\delta/n\to\delta$, marginal profits in Equation (72) do not change, and similarly, rescaling the parameters $\beta\to\beta/n$, $c\to c/n$ and assuming that the parameter of the logistically distributed error term for the marginal profits from collaborations in Definition 1 is rescaled as $\eta/\delta\to\eta n/\delta$ leaves the linking probability in Equation (14) unchanged. However, Equation (74) becomes independent of n and can be written as

$$\begin{aligned} \mathsf{F}_x(\mathbf{h}) = & \sum_{k=1}^N \mathbb{1}_{\{h_k=1\}}\left(\gamma+\alpha\langle\mathbf{h}-\mathbf{e}_k,\mathbf{u}\rangle + \beta\sum_{\mathbf{h}'\in\mathcal{H}^N:h'_k=1} \frac{z_t(\mathbf{h}-\mathbf{e}_k,\mathbf{h}')}{x_t(\mathbf{h}-\mathbf{e}_k)}\mathbb{1}_{\{x_t(\mathbf{h}-\mathbf{e}_k)>0\}}\right) x_t(\mathbf{h}-\mathbf{e}_k) \\ & - \sum_{k=1}^N \mathbb{1}_{\{h_k=0\}}\left(\gamma+\alpha\langle\mathbf{h},\mathbf{u}\rangle + \beta\sum_{\mathbf{h}'\in\mathcal{H}^N:h'_k=1} \frac{z_t(\mathbf{h},\mathbf{h}')}{x_t(\mathbf{h})}\mathbb{1}_{\{x_t(\mathbf{h})>0\}}\right) x_t(\mathbf{h}) \\ & + \lambda\sum_{k=1}^N \mathbb{1}_{\{h_k=0\}} x_t(\mathbf{h}+\mathbf{e}_k) - \lambda\sum_{k=1}^N \mathbb{1}_{\{h_k=1\}} x_t(\mathbf{h}). \end{aligned} \tag{75}$$

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

$$
\begin{aligned}
\mathbb{P}(a_{ij,t} = 1 | \mathbf{h}_{it} = \mathbf{h}, \mathbf{h}_{jt} = \mathbf{h}') &= \frac{\mathbb{P}(a_{ij,t} = 1, \mathbf{h}_{it} = \mathbf{h}, \mathbf{h}_{jt} = \mathbf{h}')}{\mathbb{P}(\mathbf{h}_{it} = \mathbf{h})\mathbb{P}(\mathbf{h}_{jt} = \mathbf{h}')} \\
&= \frac{m_t(\mathbf{h}, \mathbf{h}')}{n^2} \frac{1}{\frac{n_t(\mathbf{h})}{n}} \frac{1}{\frac{n_t(\mathbf{h}')}{n}} \\
&= \frac{m_t(\mathbf{h}, \mathbf{h}')}{n_t(\mathbf{h}) n_t(\mathbf{h}')} \\
&= \frac{z_t(\mathbf{h}, \mathbf{h}')}{x_t(\mathbf{h}) x_t(\mathbf{h}')} \\
&\equiv \xi_t(\mathbf{h}, \mathbf{h}'),
\end{aligned}
$$

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

$$
\begin{aligned}
\mathbb{P}(a_{ij,t} = 1, \mathbf{h}_{jt} = \mathbf{h}' | \mathbf{h}_{it} = \mathbf{h}) &= \mathbb{P}(a_{ij,t} = 1 | \mathbf{h}_{it} = \mathbf{h}, \mathbf{h}_{jt} = \mathbf{h}') \mathbb{P}(\mathbf{h}_{jt} = \mathbf{h}') \\
&= \frac{\mathbb{P}(a_{ij,t} = 1, \mathbf{h}_{jt} = \mathbf{h}', \mathbf{h}_{it} = \mathbf{h})}{\mathbb{P}(\mathbf{h}_{it} = \mathbf{h})\mathbb{P}(\mathbf{h}_{jt} = \mathbf{h}')} \mathbb{P}(\mathbf{h}_{jt} = \mathbf{h}') \\
&= \frac{\mathbb{P}(a_{ij,t} = 1, \mathbf{h}_{jt} = \mathbf{h}', \mathbf{h}_{it} = \mathbf{h})}{\mathbb{P}(\mathbf{h}_{it} = \mathbf{h})} \\
&= \frac{z_t(\mathbf{h}, \mathbf{h}')}{x_t(\mathbf{h})}.
\end{aligned}
$$

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 (75) as follows

$$
\begin{aligned}
\mathsf{F}_x(\mathbf{h}) = \gamma &\sum_{k \in \mathsf{S}(\mathbf{h})} x_t(\mathbf{h} - \mathbf{e}_k) + \lambda \sum_{k \in \mathsf{S}(\mathbf{h}^c)} x_t(\mathbf{h} + \mathbf{e}_k) \\
&- (\lambda |\mathsf{S}(\mathbf{h})| + \gamma |\mathsf{S}(\mathbf{h}^c)|) \, x_t(\mathbf{h}) \\
&+ \alpha (|\mathsf{S}(\mathbf{h})| - 1) \sum_{k \in \mathsf{S}(\mathbf{h})} x_t(\mathbf{h} - \mathbf{e}_k) \\
&- \alpha |\mathsf{S}(\mathbf{h})| |\mathsf{S}(\mathbf{h}^c)| x_t(\mathbf{h}) \\
&+ \beta \sum_{k \in \mathsf{S}(\mathbf{h})} \sum_{\mathbf{h}' \in \mathcal{H}^N : h'_k = 1} z_t(\mathbf{h} - \mathbf{e}_k, \mathbf{h}') \mathbb{1}_{\{x_t(\mathbf{h} - \mathbf{e}_k) > 0\}} \\
&- \beta \sum_{k \in \mathsf{S}(\mathbf{h}^c)} \sum_{\mathbf{h}' \in \mathcal{H}^N : h'_k = 1} z_t(\mathbf{h}, \mathbf{h}') \mathbb{1}_{\{x_t(\mathbf{h}) > 0\}}.
\end{aligned} \tag{76}
$$

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 $\rho n_t(\mathbf{h}) n_t(\mathbf{h}') g(\mathbf{h}, \mathbf{h}') \left(1 - \frac{m_t(\mathbf{h}, \mathbf{h}')}{n_t(\mathbf{h}) n_t(\mathbf{h}')}\right)$, 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 $g(\mathbf{h}, \mathbf{h}')$ is the probability that they want to form a link when they have the opportunity. Collecting these terms, and noting that contributions stemming from changes in the technologies $\mathbf{h}$ and $\mathbf{h}'$ of the firms happen at a rate $o(\rho)$ we get

$$
n^2 \mathsf{F}_z(\mathbf{h}, \mathbf{h}') \equiv \rho \, n_t(\mathbf{h}) n_t(\mathbf{h}') g(\mathbf{h}, \mathbf{h}') \left(1 - \frac{m_t(\mathbf{h}, \mathbf{h}')}{n_t(\mathbf{h}) n_t(\mathbf{h}')}\right) - \rho \, m_t(\mathbf{h}, \mathbf{h}') + o(\rho), \tag{77}
$$

where we have denoted by

$$g(\mathbf{h},\mathbf{h}') \equiv \frac{e^{\eta(\beta g_{\theta,\tau}(x_t)(1+b|\mathsf{S}(\mathbf{h})|)^{1-\theta}\langle \mathbf{h}^c,\mathbf{h}'\rangle - c)}}{1+e^{\eta(\beta g_{\theta,\tau}(x_t)(1+b|\mathsf{S}(\mathbf{h})|)^{1-\theta}\langle \mathbf{h}^c,\mathbf{h}'\rangle - c)}} \frac{e^{\eta(\beta g_{\theta,\tau}(x_t)(1+|\mathsf{S}(\mathbf{h}')|)^{1-\theta}\langle \mathbf{h}'^c,\mathbf{h}\rangle - c)}}{1+e^{\eta(\beta g_{\theta,\tau}(x_t)(1+|\mathsf{S}(\mathbf{h}')|)^{1-\theta}\langle \mathbf{h}'^c,\mathbf{h}\rangle - c)}}. \tag{78}$$

Dividing Equation (77) by n^2 gives

$$\mathsf{F}_z(\mathbf{h},\mathbf{h}') = \rho\, x_t(\mathbf{h})x_t(\mathbf{h}')\left[g(\mathbf{h},\mathbf{h}')\left(1-\frac{z_t(\mathbf{h},\mathbf{h}')}{x_t(\mathbf{h})x_t(\mathbf{h}')}\right)-\frac{z_t(\mathbf{h},\mathbf{h}')}{x_t(\mathbf{h})x_t(\mathbf{h}')}\right]+o(\rho). \tag{79}$$

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

$$\mathbb{P}\left(\zeta_P^n = z\right) = \mathbb{P}\left(P^n\left(t+\Delta t\right) = P + z \middle| P^n\left(t\right) = P\right).$$

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 (76) and $\mathsf{F}_z(\mathbf{h},\mathbf{h}')$ as in Equation (79). Further, we introduce the functions V_y^N, A_y^N and $A_{y,\delta}^N$ defined by

$$\begin{aligned} V_y^n(P) &\equiv \lambda_y^n \mathbb{E}[\zeta_y^n], \\ A_y^n(P) &\equiv \lambda_y^n \mathbb{E}[|\zeta_y^n|], \\ A_{y,\delta}^n(P) &\equiv \lambda_y^n \mathbb{E}[\left|\zeta_y^n I_{\{|\zeta_y^n|>\delta\}}\right|], \end{aligned}$$

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

$$\begin{aligned} V_x^n(P) &= n\mathbb{E}[\zeta_x^n] \\ &= n\sum_{\mathbf{h},\mathbf{h}'}\frac{1}{n}(e_{\mathbf{h}'}-e_{\mathbf{h}})\mathbb{P}\left(\zeta_x^n = \frac{1}{n}(e_{\mathbf{h}'}-e_{\mathbf{h}})\right) \\ &= \sum_{\mathbf{h}} e_{\mathbf{h}}\mathsf{F}_x(\mathbf{h}) = V_x(P), \end{aligned}$$

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

$$\begin{aligned}
V_z^n(P) &= n^2\mathbb{E}[\zeta_z^n] \\
&= n^2 \sum_{\mathbf{h}'',\mathbf{h}'''} \frac{1}{n^2}(e_{\mathbf{h}'',\mathbf{h}'''} - e_{\mathbf{h},\mathbf{h}'})\mathbb{P}\left(\zeta_z^n = \frac{1}{n^2}(e_{\mathbf{h}'',\mathbf{h}'''} - e_{\mathbf{h},\mathbf{h}'})\right) \\
&= \sum_{\mathbf{h},\mathbf{h}'} e_{\mathbf{h},\mathbf{h}'}\mathsf{F}_z(\mathbf{h},\mathbf{h}) = V_z(P),
\end{aligned}$$

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[43]

$$\frac{dP}{dt} = \frac{d}{dt}\begin{bmatrix} \vdots \\ x_t(\mathbf{h}) \\ \vdots \\ z_t(\mathbf{h},\mathbf{h}') \\ \vdots \end{bmatrix} = V(P) = \begin{bmatrix} V_x(P) \\ V_z(P) \end{bmatrix} = \begin{bmatrix} \vdots \\ \mathsf{F}_x(\mathbf{h}) \\ \vdots \\ \mathsf{F}_z(\mathbf{h},\mathbf{h}') \\ \vdots \end{bmatrix} \tag{80}$$

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

$$\xi_t(\mathbf{h},\mathbf{h}') \equiv \mathbb{P}(a_{ij,t} = 1|\mathbf{h}_{it} = \mathbf{h}, \mathbf{h}_{jt} = \mathbf{h}') = \frac{z_t(\mathbf{h},\mathbf{h}')}{x_t(\mathbf{h})x_t(\mathbf{h}')}, \tag{81}$$

then we have that

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

[43] See also Kurtz [1971] and Wormald [1995].

Inserting Equation (79) gives

$$\begin{aligned}\frac{d\xi_t(\mathbf{h},\mathbf{h}')}{dt} &= \rho g(\mathbf{h},\mathbf{h}')\,(1-\xi_t(\mathbf{h},\mathbf{h}')) - \rho\,\xi_t(\mathbf{h},\mathbf{h}') \\ &+ \xi_t(\mathbf{h},\mathbf{h}')\left(\frac{1}{x_t(\mathbf{h})}\frac{dx_t(\mathbf{h})}{dt} + \frac{1}{x_t(\mathbf{h}')}\frac{dx_t(\mathbf{h}')}{dt}\right) \\ &- \xi_t(\mathbf{h},\mathbf{h}')\left(\frac{1}{x_t(\mathbf{h})}\frac{dx_t(\mathbf{h})}{dt} + \frac{1}{x_t(\mathbf{h}')}\frac{dx_t(\mathbf{h}')}{dt}\right) + o(\rho) \\ &= \rho\, g(\mathbf{h},\mathbf{h}') - \rho\,(1+g(\mathbf{h},\mathbf{h}'))\,\xi_t(\mathbf{h},\mathbf{h}') + o(\rho). \end{aligned} \tag{82}$$

Moreover, inserting the definition of $\xi_t(\mathbf{h},\mathbf{h}')$ from Equation (81) into Equation (76) we get

$$\begin{aligned}\frac{dx_t(\mathbf{h})}{dt} &= \gamma \sum_{k\in \mathsf{S}(\mathbf{h})} x_t(\mathbf{h}-\mathbf{e}_k) + \lambda \sum_{k\in \mathsf{S}(\mathbf{h}^c)} x_t(\mathbf{h}+\mathbf{e}_k) \\ &- (\lambda|\mathsf{S}(\mathbf{h})| + \gamma|\mathsf{S}(\mathbf{h}^c)|)\,x_t(\mathbf{h}) \\ &+ \alpha(|\mathsf{S}(\mathbf{h})|-1)\sum_{k\in \mathsf{S}(\mathbf{h})} x_t(\mathbf{h}-\mathbf{e}_k) \\ &- \alpha|\mathsf{S}(\mathbf{h})||\mathsf{S}(\mathbf{h}^c)|x_t(\mathbf{h}) \\ &+ \beta \sum_{k\in \mathsf{S}(\mathbf{h})}\sum_{\mathbf{h}'\in\mathcal{H}^N:h'_k=1} \xi_t(\mathbf{h}-\mathbf{e}_k,\mathbf{h}')x_t(\mathbf{h}-\mathbf{e}_k)x_t(\mathbf{h}') \\ &- \beta \sum_{k\in \mathsf{S}(\mathbf{h}^c)}\sum_{\mathbf{h}'\in\mathcal{H}^N:h'_k=1} \xi_t(\mathbf{h},\mathbf{h}')x_t(\mathbf{h})x_t(\mathbf{h}'). \end{aligned} \tag{83}$$

□

Proof of Proposition 2. Let us define

$$\tilde{x}_t(s) \equiv \sum_{\mathbf{h}\in\mathcal{H}^N:|\mathsf{S}(\mathbf{h})|=s} x_t(\mathbf{h}), \quad 0\le s\le N.$$

Summation over all $\mathbf{h}\in\mathcal{H}^N$ with the property that $|\mathsf{S}(\mathbf{h})|=s$ in Equation (76) gives

$$\begin{aligned}\frac{d\tilde{x}_t(s)}{dt} &= \gamma(N-s+1)\tilde{x}_t(s-1) + \lambda(s+1)\tilde{x}_t(s+1) \\ &- (\lambda s + \gamma(N-s))\tilde{x}_t(s) \\ &+ \alpha(s-1)(N-s+1)\tilde{x}_t(s-1) \\ &- \alpha s(N-s)\tilde{x}_t(s) \\ &+ \beta \sum_{\mathbf{h}\in\mathcal{H}^N:|\mathsf{S}(\mathbf{h})|=s}\sum_{k\in \mathsf{S}(\mathbf{h})}\sum_{\mathbf{h}'\in\mathcal{H}^N:h'_k=1} \xi_t(\mathbf{h}-\mathbf{e}_k,\mathbf{h}')x_t(\mathbf{h}-\mathbf{e}_k)x_t(\mathbf{h}') \\ &- \beta \sum_{\mathbf{h}\in\mathcal{H}^N:|\mathsf{S}(\mathbf{h})|=s}\sum_{k\in \mathsf{S}(\mathbf{h}^c)}\sum_{\mathbf{h}'\in\mathcal{H}^N:h'_k=1} \xi_t(\mathbf{h},\mathbf{h}')x_t(\mathbf{h})x_t(\mathbf{h}'), \end{aligned} \tag{84}$$

where we have used the fact that

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

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

$$\tilde{z}_t(s,s') \equiv \sum_{\substack{\mathbf{h}\in\mathcal{H}^N:|\mathsf{S}(\mathbf{h})|=s\\ \mathbf{h}'\in\mathcal{H}^N:|\mathsf{S}(\mathbf{h})|=s'}} z_t(\mathbf{h},\mathbf{h}'), \quad 0\le s,s'\le N,$$

and let the average be given by

$$\bar{z}_t(s,s') \equiv \frac{1}{\binom{N}{s}}\frac{1}{\binom{N}{s'}}\sum_{\substack{\mathbf{h}\in\mathcal{H}^N:|\mathsf{S}(\mathbf{h})|=s\\ \mathbf{h}'\in\mathcal{H}^N:|\mathsf{S}(\mathbf{h})|=s'}} z_t(\mathbf{h},\mathbf{h}') = \frac{1}{\binom{N}{s}}\frac{1}{\binom{N}{s'}}\tilde{\xi}_t(s,s'), \quad 0\le s,s'\le N. \tag{85}$$

With[44]

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

and similarly

$$\sum_{\mathbf{h}\in\mathcal{H}^N:|\mathsf{S}(\mathbf{h})|=s}\sum_{k\in\mathsf{S}(\mathbf{h}^c)}\sum_{\mathbf{h}'\in\mathcal{H}^N:h'_k=1} z_t(\mathbf{h},\mathbf{h}') = (N-s+1)\sum_{s'=1}^{N}\frac{s'}{N-s'+1}\tilde{z}_t(s,s').$$

[44]If the initial conditions are such that $x_0(\mathbf{h})$, $z_0(\mathbf{h},\mathbf{h}')$ and $\xi_0(\mathbf{h},\mathbf{h}')$ 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 obtain

$$\begin{aligned}\frac{d\tilde{x}_t(s)}{dt} &= \gamma(N-s+1)\tilde{x}_t(s-1) + \lambda(s+1)\tilde{x}_t(s+1) \\ &- (\lambda s + \gamma(N-s))\tilde{x}_t(s) \\ &+ \alpha(s-1)(N-s+1)\tilde{x}_t(s-1) \\ &- \alpha s(N-s)\tilde{x}_t(s) \\ &+ \beta(N-s+1)\sum_{s'=1}^{N}\frac{s'}{N-s'+1}\tilde{z}_t(s-1,s')\mathbb{1}_{\{\tilde{x}_t(s-1)>0\}} \\ &- \beta(N-s+1)\sum_{s'=1}^{N}\frac{s'}{N-s'+1}\tilde{z}_t(s,s')\mathbb{1}_{\{\tilde{x}_t(s)>0\}},\end{aligned} \tag{86}$$

We can write Equation (86) more compactly as follows

$$\begin{aligned}\frac{d\tilde{x}_t(s)}{dt} &= \gamma(N-s+1)(\gamma+\alpha(s-1))\tilde{x}_t(s-1) + \lambda(s+1)\tilde{x}_t(s+1) \\ &- (\lambda s + \gamma(N-s) + \alpha s(N-s))\tilde{x}_t(s) \\ &+ \beta(N-s+1)\sum_{s'=1}^{N}\frac{s'}{N-s'+1}(\tilde{z}_t(s-1,s')\mathbb{1}_{\{\tilde{x}_t(s-1)>0\}} - \tilde{z}_t(s,s')\mathbb{1}_{\{\tilde{x}_t(s)>0\}}).\end{aligned} \tag{87}$$

Introducing the average

$$\bar{x}_t(s) \equiv \frac{1}{\binom{N}{s}}\sum_{\mathbf{h}\in\mathcal{H}^N:|\mathsf{S}(\mathbf{h})|=s} x_t(\mathbf{h}) = \frac{1}{\binom{N}{s}}\tilde{x}_t(s), \quad 0 \le s \le N, \tag{88}$$

and using the fact that

$$\frac{d\bar{x}_t(s)}{dt} = \frac{1}{\binom{N}{s}}\sum_{\mathbf{h}\in\mathcal{H}^N:|\mathsf{S}(\mathbf{h})|=s}\frac{dx_t(\mathbf{h})}{dt},$$

where $\binom{N}{s}$ is the number of vectors $\mathbf{h} \in \mathcal{H}^N$ with $|\mathsf{S}(\mathbf{h})| = s$, we obtain from Equation (76)

$$\begin{aligned}\frac{d\bar{x}_t(s)}{dt} &= \gamma s\bar{x}_t(s-1) + \lambda(N-s)\bar{x}_t(s+1) \\ &- (\lambda s + \gamma(N-s))\bar{x}_t(s) \\ &+ \alpha(s-1)s\bar{x}_t(s-1) - \alpha s(N-s)\bar{x}_t(s) \\ &+ \beta s\sum_{s'=1}^{N}\bar{z}_t(s-1,s')\binom{N-1}{s'-1}\mathbb{1}_{\{\bar{x}_t(s-1)>0\}} \\ &- \beta(N-s)\sum_{s'=1}^{N}\bar{z}_t(s,s')\binom{N-1}{s'-1}\mathbb{1}_{\{\bar{x}_t(s)>0\}},\end{aligned}$$

which can be written more compactly as follows

$$\begin{aligned}\frac{d\bar{x}_t(s)}{dt} &= (\gamma s + \alpha(s-1)s)\bar{x}_t(s-1) + \lambda(N-s)\bar{x}_t(s+1) \\ &- (\lambda s + \gamma(N-s) + \alpha s(N-s))\bar{x}_t(s) \\ &+ \beta\sum_{s'=1}^{N}\binom{N-1}{s'-1}(s\bar{z}_\tau(s-1,s')\mathbb{1}_{\{\bar{x}_t(s-1)>0\}} - (N-s)\bar{z}_\tau(s,s')\mathbb{1}_{\{\bar{x}_t(s)>0\}}).\end{aligned} \tag{89}$$

We further define

$$\bar{\xi}_t(s,s') \equiv \frac{1}{\binom{N}{s}} \frac{1}{\binom{N}{s'}} \sum_{\substack{\mathbf{h}\in\mathcal{H}^N:|\mathsf{S}(\mathbf{h})|=s \\ \mathbf{h}'\in\mathcal{H}^N:|\mathsf{S}(\mathbf{h})|=s'}} \frac{z_t(\mathbf{h},\mathbf{h}')}{x_t(\mathbf{h})x_t(\mathbf{h}')} \approx \frac{\bar{z}_t(s,s')}{\bar{x}_t(s)\bar{x}_t(s')}, \tag{90}$$

so that we can write Equation (89) as follows

$$\begin{aligned}\frac{d\bar{x}_t(s)}{dt} &= (\gamma s + \alpha(s-1)s)\bar{x}_t(s-1) + \lambda(N-s)\bar{x}_t(s+1) \\ &- (\lambda s + \gamma(N-s) + \alpha s(N-s))\bar{x}_t(s) \\ &+ \beta \sum_{s'=1}^{N} \binom{N-1}{s'-1} \left(s\bar{\xi}_t(s-1,s')\bar{x}_t(s-1)\bar{x}_t(s') - (N-s)\bar{\xi}_t(s,s')\bar{x}_t(s)\bar{x}_t(s') \right).\end{aligned}$$

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 (90) into Equation (18) gives

$$\frac{d\bar{\xi}_t(s,s')}{dt} = \rho\bar{g}(s,s') - \rho(1+\bar{g}(s,s'))\bar{\xi}_t(s,s') + o(\rho),$$

where we have denoted by

$$\begin{aligned}\bar{g}(s,s') &\equiv \frac{1}{\binom{N}{s}} \frac{1}{\binom{N}{s'}} \tilde{g}(s,s') \\ &\equiv \frac{1}{\binom{N}{s}} \frac{1}{\binom{N}{s'}} \sum_{\substack{\mathbf{h}\in\mathcal{H}^N:|\mathsf{S}(\mathbf{h})|=s \\ \mathbf{h}'\in\mathcal{H}^N:|\mathsf{S}(\mathbf{h}')|=s'}} g(\mathbf{h},\mathbf{h}') \\ &= \frac{1}{\binom{N}{s}} \frac{1}{\binom{N}{s'}} \sum_{\substack{\mathbf{h}\in\mathcal{H}^N:|\mathsf{S}(\mathbf{h})|=s \\ \mathbf{h}'\in\mathcal{H}^N:|\mathsf{S}(\mathbf{h}')|=s'}} \frac{e^{\eta(\beta g_{\theta,\tau}(x_t)(1+s)^{1-\theta}\langle \mathbf{h}^c,\mathbf{h}'\rangle - c)}}{1+e^{\eta(\beta g_{\theta,\tau}(x_t)(1+s)^{1-\theta}\langle \mathbf{h}^c,\mathbf{h}'\rangle - c)}} \frac{e^{\eta(\beta g_{\theta,\tau}(\bar{h}_t)(1+s')^{1-\theta}\langle \mathbf{h}'^c,\mathbf{h}\rangle - c)}}{1+e^{\eta(\beta g_{\theta,\tau}(x_t)(1+s')^{1-\theta}\langle \mathbf{h}'^c,\mathbf{h}\rangle - c)}},\end{aligned}$$

which is a symmetric matrix, $\bar{g}(s,s') = \bar{g}(s',s)$ for all $0 \leq s,s' \leq N$. □

Proof of Corollary 1. The proof follows from setting $\frac{d\tilde{x}_t(s)}{dt} = 0$ in Equation (87) and inserting $\tilde{x}(s)$ from Equation (25). □

Proof of Lemma 1. It follows immediately, that the fixed points for $\bar{\xi}(s,s')$ of Equation (24) are given by

$$\bar{\xi}(s,s') = \frac{\bar{g}(s,s')}{1+\bar{g}(s,s')}.$$

Moreover, the fixed points from Equation (23) satisfy the following recursive equation

$$\begin{aligned}&\left(\lambda s + (N-s)(\gamma+\alpha s) + \beta(N-s)\sum_{s'=1}^{N}\binom{N-1}{s'-1}\bar{\xi}(s,s')\bar{x}(s') \right)\bar{x}(s) \\ &\qquad = \left(s(\gamma+\alpha(s-1)) + \beta s \sum_{s'=1}^{N}\binom{N-1}{s'-1}\bar{\xi}(s-1,s')\bar{x}(s') \right)\bar{x}(s-1) + \lambda(N-s)\bar{x}(s+1).\end{aligned}$$

For $s=0$ we obtain

$$\left(\gamma + \beta \sum_{s'=1}^{N}\binom{N-1}{s'-1}\bar{\xi}(0,s')\bar{x}(s') \right)\bar{x}(0) = \lambda\bar{x}(1). \tag{91}$$

Similarly, for $s = 1$ we get

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

Using Equation (91), this can be written as

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

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

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

Hence, we get

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

□

Proof of Corollary 2. The proof follows directly from Lemma 1 for $\beta = 0$ and the normalization condition $1 = \sum_{s=0}^{N}\binom{N}{s}\bar{x}(s)$. □

Proof of Proposition 3. In the limit of $\rho \to \infty$ and $\beta, \gamma \to 0$ we have that

$$\bar{x}(s) = \bar{x}(0)\left(\frac{\beta}{\lambda}\right)^{s}\prod_{k=0}^{s-1}\sum_{s'=1}^{N}\hat{\xi}(k,s')\bar{x}(s'),$$

where we have denoted by $\hat{\xi}(k,s') = \binom{N-1}{s'-1}\bar{\xi}(k,s')$. 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

$$\bar{x}(1) = \bar{x}(0)\frac{\beta}{\lambda}\hat{\xi}(0,1)\bar{x}(1).$$

When $\bar{x}(1) > 0$ (and consequently $\bar{x}(0) < 1$) we must have that

$$\bar{x}(0) = \frac{\lambda}{\beta\hat{\xi}(0,1)}.$$

As $\bar{x}(0) < 1$ it must hold that

$$\beta > \beta^c \equiv \frac{\lambda}{\hat{\xi}(0,1)} = \frac{\lambda(1+\bar{g}(0,1))}{\bar{g}(0,1)}.$$

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

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

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 $Z_{s,s'} \equiv \frac{d\bar{\xi}(s,s')}{dt}$. 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

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

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 $\mu_s = \frac{\partial Z_{s,s}}{\partial \bar{\xi}(s,s)} = -\rho(1 + \bar{g}(s,s))$. 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

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

This can be written as

$$\frac{\partial F_s}{\partial \bar{x}(k)} = F^{(0)}_{sk} + \beta F^{(1)}_{sk},$$

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

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

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

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

This is an upper triangular matrix with eigenvalues given by $0, -\lambda + \beta\bar{\xi}(0,1), -2\lambda, \ldots, -N\lambda$. All eigenvalues are non-positive if $\beta < \frac{\lambda}{\bar{\xi}(0,1)}$, 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

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

and

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

and

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

and

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

This concludes the proof. □

Proof of Proposition 4. We have that

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

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

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

Using the Taylor expansion

$$\prod_{k=0}^{s}(a_k + \beta b_k) = \prod_{k=0}^{s}a_k + \beta\sum_{k=0}^{s}\prod_{l\neq k}^{s}a_l b_k + O(\beta^2),$$

we get

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

We then obtain

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

From which we get

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

Further, denoting by

$$a_s \equiv \tilde{\gamma}^s\prod_{k=0}^{s-1}(1+\tilde{\alpha}k)$$

and

$$b_s \equiv \tilde{\gamma}^s\sum_{k=0}^{s-1}\prod_{l\neq k}^{s-1}(1+\tilde{\alpha}l)\sum_{s'=1}^{N}\hat{\xi}(k,s')\tilde{\gamma}^{s'}\prod_{k'=0}^{s'-1}(1+\tilde{\alpha}k')$$

we can write

$$\bar{x}(s) = \bar{x}(0)a_s + \tilde{\beta}\bar{x}(0)^2 b_s.$$

Moreover, from the normalization condition $1 = \sum_{s=0}^{N}\binom{N}{s}\bar{x}(s)$ we obtain

$$1 = \sum_{s=0}^{N}\binom{N}{s}\bar{x}(s) = \bar{x}(0)\sum_{s=0}^{N}\binom{N}{s}a_s + \tilde{\beta}\bar{x}(0)^2\sum_{s=0}^{N}\binom{N}{s}b_s.$$

Denoting by $a \equiv \sum_{s=0}^{N}\binom{N}{s}a_s$ and $b \equiv \sum_{s=0}^{N}\binom{N}{s}b_s$ we find that

$$\bar{x}(0) = \frac{2}{a+\sqrt{a^2+4\beta b}}.$$

From a first order Taylor expansion we then find

$$\bar{x}(0) = \frac{1}{a} - \tilde{\beta}\frac{b}{a^3} + O(\beta^2),$$

and

$$\tilde{\beta}\bar{x}(0)^2 = \frac{\tilde{\beta}}{a^2} + O(\beta^2),$$

so that

$$\bar{x}(s) = \frac{1}{a}a_s + \tilde{\beta}\frac{1}{a^2}\left(b_s - \frac{b}{a}a_s\right).$$

Further, observe that

$$\frac{a_s}{a} = \frac{\prod_{k=0}^{s-1}(1+\tilde{\alpha}k)}{\sum_{s'=0}^{N}\binom{N}{s'}\prod_{k=0}^{s'-1}(1+\tilde{\alpha}k)} = \frac{\prod_{k=0}^{s-1}\frac{\gamma+\alpha k}{\lambda}}{\sum_{s'=0}^{N}\binom{N}{s'}\prod_{k=0}^{s'-1}\frac{\gamma+\alpha k}{\lambda}} = \bar{x}(s)|_{\beta=0} \equiv \bar{x}_0(s).$$

Hence, we can write

$$\bar{x}(s) = \bar{x}_0(s) + \tilde{\beta}\frac{b_s - b\bar{x}_0(s)}{a^2}.$$

This is

$$\bar{x}(s) = \bar{x}_0(s) + \frac{\beta}{\gamma}\frac{b}{a^2}\left(\frac{b_s}{b} - \bar{x}_0(s)\right) + O\left(\frac{\beta}{\gamma}\right)^2,$$

where $\bar{x}_0(s) \equiv \bar{x}(s)|_{\beta=0}$ is given in Equation (30),

$$b_s = \sum_{k=0}^{s-1} \prod_{l \neq k}^{s-1} \frac{\gamma + \alpha l}{\lambda} \sum_{s'=1}^{N} \binom{N-1}{s'-1} \frac{\bar{g}(k,s')}{1+\bar{g}(k,s')} \prod_{k''=0}^{s'-1} \frac{\gamma + \alpha k''}{\lambda},$$

$b = \sum_{s'=0}^{N} \binom{N}{s'} b_{s'}$ and $a = \sum_{k=0}^{N} \binom{N}{k} \prod_{l=0}^{k-1} \frac{\gamma+\alpha l}{\lambda}$ for all $s = 0, \dots, N$. In the following we denote by $F_s \equiv \frac{d\bar{x}(s)}{dt}$ and $Z_{s,s'} \equiv \frac{d\bar{\xi}(s,s')}{dt}$. 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

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

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 $\mu_s = \frac{\partial Z_{s,s}}{\partial \xi(s,s)} = -\rho(1 + \bar{g}(s,s))$. In order to compute the eigenvalues of the first, observe that

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

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]

$$\mu_i(\mathbf{J}_{11}) = \mu_i(\mathbf{F}^{(0)}) + \beta\frac{\mathbf{w}_i^\top \mathbf{F}^{(1)}\mathbf{v}_i}{\mathbf{w}_i^\top \mathbf{v}_i} + O(\beta^2) \tag{92}$$

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 [cf. Elaydi, 2005]

$$k(\gamma+\alpha(k-1))v_k - (\lambda k + (N-k)(\gamma+\alpha k)+\mu)v_{k+1} + \lambda(N-k)v_{k+2} = 0,$$

with the boundary conditions $v_0 = v_{N+2} = 0$ for $k = 0, \ldots, N-1$. 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 (92) are real up to the first order terms in β, showing that the eigenvalues of the Jacobian are real.

In the following we provide an explicit expression for $N = 1, 2$. 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

$$\mu_1 = 0, \quad \mu_2 = -(\gamma+\lambda).$$

The corresponding left eigenvector is given by

$$\begin{aligned}\mathbf{v}_1 &= \left[\frac{\lambda}{\gamma}, 1\right]^\top \\ \mathbf{v}_2 &= [-1, 1]^\top,\end{aligned}$$

while the right eigenvector is

$$\begin{aligned}\mathbf{w}_1 &= [1, 1]^\top \\ \mathbf{w}_2 &= \left[-\frac{\lambda}{\gamma}, 1\right]^\top.\end{aligned}$$

In the case of $N = 2$ we get the three eigenvalues

$$\mu_1 = 0, \quad \mu_{2,3} = -\frac{1}{2}\left(\alpha + 3(\gamma+\lambda) \pm \sqrt{\alpha^2 - 2\alpha(\gamma-3\lambda)+(\gamma+\lambda)^2}\right).$$

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

$$\begin{aligned}\mathbf{v}_1 &= \left[\frac{\lambda^2}{\gamma(\alpha+\gamma)}, \frac{\lambda}{\alpha+\gamma}, 1\right]^\top \\ \mathbf{v}_2 &= \left[\frac{-\alpha+\gamma-\lambda+\sqrt{\alpha^2-2\alpha(\gamma-3\lambda)+(\gamma+\lambda)^2}}{2(\alpha+\gamma)}, -\frac{\alpha+3\gamma-\lambda+\sqrt{\alpha^2-2\alpha(\gamma-3\lambda)+(\gamma+\lambda)^2}}{4(\alpha+\gamma)}, 1\right]^\top \\ \mathbf{v}_3 &= \left[-\frac{\alpha-\gamma+\lambda+\sqrt{\alpha^2-2\alpha(\gamma-3\lambda)+(\gamma+\lambda)^2}}{2(\alpha+\gamma)}, \frac{-\alpha-3\gamma+\lambda+\sqrt{\alpha^2-2\alpha(\gamma-3\lambda)+(\gamma+\lambda)^2}}{4(\alpha+\gamma)}, 1\right]^\top,\end{aligned}$$

while the right eigenvectors are

$$\mathbf{w}_1 = [1,2,1]^\top$$

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

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

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

Proof of Proposition 5. We first show that the eigenvalues of the Jacobian $\mathbf{J}$ are all negative evaluated at $x = 0$ if the spillover parameter β is below a threshold β^c. The Jacobian is given by

$$\mathbf{J} = \begin{bmatrix} (1-2x)z\beta-\gamma-\lambda & -(x-1)x\beta \\ \frac{2b^2 e^{\eta\left(\frac{2b\beta\tau}{bx+1}-c\right)}(z-1)\beta\eta\rho\tau}{(1+e^{c\eta})\left(1+e^{\eta\left(\frac{2b\beta\tau}{bx+1}-c\right)}\right)^2(bx+1)^2} & -\left(1+\frac{e^{\eta\left(\frac{2b\beta\tau}{bx+1}-c\right)}}{(1+e^{c\eta})\left(1+e^{\eta\left(\frac{2b\beta\tau}{bx+1}-c\right)}\right)}\right)\rho \end{bmatrix}.$$

Evaluated at $x = 0$ this simplifies to

$$\mathbf{J} = \begin{bmatrix} z\beta-\gamma-\lambda & 0 \\ \frac{2b^2 e^{\eta(c+2b\beta\tau)}(z-1)\beta\eta\rho\tau}{(1+e^{c\eta})(e^{c\eta}+e^{2b\beta\eta\tau})^2} & \left(-1-\frac{e^{2b\beta\eta\tau}}{(1+e^{c\eta})(e^{c\eta}+e^{2b\beta\eta\tau})}\right)\rho \end{bmatrix}.$$

The eigenvalues of the Jacobian are given by $\mu_{1,2} = \frac{1}{2}(\mathrm{tr}(\mathbf{J}) \pm \sqrt{\mathrm{tr}(\mathbf{J})^2 - 4\det(\mathbf{J})})$, and we obtain complex eigenvalues if $\mathrm{tr}(\mathbf{J})^2 < 4\det(\mathbf{J})$. For the trace we get

$$\mathrm{tr}(\mathbf{J}) = -\frac{\rho e^{2b\beta\eta\tau}}{(e^{c\eta}+1)(e^{2b\beta\eta\tau}+e^{c\eta})} - \gamma - \lambda + \beta n z - \rho,$$

and for the determinant we obtain

$$\det(\mathbf{J}) = -\frac{\rho\left(2e^{2b\beta\eta\tau}+e^{\eta(2b\beta\tau+c)}+e^{c\eta}+e^{2c\eta}\right)(-\gamma-\lambda+\beta n z)}{(e^{c\eta}+1)(e^{2b\beta\eta\tau}+e^{c\eta})}$$

Inserting into $\mu_{1,2} = \frac{1}{2}(\mathrm{tr}(\mathbf{J}) \pm \sqrt{\mathrm{tr}(\mathbf{J})^2 - 4\det(\mathbf{J})})$ delivers the eigenvalues

$$\mu_1 = -\gamma - \lambda + \beta z,$$
$$\mu_2 = -\frac{\rho\left(2e^{2b\beta\eta\tau}+e^{2b\beta\eta\tau+c\eta}+e^{c\eta}+e^{2c\eta}\right)}{(e^{c\eta}+1)(e^{2b\beta\eta\tau}+e^{c\eta})}.$$

From these eigenvalues we find that they are real and negative if $\beta < \beta^c = \frac{\lambda+\gamma}{z}$. Hence, we have that $x_1 = 0$ is an asymptotically stable fixed point in the limit of $\gamma \to 0$ if

$$\beta < \beta^c = \frac{\lambda(1+g)}{g}, \tag{93}$$

where g is given by

$$g = \begin{cases} \frac{e^{\eta(\beta b\tau-2c)}}{(1+e^{\eta(\beta b\tau-c)})(1+e^{-\eta c})} & \text{if } \theta = 1, \\ \frac{e^{\eta(2\beta\tau b-c)}}{1+e^{\eta(2\beta\tau b-c)}}\frac{e^{-\eta c}}{1+e^{-\eta c}} & \text{if } \theta = 0, \end{cases} \tag{94}$$

from which we get

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

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

Next, we consider the limit $\eta \to 0$. We then have that $g = \frac{1}{4}$ and it follows that in the stationary state $z = \frac{1}{5}$. The steady state for x then solves

$$\gamma - \frac{1}{5}x(5(\gamma + \lambda) + \beta(x-1)) = 0,$$

from which we obtain

$$x = \frac{\beta - 5(\gamma+\lambda) + \sqrt{(\beta - 5(\gamma+\lambda))^2 + 20\beta\gamma}}{2\beta}.$$

From the above equation we find that

$$\frac{\partial x}{\partial \gamma} = \frac{1}{2\beta}\left(\frac{5(\beta + 5(\gamma+\lambda))}{\sqrt{(\beta - 5(\gamma+\lambda))^2 + 20\beta\gamma}} - 5\right).$$

This derivative is non-negative for all $\beta > 0$ and attains its unique maximum at $\beta = 5(\gamma+\lambda)$. Next, for $\eta \to 0$ we obtain the Jacobian

$$\mathbf{J} = \begin{bmatrix} (1-2x)z\beta - \gamma - \lambda & -(x-1)x\beta \\ 0 & -\frac{5\rho}{4} \end{bmatrix}.$$

The trace and determinant are given by

$$\mathrm{tr}(\mathbf{J}) = -\gamma - \lambda + \beta(1-2x)z - \frac{5\rho}{4},$$

and

$$\det(\mathbf{J}) = \frac{5}{4}\rho(\gamma + \lambda + \beta(2x-1)z),$$

from which we get the eigenvalues

$$\mu_1 = -\gamma - \lambda + \beta(1-2x)z,$$
$$\mu_2 = -\frac{5\rho}{4}.$$

Inserting the fixed point delivers

$$\mu_1 = -\frac{1}{5}\sqrt{(\beta - 5(\gamma+\lambda))^2 + 20\beta\gamma},$$

which shows that the Jacobian possesses only real and negative eigenvalues.

We next consider the case of $\theta = 1$. The stationary state is given by

$$z = \frac{g}{1+g}, \tag{95}$$

and

$$\frac{\beta g}{1+g}x^2 - \left(\frac{\beta g}{1+g} - \lambda - \gamma\right)x - \gamma = 0,$$

with the solutions

$$x = \frac{g\beta - \gamma - g\gamma - \lambda - g\lambda + \sqrt{4g(1+g)\beta\gamma + (\gamma + \lambda + g(\gamma + \lambda - \beta))^2}}{2g\beta}. \tag{96}$$

In the limit of $\gamma \to 0$ we obtain the trivial solution $x = 0$ and the solution $x = 1 - \frac{\lambda(1+g)}{\beta g}$. Inserting g gives

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

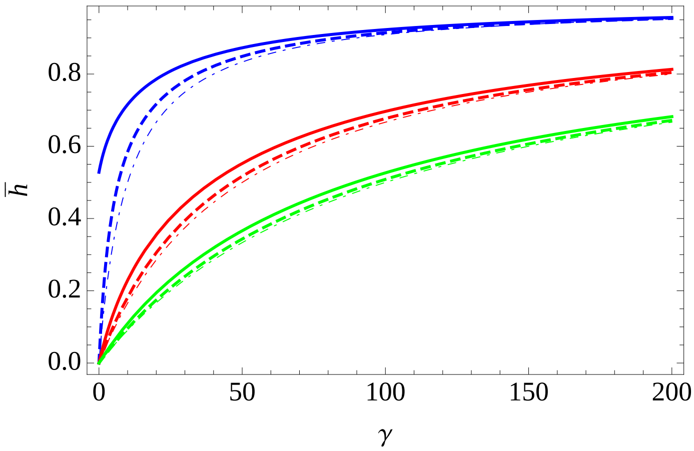


Figure 16: The average stock of knowledge $\bar{h}$ as a function of γ for $\lambda \in \{10, 50, 100\}$. The solid line indicates the exact solution, the dashed line the first order approximation in β and the dashed-dotted line the solution for $\beta = 0$. The parameters used are $b = 10$, $\beta = 100$, $c = 1$, $\eta = 1$ and $\tau = 0.01$.

We then find that

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

Moreover, we find that

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

which is positive if $c\left(e^{b\beta\eta\tau} < 2e^{c\eta} + 1\right) + b\beta\tau\left(e^{c\eta} + 1\right)$. Further, in the limit of small β we obtain

$$x = \frac{\gamma}{\gamma + \lambda} + \beta\frac{\gamma\lambda\bar{g}(0,1)}{(\gamma + \lambda)^3(1 + \bar{g}(0,1))} + O\left(\beta^2\right) = \frac{\gamma}{\gamma + \lambda} + \beta\frac{\gamma\lambda}{\left(2e^{c\eta} + e^{2c\eta} + 2\right)(\gamma + \lambda)^3} + O\left(\beta^2\right).$$

An illustration can be seen in Figure 16. Next we analyze the stability of the fixed point. The Jacobian is

$$\mathbf{J} = \begin{bmatrix} (1 - 2x)z\beta - \gamma - \lambda & (1 - x)x\beta \\ 0 & -\rho(1 + g) \end{bmatrix}.$$

Using the fact that x at $\beta = 0$ is given by $\frac{\gamma}{\gamma + \lambda}$ we can write

$$\Delta x = x - \frac{\gamma}{\gamma + \lambda} = \beta\frac{\gamma\lambda}{\left(2e^{c\eta} + e^{2c\eta} + 2\right)(\gamma + \lambda)^3} + O\left(\beta^2\right).$$

From this expression we then find

$$\begin{aligned}
\frac{\partial \Delta x}{\partial \gamma} &= \beta \frac{\lambda(\lambda-2\gamma)}{\left(2e^{c\eta}+e^{2c\eta}+2\right)(\gamma+\lambda)^4} + O\left(\beta^2\right) \\
\frac{\partial \Delta x}{\partial \lambda} &= \beta \frac{\gamma(\gamma-2\lambda)}{\left(2e^{c\eta}+e^{2c\eta}+2\right)(\gamma+\lambda)^4} + O\left(\beta^2\right) \\
\frac{\partial \Delta x}{\partial c} &= \beta \frac{2\gamma\eta\lambda e^{c\eta}\left(e^{c\eta}+1\right)}{\left(2e^{c\eta}+e^{2c\eta}+2\right)^2(\gamma+\lambda)^3} + O\left(\beta^2\right) \\
\frac{\partial \Delta x}{\partial \eta} &= \beta - \frac{2c\gamma\lambda e^{c\eta}\left(e^{c\eta}+1\right)}{\left(2e^{c\eta}+e^{2c\eta}+2\right)^2(\gamma+\lambda)^3} + O\left(\beta^2\right).
\end{aligned}$$

The trace of the Jacobian is $\mathrm{tr}(\mathbf{J}) = (1-2x)z\beta - \gamma - \lambda - \rho(1+g)$ and its determinant is $\det(\mathbf{J}) = ((2x-1)z\beta + \gamma + \lambda)\rho(1+g)$. The eigenvalues of the Jacobian are given by $\mu_{1,2} = \frac{1}{2}(\mathrm{tr}(\mathbf{J}) \pm \sqrt{\mathrm{tr}(\mathbf{J})^2 - 4\det(\mathbf{J})})$, so that we obtain complex eigenvalues if $\mathrm{tr}(\mathbf{J})^2 < 4\det(\mathbf{J})$, or equivalently

$$\left(\sqrt{4z\beta\gamma + (\gamma+\lambda-z\beta)^2} + \rho(1+g)\right)^2 < 4\sqrt{4z\beta\gamma + (\gamma+\lambda-z\beta)^2}\rho(1+g),$$

where $z = g/(1+g)$. Denoting by $y \equiv \sqrt{4z\beta\gamma + (\gamma+\lambda-z\beta)^2}$, this can be written as follows

$$(y + \rho(1+g))^2 < 4y\rho(1+g).$$

This is equivalent to $(y-\rho(1+g))^2 < 0$ which can never be satisfied for any admissible choice of the model's parameter values. Hence, when $\theta = 1$ this system does not show oscillatory behavior. Moreover, for the eigenvalues we obtain

$$\begin{aligned}
\mu_1 &= -\frac{\sqrt{(\gamma + g(-\beta+\gamma+\lambda)+\lambda)^2 + 4\beta\gamma g(g+1)}}{g+1} \\
\mu_2 &= -\rho(1+g),
\end{aligned}$$

which are both real and negative.

In the following we compute the fixed points for small values of τ when $\gamma = 0$ and $\theta = 0$. From

$$g \equiv \bar{g}(0,1) = \frac{e^{c(-\eta)}e^{\eta\left(\frac{2\beta b\tau}{bx+1}-c\right)}}{\left(e^{c(-\eta)}+1\right)\left(e^{\eta\left(\frac{2\beta b\tau}{bx+1}-c\right)}+1\right)},$$

we obtain for the stationary state

$$z = \frac{g}{1+g} = \frac{1}{e^{\eta\left(c-\frac{2b\beta\tau}{bx+1}\right)}\left(e^{\frac{2b\beta\eta\tau}{bx+1}}+e^{c\eta}+1\right)+2} = \frac{2b\beta\eta\tau e^{c\eta}\left(e^{c\eta}+1\right)}{(bx+1)\left(2e^{c\eta}+e^{2c\eta}+2\right)^2} + \frac{1}{2e^{c\eta}+e^{2c\eta}+2} + O\left(\tau^2\right).$$

Dropping terms of the order $O\left(\tau^2\right)$ and inserting into the stationary state condition $x(\beta nz - \lambda) - \beta nx^2 z = 0$ gives the solutions indicated in Equations (38) and (39).

Finally, in the case of $\eta \to \infty$ we have that $g = 0$ and the fixed point is given by $\frac{\gamma}{\gamma+\lambda}$ and $z = 0$. If $c > \frac{2b\beta\tau}{b+1}$ then the eigenvalues of the Jacobian are given by $\mu_1 = -(\gamma+\lambda)$ and $\mu_2 = -\rho$, which are both real and negative. □

Proof of Proposition 6. The stationary solution satisfies

$$\begin{aligned}
0 &= \gamma - \beta x^2 z_1 + x(-\gamma - \lambda + \beta z_1) \\
0 &= g_1\rho - (g_1+1)\rho z_1 + \frac{(1-x)(z_2-z_1)(\gamma+\beta x z_1)}{x} + \frac{\lambda x(z_3-z_1)}{1-x} \\
0 &= g_2\rho - (g_2+1)\rho z_2 + \frac{2\lambda x(z_1-z_2)}{1-x} \\
0 &= g_3\rho - (g_3+1)\rho z_3 + \frac{2(1-x)(z_1-z_3)(\gamma+\beta x z_1)}{x}
\end{aligned}$$

From this system of equations we obtain

$$\begin{aligned} x &= \frac{-\gamma-\lambda+\sqrt{(\gamma+\lambda-\beta z_1)^2+4\beta\gamma z_1}+\beta z_1}{2\beta z_1} \\ z_2 &= \frac{g_2\rho-g_2\rho x+2\lambda x z_1}{2\lambda x-(g_2+1)\rho(x-1)} \\ z_3 &= \frac{2(x-1)z_1(\gamma+\beta x z_1)-g_3\rho x}{2(x-1)(\gamma+\beta x z_1)-(g_3+1)\rho x}, \end{aligned}$$

and

$$\begin{aligned} 2g_1-2(1+g_1)z_1 = &\frac{2(g_3(-1+z_1)+z_1)\lambda\left(z_1\beta-\gamma-\lambda+A(z_1)\right)}{\left(z_1\beta+\gamma+\lambda-A(z_1)\right)(2\lambda+\rho+g_3\rho)} \\ &+\frac{\left(g_2(-1+z_1)+z_1\right)\left(z_1\beta+\gamma+\lambda-A(z_1)\right)^2\left(z_1\beta+\gamma-\lambda+A(z_1)\right)}{\left(z_1\beta-\gamma-\lambda+A(z_1)\right)\left(z_1\beta(2\lambda+\rho+g_2\rho)+\left(-\gamma-\lambda+A(z_1)\right)(2\lambda-(1+g_2)\rho)\right)}, \end{aligned}$$

with $A(z_1)\equiv\sqrt{4z_1\beta\gamma+(-z_1\beta+\gamma+\lambda)^2}$. In the case of $\gamma=0$ the non-trivial solution is then given by

$$\begin{aligned} x &= 1-\frac{\lambda}{\beta z_1} \\ z_2 &= \frac{g_2\rho-g_2\rho x+2\lambda x z_1}{2\lambda x-(g_2+1)\rho(x-1)} \\ z_3 &= \frac{g_3\rho-2\beta(x-1)z_1^2}{(g_3+1)\rho-2\beta(x-1)z_1}, \end{aligned}$$

while z_1 is the root of

$$g_1(1-z_1)-\frac{\lambda(g_2(z_1-1)+z_1)}{g_2\rho-2\lambda+\rho+2\beta z_1}+\frac{(g_3(z_1-1)+z_1)(\lambda-\beta z_1)}{g_3\rho+2\lambda+\rho}-z_1=0.$$

A (cumbersome) closed form solution to this equation exists and can be obtained upon request from the author. In first order of large ρ the above equation can be written as

$$0=\frac{1}{\rho}\left(\frac{(g_3(z_1-1)+z_1)(\lambda-\beta z_1)}{g_3+1}-\frac{\lambda(g_2(z_1-1)+z_1)}{g_2+1}\right)+O\left(\frac{1}{\rho^2}\right).$$

The solution to this equation is given by

$$\begin{aligned} z_1 = &-\frac{1}{2\beta(g_2+1)(g_3+1)} \\ &\times\Big(-\sqrt{(g_2+1)\left(4\beta(g_3+1)(g_1(g_2+1)(g_3+1)\rho+\lambda(g_2-g_3))\right)+(g_2+1)(\beta g_3-(g_1+1)(g_3+1)\rho)^2)} \\ &+(g_1+1)g_2(g_3+1)\rho+g_1g_3\rho+g_1\rho-\beta g_2g_3-\beta g_3+g_3\rho+\rho\Big). \end{aligned}$$

The critical level β^c is obtained from setting $1-\frac{\lambda}{\beta z_1}=0$. This yields

$$\beta^c=\frac{(g_2+1)\lambda(g_1\rho+\lambda+\rho)}{g_1g_2\rho+g_1\rho+g_2\lambda},$$

with the limit

$$\lim_{\rho\to\infty}\beta^c=\frac{\lambda(1+g_1)}{g_1}.$$

Inserting g_1 and g_2 into β^c yields

$$\beta^c=\frac{\lambda\left(2e^{c\eta}+e^{2c\eta}+2\right)\left((\lambda+2\rho)e^{b\beta\eta\tau}+(\lambda+\rho)e^{\eta(b\beta\tau+c)}+e^{c\eta}(\lambda+\rho)+e^{2c\eta}(\lambda+\rho)\right)}{\lambda\left(e^{c\eta}+1\right)\left(e^{b\beta\eta\tau}+e^{c\eta}\right)+\rho\left(2e^{c\eta}+e^{2c\eta}+2\right)e^{b\beta\eta\tau}}.$$

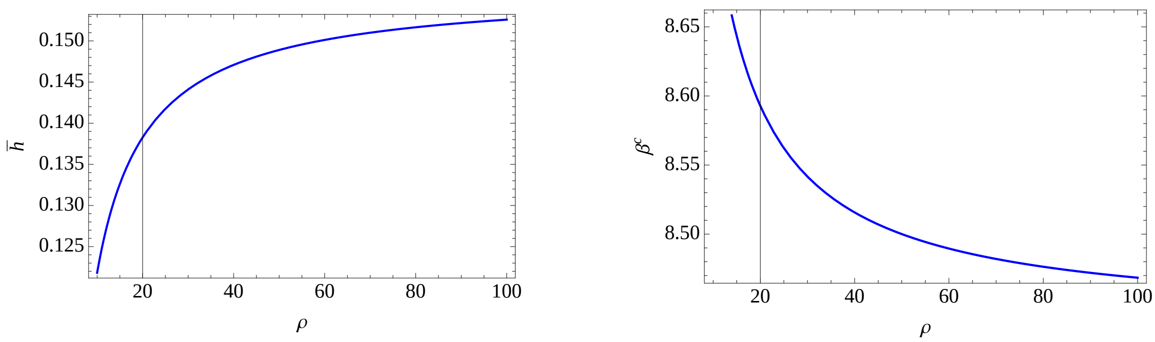


Figure 17: (Left panel) The average stock of knowledge $\bar{h}$, and (right panel) the critical threshold β^c as a function of ρ for $b = 10$, $\beta = 10$, $c = 1$, $\eta = 1$, $\lambda = 1$ and $\tau = 0.01$.

For the derivative we then obtain

$$\frac{\partial \beta^c}{\partial \rho} = -\frac{\lambda^2 e^{c\eta}\left(e^{c\eta}+1\right)^2\left(2e^{c\eta}+e^{2c\eta}+2\right)\left(e^{b\beta\eta\tau}-1\right)\left(e^{b\beta\eta\tau}+e^{c\eta}\right)}{\left((\lambda+2\rho)e^{b\beta\eta\tau}+(\lambda+2\rho)e^{\eta(b\beta\tau+c)}+\rho e^{b\beta\eta\tau+2c\eta}+\lambda e^{c\eta}+\lambda e^{2c\eta}\right)^2} < 0.$$

We next compute the Jacobian. For $\theta = 1$, it is given by

$$\mathbf{J} = \begin{bmatrix} z_1(\beta-2x\beta)-\gamma-\lambda-\mu & -(x-1)x\beta & 0 & \\ \frac{(z_1-z_2)\left(z_1\beta x^2+\gamma\right)}{x^2}+\frac{(z_3-z_1)\lambda}{(x-1)^2} & 2(x-1)z_1\beta+z_2(\beta-x\beta)+\gamma+\frac{x\lambda}{x-1}-\mu-(g_1+1)\rho-\frac{\gamma}{x} & -\frac{(x-1)(xz_1\beta+\gamma)}{x} & \\ \frac{2(z_1-z_2)\lambda}{(x-1)^2} & -\frac{2x\lambda}{x-1} & \frac{2x\lambda}{x-1}-\mu-(g_2+1)\rho & \\ -\frac{2(z_1-z_3)\left(z_1\beta x^2+\gamma\right)}{x^2} & -\frac{2(x-1)(x(2z_1-z_3)\beta+\gamma)}{x} & 0 & 2(x-1 \end{bmatrix}$$

In the case of $\gamma = 0$ we get

$$\mathbf{J} = \begin{bmatrix} -\lambda-\mu & (1-x)x\beta & 0 & 0 \\ 0 & z_2(\beta-x\beta)+\frac{x\lambda}{x-1}-\mu-(g_1+1)\rho & 0 & \frac{x\lambda}{1-x} \\ -\frac{2z_2\lambda}{(x-1)^2} & -\frac{2x\lambda}{x-1} & \frac{2x\lambda}{x-1}-\mu-(g_2+1)\rho & 0 \\ 0 & 0 & 0 & -\mu-(g_3+1)\rho \end{bmatrix}.$$

An LU decomposition of $\mathbf{J}-\mu\mathbf{I}_4 = \mathbf{LU}$ yields [cf. Horn and Johnson, 1990]

$$\mathbf{L} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ -\frac{2z_2\lambda}{(x-1)^2(-\lambda-\mu)} & \frac{\frac{2(1-x)xz_2\beta\lambda}{(x-1)^2(-\lambda-\mu)}-\frac{2x\lambda}{x-1}}{z_2(\beta-x\beta)+\frac{x\lambda}{x-1}-\mu-(g_1+1)\rho} & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

and

$$\mathbf{U} = \begin{bmatrix} -\lambda-\mu & (1-x)x\beta & 0 & 0 \\ 0 & z_2(\beta-x\beta)+\frac{x\lambda}{x-1}-\mu-(g_1+1)\rho & 0 & \frac{x\lambda}{1-x} \\ 0 & 0 & \frac{2x\lambda}{x-1}-\mu-(g_2+1)\rho & -\frac{x\lambda\left(\frac{2(1-x)xz_2\beta\lambda}{(x-1)^2(-\lambda-\mu)}-\frac{2x\lambda}{x-1}\right)}{(1-x)\left(z_2(\beta-x\beta)+\frac{x\lambda}{x-1}-\mu-(g_1+1)\rho\right)} \\ 0 & 0 & 0 & -\mu-(g_3+1)\rho \end{bmatrix}.$$

Using the fact that $\det(\mathbf{J}-\mu\mathbf{I}_4) = \det(\mathbf{L})\det(\mathbf{U})$, and that the determinant of a triangular matrix is given by the product of the diagonal entries, gives the following characteristic polynomial

$$\frac{1}{(x-1)^2}(\lambda+\mu)(g_3\rho+\mu+\rho)(-(g_2+1)\rho(x-1)+\mu+2\lambda x-\mu x)$$
$$\times\left(g_1\rho+x(-(g_1+1)\rho+\lambda-\mu)+\mu+\rho+\beta\left(-(x-1)^2\right)z_2\right)=0.$$

From this equation we obtain the eigenvalues

$$\mu_1 = -\lambda$$
$$\mu_2 = -(g_3+1)\rho$$
$$\mu_3 = \frac{2\lambda x}{x-1} - (g_2+1)\rho$$
$$\mu_4 = \frac{(g_1+1)\rho(1-x) - \lambda x + \beta(1-x)^2 z_2}{1-x}.$$

These eigenvalues are all real and negative for large ρ. □

Proof of Proposition 7. Note that when $N = 2$ with $s \in \{0, 1, 2\}$ we obtain from Equation (23)

$$\begin{aligned}
\frac{d\bar{x}_t(0)}{dt} &= 2\left(\lambda\bar{x}_t(1) - \gamma\bar{x}_t(0) - \beta(\bar{\xi}_t(0,1)\bar{x}_t(0)\bar{x}_t(1) + \bar{\xi}_t(0,2)\bar{x}_t(0)\bar{x}_t(2))\right)\\
\frac{d\bar{x}_t(1)}{dt} &= \gamma\bar{x}_t(0) + \lambda\bar{x}_t(2) - (\lambda+\gamma+\alpha)\bar{x}_t(1) + \beta\left(\bar{\xi}_t(0,1)\bar{x}_t(0)\bar{x}_t(1) - \bar{\xi}_t(1,1)\bar{x}_t(1)^2\right.\\
&\quad \left.+\bar{\xi}_t(0,2)\bar{x}_t(0)\bar{x}_t(2) - \bar{\xi}_t(1,2)\bar{x}_t(1)\bar{x}_t(2)\right)\\
\frac{d\bar{x}_t(2)}{dt} &= 2\left((\gamma+\alpha)\bar{x}_t(1) - \lambda\bar{x}_t(2) + \beta\left(\bar{\xi}_t(1,1)\bar{x}_t(1)^2 + \bar{\xi}_t(1,2)\bar{x}_t(1)\bar{x}_t(2)\right)\right),
\end{aligned} \tag{97}$$

and from Equation (24) we obtain

$$\begin{aligned}
\frac{d\bar{\xi}_t(0,1)}{dt} &= \frac{1}{2}\rho\tilde{g}(0,1) - \rho\left(1 + \frac{1}{2}\tilde{g}(0,1)\right)\bar{\xi}_t(0,1)\\
\frac{d\bar{\xi}_t(0,2)}{dt} &= \rho\tilde{g}(0,2) - \rho\left(1 + \tilde{g}(0,2)\right)\bar{\xi}_t(0,2)\\
\frac{d\bar{\xi}_t(1,1)}{dt} &= \frac{1}{4}\rho\tilde{g}(1,1) - \rho\left(1 + \frac{1}{4}\tilde{g}(1,1)\right)\bar{\xi}_t(1,1)\\
\frac{d\bar{z}_t(1,2)}{dt} &= \frac{1}{2}\rho\tilde{g}(1,2) - \rho\left(1 + \frac{1}{2}\tilde{g}(1,2)\right)\bar{\xi}_t(1,2).
\end{aligned} \tag{98}$$

From the definition in Equation (22) we find that

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

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

$$
\begin{aligned}
\lim_{\eta\to 0}\tilde{g}(0,1) &= \frac{1}{2}, \\
\lim_{\eta\to 0}\tilde{g}(0,2) &= \frac{1}{4}, \\
\lim_{\eta\to 0}\tilde{g}(1,1) &= 1, \\
\lim_{\eta\to 0}\tilde{g}(1,2) &= \frac{1}{2},
\end{aligned}
$$

while in the limit of $\eta \to \infty$ we obtain

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

We now proof part (i) of the proposition. Denote by $x_1 \equiv \lim_{t\to\infty}\bar{x}_t(1), x_2 \equiv \lim_{t\to\infty}\bar{x}_t(2)$ and $z_1 \equiv \lim_{t\to\infty}\bar{\xi}_t(0,1), z_2 \equiv \lim_{t\to\infty}\bar{\xi}_t(0,2), z_3 \equiv \lim_{t\to\infty}\bar{\xi}_t(1,1), z_4 \equiv \lim_{t\to\infty}\bar{\xi}_t(1,2)$ and also let $g_1 \equiv \tilde{g}(0,1)$, $g_2 \equiv \tilde{g}(0,2)$, $g_3 \equiv \tilde{g}(1,1)$ and $g_4 \equiv \tilde{g}(1,2)$. 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

$$\mathbf{J}_{11} = \begin{bmatrix} (1-4x_1-x_2)z_1\beta - 2\beta(x_2z_2+x_1z_3) - x_2z_4\beta - 3\gamma - \lambda & z_2\beta(1-2x_2) - x_1(z_1+2z_2+z_4)\beta - \gamma + \lambda \\ 2(2x_1z_3\beta + x_2z_4\beta + \gamma) & 2x_1z_4\beta - 2\lambda \end{bmatrix}$$

$$\mathbf{J}_{12} = \begin{bmatrix} x_1(1-2x_1-x_2)\beta & x_2(1-2x_1-x_2)\beta & -x_1^2\beta & -x_1x_2\beta \\ 0 & 0 & 2x_1^2\beta & 2x_1x_2\beta \end{bmatrix}$$

$$\mathbf{J}_{22} = \begin{bmatrix} -\rho\left(1+\frac{1}{2}g_1\right) & 0 & 0 & 0 \\ 0 & -\rho(1+g_2) & 0 & 0 \\ 0 & 0 & -\rho\left(1+\frac{1}{4}g_3\right) & 0 \\ 0 & 0 & 0 & -\rho\left(1+\frac{1}{2}g_4\right) \end{bmatrix}.$$

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

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

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

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

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

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

Hence, in both cases when $x_1 = x_2 = 0$ we have that[45]

$$\det(\mathbf{J} - \mu\mathbf{I}_6) = \det(\mathbf{J}_{11} - \mu\mathbf{I}_2)\det(\mathbf{J}_{22} - \mu\mathbf{I}_4).$$

We have that

$$\det(\mathbf{J}_{22} - \mu\mathbf{I}_4) = \frac{1}{16}(2\mu + g_1\rho)(\mu + g_2\rho)(4\mu + g_3\rho)(2\mu + g_4\rho),$$

The roots of the characteristic polynomial give us the eigenvalues $\mu_3 = -\rho\left(1+\frac{1}{2}g_1\right)$, $\mu_4 = -\rho(1+g_2)$, $\mu_5 = -\rho\left(1+\frac{1}{4}g_3\right)$, $\mu_6 = -\rho\left(1+\frac{1}{2}g_4\right)$. These are all negative an real. Further we have that at $x_1 = x_2 = 0$

$$\det(\mathbf{J}_{11} - \mu\mathbf{I}_2) = \alpha(2\gamma + \mu) + (\gamma + \lambda + \mu)(2(\gamma + \lambda) + \mu) - \beta(z_1(2\lambda + \mu) + 2z_2(\alpha + \gamma)).$$

[45] 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.

Setting $\alpha = \gamma = 0$ this simplifies to

$$\det(\mathbf{J}_{11} - \mu \mathbf{I}_2) = \lambda + \mu - \beta z_1)2\lambda + \mu$$

The roots are given by

$$\mu_1 = \beta z_1 - \lambda$$
$$\mu_2 = -2\lambda.$$

Inserting the stationary state

$$z_1 = \frac{g_1}{g_1 + 2}$$

delivers Equation (35). This completes the proof of part (i) of the proposition.

Next, we consider part (ii) of the proposition. For small values of τ we have that

$$\begin{aligned}
z_1 &= \frac{g_1}{g_1+2} = \frac{e^{b\beta\eta\tau}}{2e^{b\beta\eta\tau} + e^{\eta(b\beta\tau+c)} + e^{c\eta} + e^{2c\eta}} = \frac{b\beta\eta\tau e^{c\eta}\left(e^{c\eta}+1\right)}{\left(2e^{c\eta}+e^{2c\eta}+2\right)^2} + \frac{1}{2e^{c\eta}+e^{2c\eta}+2} + O\left(\tau^2\right) \\
z_2 &= \frac{g_2}{g_2+1} = \frac{e^{2b\beta\eta\tau}}{2e^{2b\beta\eta\tau} + e^{\eta(2b\beta\tau+c)} + e^{c\eta} + e^{2c\eta}} = \frac{2b\beta\eta\tau e^{c\eta}\left(e^{c\eta}+1\right)}{\left(2e^{c\eta}+e^{2c\eta}+2\right)^2} + \frac{1}{2e^{c\eta}+e^{2c\eta}+2} + O\left(\tau^2\right) \\
z_3 &= \frac{g_3}{g_3+4} = \frac{\frac{e^{2b\beta\eta\tau}}{(e^{b\beta\eta\tau}+e^{c\eta})^2} + \frac{1}{(e^{c\eta}+1)^2}}{\frac{e^{2b\beta\eta\tau}}{(e^{b\beta\eta\tau}+e^{c\eta})^2} + \frac{1}{(e^{c\eta}+1)^2} + 2} = \frac{b\beta\eta\tau e^{c\eta}\left(e^{c\eta}+1\right)}{\left(2e^{c\eta}+e^{2c\eta}+2\right)^2} + \frac{1}{2e^{c\eta}+e^{2c\eta}+2} + O\left(\tau^2\right) \\
z_4 &= \frac{g_4}{g_4+2} = \frac{e^{b\beta\eta\tau}}{2e^{b\beta\eta\tau} + e^{\eta(b\beta\tau+c)} + e^{c\eta} + e^{2c\eta}} = \frac{b\beta\eta\tau e^{c\eta}\left(e^{c\eta}+1\right)}{\left(2e^{c\eta}+e^{2c\eta}+2\right)^2} + \frac{1}{2e^{c\eta}+e^{2c\eta}+2} + O\left(\tau^2\right).
\end{aligned}$$

In the case of $\alpha = \gamma = 0$, the stationary state solves the equations

$$\begin{aligned}
0 &= \beta\left(x_1^2(-z_3) + x_1 z_1(-2x_1 - x_2 + 1) + x_2 z_2(-2x_1 - x_2 + 1) - x_1 x_2 z_4\right) - x_1\lambda + \gamma(-2x_1 - x_2 + 1) + \lambda x_2 \\
0 &= 2\left(\beta\left(x_1^2 z_3 + x_1 x_2 z_4\right) - \lambda x_2\right).
\end{aligned}$$

For x_2 we get

$$x_2 = \frac{\beta x_1^2 z_1}{\lambda - \beta x_1 z_1},$$

while solving for x_1 gives

$$\begin{aligned}
x_1 = {} & \frac{1}{3\sqrt[3]{A}\beta^2 z_1^2(4\beta(3\beta z_1(z_1 - z_2) + 2z_1 - z_2) + 1)} \\
& \times (-A^{2/3} + 2\sqrt[3]{A}\beta z_1(\lambda + \beta(\beta z_1^2(6\beta(z_1 - z_2) + 1) + \lambda(12\beta z_1(z_1 - z_2) + 9z_1 - 2z_2))) \\
& - 4\beta^6 z_1^6(6\beta(z_1 - z_2) + 1)^2 + 4\beta^4\lambda z_1^4(\beta(12(3\beta^2 z_1(z_1 - z_2)(2z_1 + z_2) \\
& + \beta(4z_1^2 - 2z_1 z_2 + z_2^2) + z_1) - 5z_2) + 1) - \beta^2\lambda^2 z_1^2(4\beta(144\beta^3 z_1^2(z_1 - z_2)^2 \\
& + 48\beta^2 z_1(3z_1^2 - 4z_1 z_2 + z_2^2) + \beta(48z_1^2 - 27z_1 z_2 + 4z_2^2) + 6z_1 - z_2) + 1))),
\end{aligned}$$

where $A = B + 6\sqrt{3C}$ with

$$\begin{aligned}
B = {} & \beta^3 z_1^3(-8\beta^6 z_1^6(6\beta(z_1 - z_2) + 1)^3 + 12\beta^4\lambda z_1^4(\beta(6(36\beta^3 z_1(z_1 - z_2)^2(2z_1 + z_2) \\
& + 6\beta^2(z_1 - z_2)(10z_1^2 - 3z_1 z_2 + 2z_2^2) + \beta(20z_1^2 - 21z_1 z_2 + 7z_2^2) + 3z_1) - 11z_2) + 1) \\
& - 6\beta^2\lambda^2 z_1^2(\beta(432\beta^4 z_1^2(z_1 - z_2)^2(5z_1 - 2z_2) + 72\beta^3 z_1(z_1 - z_2)(35z_1^2 - 30z_1 z_2 - 2z_2^2) \\
& + 12\beta^2(93z_1^3 - 112z_1^2 z_2 + 9z_1 z_2^2 + 4z_2^3) + 2\beta(117z_1^2 - 78z_1 z_2 - 4z_2^2) + 24z_1 - 7z_2) + 1) \\
& + \lambda^3(2\beta(432\beta^5 z_1^3(z_1 - z_2)^2(11z_1 + 16z_2) + 216\beta^4 z_1^2(z_1 - z_2)(33z_1^2 + 19z_1 z_2 - 16z_2^2) \\
& + 72\beta^3 z_1(60z_1^3 - 27z_1^2 z_2 - 23z_1 z_2^2 + 8z_2^3) + 2\beta^2(3z_1 - 2z_2)(225z_1^2 + 33z_1 z_2 - 8z_2^2) \\
& + 3\beta(75z_1^2 - 30z_1 z_2 - 4z_2^2) + 18z_1 - 3z_2) + 1)),
\end{aligned}$$

and

$$
\begin{aligned}
C = {} & \beta^8\lambda^2 z_1^6(4\beta(3\beta z_1(z_1-z_2)+2z_1-z_2)+1)^2(1296\beta^{10}z_1^8 z_2^2(z_1-z_2)^2 \\
& + 432\beta^9 z_1^8 z_2(z_1-z_2)(6\lambda z_1 - 18\lambda z_2 + z_2) + 36\beta^8 z_1^6(36\lambda^2(z_1-z_2)^2(z_1^2-10z_1z_2+4z_2^2)+z_1^2z_2^2 \\
& + 24\lambda z_2(z_1^3-5z_1^2z_2+3z_1z_2^2-z_2^3)) + 72\beta^7\lambda z_1^5(z_1z_2(z_1^2-11z_1z_2+5z_2^2)+6\lambda(z_1-z_2)(3z_1^3-22z_1^2z_2 \\
& + 31z_1z_2^2-4z_2^3) - 36\lambda^2 z_1(z_1-10z_2)(z_1-z_2)^2) + 12\beta^6\lambda z_1^4(3\lambda(12\lambda(15\lambda-8)+13)z_1^4 \\
& - 6\lambda(6\lambda(126\lambda-73)+37)z_1^3z_2 + (18\lambda(414\lambda^2-226\lambda+29)-5)z_1^2z_2^2 - 24\lambda(72\lambda(2\lambda-1)+7)z_1z_2^3 + 12\lambda z_2^4) \\
& + 24\beta^5\lambda^2 z_1^4((6\lambda(75\lambda-13)+3)z_1^3 + (9(55-336\lambda)\lambda-10)z_1^2z_2 + (3726\lambda^2-681\lambda+31)z_1z_2^2 \\
& - 2(3\lambda(192\lambda-35)+2)z_2^3) + \beta^4\lambda^2 z_1^3((12\lambda(597\lambda-43)+4)z_1^3 - 4(9144\lambda^2-507\lambda+1)z_1^2z_2 \\
& + (12\lambda(2703\lambda-253)+37)z_1z_2^2 + 432\lambda(1-16\lambda)z_2^3) + 2\beta^3\lambda^3 z_1^2(6(199\lambda-6)z_1^3 \\
& + (79-4350\lambda)z_1^2z_2 + 2(1506\lambda-71)z_1z_2^2 + 4(1-96\lambda)z_2^3) + \beta^2\lambda^3 z_1((413\lambda-4)z_1^3 + 4(1-250\lambda)z_1^2z_2 \\
& + 2(326\lambda-5)z_1z_2^2 - 32\lambda z_2^3) + 2\beta\lambda^4 z_1(17z_1^2-26z_1z_2+20z_2^2) + \lambda^4(z_1^2-z_1z_2+z_2^2)).
\end{aligned}
$$

To determine the stability of the fixed points, observe that in the case of $\tau \to 0$ the block element $\mathbf{J}_{21}$ of the Jacobian $\mathbf{J}$ is an all zero matrix, so 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 know that the characteristic polynomial $\det(\mathbf{J}_{22} - \mu \mathbf{I}_4)$ has only real eigenvalues. Moreover, in the case of $\gamma = 0$ and $c = 0$ where $z_1 = z_2 = z_3 = z_4 = \frac{1}{5}$ we have that

$$
\begin{aligned}
\det(\mathbf{J}_{11} - \mu \mathbf{I}_2) = {} & \frac{1}{25}\left(25(\alpha\mu + (\lambda+\mu)(2\lambda+\mu)) + 2\beta^2(x_1+x_2)(2x_1+2x_2-1)\right. \\
& \left. +5\beta(-2\lambda - \mu + \alpha(6x_1+4x_2-2) + 2(3\lambda+2\mu)(x_1+x_2))\right).
\end{aligned}
$$

The roots are given by

$$\mu_{1,2} = \frac{1}{10}\left(-5\alpha - 15\lambda \pm \sqrt{25\alpha^2 + 150\alpha\lambda + (5\lambda+\beta)^2 - 10\alpha\beta(8x_1+4x_2-3)} + \beta(-4x_1-4x_2+1)\right).$$

The eigenvalues are real if the term under the square root is positive. Using the fact that $x_1 = \frac{5\lambda(\beta-5\lambda)}{\beta^2}$ and $x_2 = \frac{(\beta-5\lambda)^2}{\beta^2}$ we find that this is equivalent to

$$25\alpha^2 - \frac{10\alpha(\beta-20\lambda)(5\lambda+\beta)}{\beta} + (5\lambda+\beta)^2 > (5\lambda+\beta)\frac{200\alpha\lambda + \beta^2 + 5\beta(\lambda-2\alpha)}{\beta} > 0,$$

which holds if $\alpha < \frac{\lambda}{2}$.

In the following we prove part (iii) of the proposition. In the case of $\eta \to 0$ we have that $z_1 = z_2 = z_3 = z_4 = \frac{1}{5}$. Setting also $\gamma = 0$ we obtain for the fixed points

$$
\begin{aligned}
0 &= \frac{2}{5}\beta x_1(x_1+x_2) + 2\alpha x_1 - 2\lambda x_2, \\
0 &= \frac{1}{5}\left(-3\beta x_1^2 + x_1(-5\alpha - 5\lambda + \beta - 4\beta x_2) + x_2(5\lambda + (\beta - \beta x_2))\right).
\end{aligned}
$$

The solutions are given by $x_1 = x_2 = 0$, or

$$x_2 = \frac{2\alpha(A - 5\alpha) + \lambda(-25\alpha + A - \beta)}{2\alpha\beta},$$

and

$$
\begin{aligned}
x_1 = \frac{1}{6\alpha\beta}\Big[& 15\alpha^2 + 45\alpha\lambda - 2A(2\alpha+\lambda) \\
&+\alpha\left(100\alpha\lambda^3 + \lambda^2\left(-175\alpha^2 + 2\beta(\beta - A) + 10\alpha\beta\right) + \alpha^2\left(25\alpha^2 + \beta(4A+5\beta) + 10\alpha\beta\right)\right. \\
&\left. +2\alpha\lambda\left((\beta-5\alpha)^2 - A\beta\right)\right)^{\frac{1}{2}} + \alpha^2\beta + 2\alpha\beta\lambda\Big],
\end{aligned}
$$

where we have denoted by $A \equiv \sqrt{100\alpha\lambda + (5\alpha+\beta n)^2}$. Since the block element $\mathbf{J}_{21}$ of the Jacobian $\mathbf{J}$ is an all zero matrix when $\eta \to 0$, the same reasoning as above shows that the Jacobian has only real eigenvalues.

Finally, we give a proof of part (iv) of the proposition. When $\beta b\tau > c$ in the case of $\theta = 1$ and $\frac{2b(b+1)\beta\tau}{4b+1} > c$ in the case of $\theta = 0$ then in the limit of $\eta \to \infty$ starting from an empty graph $\bar{K}_n$ we have that

$$
\begin{aligned}
z_1 &= z_2 = z_4 = 0, \\
z_3 &= \frac{1}{3}.
\end{aligned}
$$

The fixed points of Equation (43) then satisfy

$$
\begin{aligned}
0 &= 2\left(\frac{1}{3}\beta x_1^2 + x_1(\alpha+\gamma) - \lambda x_2\right) \\
0 &= \gamma - \frac{1}{3}\beta x_1^2 - x_1(\alpha + 3\gamma + \lambda) - \gamma x_2 + \lambda x_2
\end{aligned}
$$

and the solution is given by

$$
x_1 = \frac{-3\alpha\gamma - 3(\gamma+\lambda)^2 + \sqrt{12\beta\gamma^2\lambda + 9\left(\alpha\gamma + (\gamma+\lambda)^2\right)^2}}{2\beta\gamma},
$$

and

$$
\begin{aligned}
x_2 =& \frac{1}{12\beta\gamma^2\lambda}\left(3\alpha\gamma + 3\gamma^2 - 6\gamma\lambda - 3\lambda^2 + \sqrt{12\beta\gamma^2\lambda + 9\left(\alpha\gamma + (\gamma+\lambda)^2\right)^2}\right), \\
&\times\left(-3\alpha\gamma - 3\left(\gamma+\lambda\right)^2 + \sqrt{12\beta\gamma^2\lambda + 9\left(\alpha\gamma + (\gamma+\lambda)^2\right)^2}\right).
\end{aligned}
$$

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 $\mathbf{J}$ is an all zero matrix. The eigenvalues of the Jacobian are then determined by

$$
\det(\mathbf{J} - \mu\mathbf{I}_6) = \det(\mathbf{J}_{11} - \mu\mathbf{I}_2)\det(\mathbf{J}_{22} - \mu\mathbf{I}_4).
$$

We know that the characteristic polynomial $\det(\mathbf{J}_{22} - \mu\mathbf{I}_4)$ has only real eigenvalues. Moreover, when $z_1 = z_2 = z_4 = 0$ and $z_3 = \frac{1}{3}$ we have that

$$
\det(\mathbf{J}_{11} - \mu\mathbf{I}_2) = \alpha(2\gamma+\mu) + (\gamma+\lambda+\mu)(2(\gamma+\lambda)+\mu) + \frac{2}{3}\beta x_1(2\gamma+\mu).
$$

The roots give us the eigenvalues

$$
\mu_{1,2} = \frac{1}{6}\left(-3\alpha - 9\gamma - 9\lambda - \sqrt{9\lambda^2 + 18\lambda(3\alpha+\gamma+2\beta \mathrm{x1}) + (3\alpha - 3\gamma + 2\beta x_1)^2} - 2\beta x_1\right).
$$

Since the term under the square root cannot be negative, we find that the eigenvalues are real. In the case of $\theta = 0$ and $\frac{2b(b+1)\beta\tau}{4b+1} < c$, we have that $z_1 = z_2 = z_3 = z_4 = 0$ and the characteristic polynomial is given by

$$
\det(\mathbf{J}_{11} - \mu\mathbf{I}_2) = \alpha(2\gamma+\mu) + (\gamma+\lambda+\mu)(2(\gamma+\lambda)+\mu).
$$

The roots give us the eigenvalues

$$\mu_{1,2} = \frac{1}{2}\left(\pm\sqrt{\alpha^2 - 2\alpha(\gamma - 3\lambda) + (\gamma+\lambda)^2} - \alpha - 3\gamma - 3\lambda\right).$$

These eigenvalues are real if $\lambda > \frac{\gamma}{3}$. □

Proof of Proposition 11. 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 (99):

$$\begin{aligned}
n^2\mathsf{F}_z(\mathbf{h},\mathbf{h}') \equiv\; & n_t(\mathbf{h})n_t(\mathbf{h}')\rho\, g(\mathbf{h},\mathbf{h}')\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}') \quad A \\
& + \sum_{k=1}^{N} \mathbb{1}_{\{h_k=1\}} \left(\gamma + \alpha\langle \mathbf{h}-\mathbf{e}_k, \mathbf{u}\rangle\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\langle \mathbf{h}, \mathbf{u}\rangle\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})} \quad B \\
& + \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{e}_k,\mathbf{h}')}{n_t(\mathbf{h}-\mathbf{e}_k)} + \mathbb{1}_{\{\mathbf{h}''=\mathbf{h}'\}}\left(1 + \frac{\tau_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}')}{x_t(\mathbf{h})} + \mathbb{1}_{\{\mathbf{h}''=\mathbf{h}'\}}\left(1 + \frac{\tau_t(\mathbf{h}',\mathbf{h},\mathbf{h}')}{m_t(\mathbf{h},\mathbf{h}')}\right)\right] \quad C \\
& + \sum_{k=1}^{N} \mathbb{1}_{\{h_k'=1\}} \left(\gamma + \alpha\langle \mathbf{h}'-\mathbf{e}_k, \mathbf{u}\rangle\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\langle \mathbf{h}', \mathbf{u}\rangle\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}')} \quad B' \\
& + \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{\tau_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{\tau_t(\mathbf{h},\mathbf{h}',\mathbf{h})}{m_t(\mathbf{h}',\mathbf{h})}\right)\right]. \quad C'
\end{aligned} \tag{99}$$

We now explain each of the terms on the RHS in Equation (99). 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 $\rho n_t(\mathbf{h})n_t(\mathbf{h}')g(\mathbf{h},\mathbf{h}')\left(1 - \frac{m_t(\mathbf{h},\mathbf{h}')}{n_t(\mathbf{h})n_t(\mathbf{h}')}\right)$, 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 $g(\mathbf{h},\mathbf{h}')$ 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.

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 trough 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 $(\gamma + \alpha\langle\mathbf{h}-\mathbf{e}_k,\mathbf{u}\rangle)\, n_t(\mathbf{h}-\mathbf{e}_k)$. 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 $k = 1,\dots,N$ 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 $(\gamma + \alpha\langle\mathbf{h},\mathbf{u}\rangle)\, n_t(\mathbf{h})$ 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 $\frac{\tau_t(\mathbf{h}',\mathbf{h}-\mathbf{e}_k,\mathbf{h}')}{m_t(\mathbf{h}-\mathbf{e}_k,\mathbf{h}')}$. 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 $\frac{\tau_t(\mathbf{h}',\mathbf{h},\mathbf{h}')}{m_t(\mathbf{h},\mathbf{h}')}$. 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.

In the following we make a *pair approximation* as in Gross et al. [2006]; Keeling and Eames [2005]:

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

In particular, we then obtain

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

Introducing the rescaled variables $\beta \to \beta/n$, $c \to c/n$, $\delta/n \to \delta$ ($\beta \to \beta/n$, $c \to c/n$ and $\eta/\delta \to \eta n/\delta$, respectively), and dividing by n^2 we can write Equation (99) as

$$
\begin{aligned}
\mathsf{F}_z(\mathbf{h},\mathbf{h}') &= x_t(\mathbf{h})x_t(\mathbf{h}')\left[\rho\, g(\mathbf{h},\mathbf{h}')\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]\\
&+\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)\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\langle\mathbf{h},\mathbf{u}\rangle\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})}\\
&+\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]\\
&-\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\langle\mathbf{h}'-\mathbf{e}_k,\mathbf{u}\rangle\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\langle\mathbf{h}',\mathbf{u}\rangle\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}')}\\
&+\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]\\
&-\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],
\end{aligned}
\tag{100}
$$

Equation (100) can be further written as

$$
\begin{aligned}
&\mathsf{F}_z(\mathbf{h},\mathbf{h}') = x_t(\mathbf{h})x_t(\mathbf{h}')\left[\rho\, g(\mathbf{h},\mathbf{h}')\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]\\
&+\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)\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\langle\mathbf{h},\mathbf{u}\rangle\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})}\\
&+\beta\sum_{k=1}^{N}\mathbb{1}_{\{h_k=1\}}\sum_{\mathbf{h}''\in\mathcal{H}^N:h''_k=1}\frac{z_t(\mathbf{h}-\mathbf{e}_k,\mathbf{h}'')z_t(\mathbf{h}-\mathbf{e}_k,\mathbf{h}')}{x_t(\mathbf{h}-\mathbf{e}_k)}+\frac{\beta}{n}\sum_{k=1}^{N}\mathbb{1}_{\{h_k=1\}}\mathbb{1}_{\{h'_k=1\}}z_t(\mathbf{h}-\mathbf{e}_k,\mathbf{h}')\\
&-\beta\sum_{k=1}^{N}\mathbb{1}_{\{h_k=0\}}\sum_{\mathbf{h}''\in\mathcal{H}^N:h''_k=1}\frac{z_t(\mathbf{h},\mathbf{h}'')z_t(\mathbf{h},\mathbf{h}')}{x_t(\mathbf{h})}-\frac{\beta}{n}\sum_{k=1}^{N}\mathbb{1}_{\{h_k=0\}}\mathbb{1}_{\{h'_k=1\}}z_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)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\langle\mathbf{h}',\mathbf{u}\rangle\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}')}\\
&+\beta\sum_{k=1}^{N}\mathbb{1}_{\{h'_k=1\}}\sum_{\mathbf{h}''\in\mathcal{H}^N:h''_k=1}\frac{z_t(\mathbf{h}'-\mathbf{e}_k,\mathbf{h}'')z_t(\mathbf{h},\mathbf{h}'-\mathbf{e}_k)}{x_t(\mathbf{h}'-\mathbf{e}_k)}+\frac{\beta}{n}\sum_{k=1}^{N}\mathbb{1}_{\{h'_k=1\}}\mathbb{1}_{\{h_k=1\}}z_t(\mathbf{h}'-\mathbf{e}_k,\mathbf{h})\\
&-\beta\sum_{k=1}^{N}\mathbb{1}_{\{h'_k=0\}}\sum_{\mathbf{h}''\in\mathcal{H}^N:h''_k=1}\frac{z_t(\mathbf{h}',\mathbf{h}'')z_t(\mathbf{h}',\mathbf{h})}{x_t(\mathbf{h}')}-\frac{\beta}{n}\sum_{k=1}^{N}\mathbb{1}_{\{h'_k=0\}}\mathbb{1}_{\{h_k=1\}}z_t(\mathbf{h}',\mathbf{h}). \qquad (101)
\end{aligned}
$$

From Equation (101) we obtain

$$
\begin{aligned}
\mathsf{F}_z(\mathbf{h},\mathbf{h}') &= x_t(\mathbf{h})x_t(\mathbf{h}')\left[\rho\, g(\mathbf{h},\mathbf{h}')\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})}\left[\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)-\sum_{k=1}^{N}\mathbb{1}_{\{h_k=0\}}\left(\gamma+\alpha\langle\mathbf{h},\mathbf{u}\rangle\right)x_t(\mathbf{h})\right. \\
&\left.+\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})\right] \\
&+\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) \\
&+\beta\sum_{k=1}^{N}\mathbb{1}_{\{h_k=1\}}\left[\sum_{\mathbf{h}''\in\mathcal{H}^N:h''_k=1}\frac{z_t(\mathbf{h}-\mathbf{e}_k,\mathbf{h}'')z_t(\mathbf{h},\mathbf{h}')}{x_t(\mathbf{h})}\right. \\
&\left.+\mathbb{1}_{\{h'_k=1\}}\frac{z_t(\mathbf{h}-\mathbf{e}_k,\mathbf{h}')}{n}+\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}'')\right] \\
&-\beta\sum_{k=1}^{N}\mathbb{1}_{\{h_k=0\}}\left[\sum_{\mathbf{h}''\in\mathcal{H}^N:h''_k=1}\frac{z_t(\mathbf{h},\mathbf{h}'')z_t(\mathbf{h},\mathbf{h}')}{x_t(\mathbf{h})}+\mathbb{1}_{\{h'_k=1\}}\frac{z_t(\mathbf{h},\mathbf{h}')}{n}\right] \\
&+\frac{z_t(\mathbf{h},\mathbf{h}')}{x_t(\mathbf{h}')}\left[\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)-\sum_{k=1}^{N}\mathbb{1}_{\{h'_k=0\}}\left(\gamma+\alpha\langle\mathbf{h}',\mathbf{u}\rangle\right)x_t(\mathbf{h}')\right. \\
&\left.+\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}')\right] \\
&+\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) \\
&+\beta\sum_{k=1}^{N}\mathbb{1}_{\{h'_k=1\}}\left[\sum_{\mathbf{h}''\in\mathcal{H}^N:h''_k=1}\frac{z_t(\mathbf{h}'-\mathbf{e}_k,\mathbf{h}'')z_t(\mathbf{h},\mathbf{h}')}{x_t(\mathbf{h}')}\right. \\
&\left.+\mathbb{1}_{\{h_k=1\}}\frac{z_t(\mathbf{h}'-\mathbf{e}_k,\mathbf{h})}{n}+\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}'')\right] \\
&-\beta\sum_{k=1}^{N}\mathbb{1}_{\{h'_k=0\}}\left[\sum_{\mathbf{h}''\in\mathcal{H}^N:h''_k=1}\frac{z_t(\mathbf{h}',\mathbf{h}'')z_t(\mathbf{h}',\mathbf{h})}{x_t(\mathbf{h}')}+\mathbb{1}_{\{h_k=1\}}\frac{z_t(\mathbf{h}',\mathbf{h})}{n}\right].
\end{aligned}
\tag{102}
$$

This can be written as follows

$$
\begin{aligned}
\mathsf{F}_z(\mathbf{h},\mathbf{h}') &= x_t(\mathbf{h})x_t(\mathbf{h}')\left[\rho\, g(\mathbf{h},\mathbf{h}')\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)\\
&+\beta\sum_{k=1}^{N}\mathbb{1}_{\{h_k=1\}}\left[\mathbb{1}_{\{h'_k=1\}}\frac{z_t(\mathbf{h}-\mathbf{e}_k,\mathbf{h}')}{n}+\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}'')\right]\\
&-\frac{\beta}{n}\sum_{k=1}^{N}\mathbb{1}_{\{h_k=0\}}\mathbb{1}_{\{h'_k=1\}}z_t(\mathbf{h},\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)\\
&+\beta\sum_{k=1}^{N}\mathbb{1}_{\{h'_k=1\}}\left[\mathbb{1}_{\{h_k=1\}}\frac{z_t(\mathbf{h}'-\mathbf{e}_k,\mathbf{h})}{n}+\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}'')\right]\\
&-\frac{\beta}{n}\sum_{k=1}^{N}\mathbb{1}_{\{h'_k=0\}}\mathbb{1}_{\{h_k=1\}}z_t(\mathbf{h}',\mathbf{h}). \qquad (103)
\end{aligned}
$$

This is

$$
\begin{aligned}
\mathsf{F}_z(\mathbf{h},\mathbf{h}') &= x_t(\mathbf{h})x_t(\mathbf{h}')\left[\rho\, g(\mathbf{h},\mathbf{h}')\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)\\
&+\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)\\
&+\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{\beta}{n}\sum_{k=1}^{N}\mathbb{1}_{\{h_k=1\}}\mathbb{1}_{\{h'_k=1\}}z_t(\mathbf{h}-\mathbf{e}_k,\mathbf{h}')+\frac{\beta}{n}\sum_{k=1}^{N}\mathbb{1}_{\{h'_k=1\}}\mathbb{1}_{\{h_k=1\}}z_t(\mathbf{h}'-\mathbf{e}_k,\mathbf{h})\\
&-\frac{\beta}{n}z_t(\mathbf{h},\mathbf{h}')\sum_{k=1}^{N}\left(\mathbb{1}_{\{h_k=0\}}\mathbb{1}_{\{h'_k=1\}}+\mathbb{1}_{\{h'_k=0\}}\mathbb{1}_{\{h_k=1\}}\right). \qquad (104)
\end{aligned}
$$

Using the fact that

$$\sum_{k=1}^{N} \mathbb{1}_{\{h_k=0\}} \mathbb{1}_{\{h'_k=1\}} = \langle \mathbf{h}^c, \mathbf{h}' \rangle,$$
$$\sum_{k=1}^{N} \mathbb{1}_{\{h'_k=0\}} \mathbb{1}_{\{h_k=1\}} = \langle \mathbf{h}'^c, \mathbf{h} \rangle,$$

we obtain

$$\begin{aligned}
\mathsf{F}_z(\mathbf{h}, \mathbf{h}') &= x_t(\mathbf{h}) x_t(\mathbf{h}') \left[\rho\, g(\mathbf{h}, \mathbf{h}') \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) \\
&+ \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) \\
&+ \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{\beta}{n} \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) - \frac{\beta}{n} z_t(\mathbf{h}, \mathbf{h}') (\langle \mathbf{h}^c, \mathbf{h}' \rangle + \langle \mathbf{h}'^c, \mathbf{h} \rangle). \qquad (105)
\end{aligned}$$

Using the fact that $\operatorname*{plim}_{n\to\infty} \frac{dz_t(\mathbf{h},\mathbf{h}')}{dt} = \mathsf{F}_z(\mathbf{h}, \mathbf{h}')$,[46] and dropping terms of the order $O\left(\frac{1}{n}\right)$ in Equation (105)

[46]The same argument as in the proof of Theorem 1 holds. In particular, note that the RHS of Equation (105) 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.

yields

$$
\begin{aligned}
\frac{dz_t(\mathbf{h},\mathbf{h}')}{dt} &= x_t(\mathbf{h})x_t(\mathbf{h}')\left[\rho\, g(\mathbf{h},\mathbf{h}')\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)\\
&+\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)\\
&+\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}''). \qquad (106)
\end{aligned}
$$

Equations (74) and (107) now represent a closed system for the variables $x_t(\mathbf{h})$ and $z_t(\mathbf{h},\mathbf{h}')$. We next introduce the variable

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

for which we have that

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

Using Equation (107) we then get

$$\begin{aligned}
\frac{d\xi_t(\mathbf{h},\mathbf{h}')}{dt} &= \rho g(\mathbf{h},\mathbf{h}')\,(1-\xi_t(\mathbf{h},\mathbf{h}')) - \rho\,\xi_t(\mathbf{h},\mathbf{h}') \\
&+ \xi_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) \\
&- \xi_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) \\
&+ \sum_{k=1}^{N} \mathbb{1}_{\{h_k=1\}}\,(\gamma+\alpha\langle \mathbf{h}-\mathbf{e}_k,\mathbf{u}\rangle)\,\frac{x_t(\mathbf{h}-\mathbf{e}_k)}{x_t(\mathbf{h})}\left(\frac{z_t(\mathbf{h}-\mathbf{e}_k,\mathbf{h}')}{x_t(\mathbf{h}-\mathbf{e}_k)x_t(\mathbf{h}')} - \frac{z_t(\mathbf{h},\mathbf{h}')}{x_t(\mathbf{h})x_t(\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(\frac{z_t(\mathbf{h}+\mathbf{e}_k,\mathbf{h}')}{x_t(\mathbf{h}+\mathbf{e}_k)x_t(\mathbf{h}')} - \frac{z_t(\mathbf{h},\mathbf{h}')}{x_t(\mathbf{h})x_t(\mathbf{h}')}\right) \\
&+ \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)x_t(\mathbf{h}')} - \frac{z_t(\mathbf{h},\mathbf{h}')}{x_t(\mathbf{h})x_t(\mathbf{h}')}\right) \sum_{\mathbf{h}''\in\mathcal{H}^N : h''_k=1} \frac{z_t(\mathbf{h}-\mathbf{e}_k,\mathbf{h}'')}{x_t(\mathbf{h})} \\
&+ \sum_{k=1}^{N} \mathbb{1}_{\{h'_k=1\}}\,(\gamma+\alpha\langle \mathbf{h}'-\mathbf{e}_k,\mathbf{u}\rangle)\,\frac{x_t(\mathbf{h}'-\mathbf{e}_k)}{x_t(\mathbf{h}')}\left(\frac{z_t(\mathbf{h},\mathbf{h}'-\mathbf{e}_k)}{x_t(\mathbf{h}'-\mathbf{e}_k)x_t(\mathbf{h})} - \frac{z_t(\mathbf{h},\mathbf{h}')}{x_t(\mathbf{h}')x_t(\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(\frac{z_t(\mathbf{h},\mathbf{h}'+\mathbf{e}_k)}{x_t(\mathbf{h}'+\mathbf{e}_k)x_t(\mathbf{h})} - \frac{z_t(\mathbf{h},\mathbf{h}')}{x_t(\mathbf{h}')x_t(\mathbf{h})}\right) \\
&+ \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)x_t(\mathbf{h})} - \frac{z_t(\mathbf{h},\mathbf{h}')}{x_t(\mathbf{h}')x_t(\mathbf{h})}\right) \sum_{\mathbf{h}''\in\mathcal{H}^N : h''_k=1} \frac{z_t(\mathbf{h}'-\mathbf{e}_k,\mathbf{h}'')}{x_t(\mathbf{h}')}. \qquad (107)
\end{aligned}$$

This is

$$\begin{aligned}
\frac{d\xi_t(\mathbf{h},\mathbf{h}')}{dt} &= \rho g(\mathbf{h},\mathbf{h}')\,(1-\xi_t(\mathbf{h},\mathbf{h}')) - \rho\,\xi_t(\mathbf{h},\mathbf{h}') \\
&+ \sum_{k=1}^{N} \mathbb{1}_{\{h_k=1\}}\,(\gamma+\alpha\langle \mathbf{h}-\mathbf{e}_k,\mathbf{u}\rangle)\,\frac{x_t(\mathbf{h}-\mathbf{e}_k)}{x_t(\mathbf{h})}\,(\xi_t(\mathbf{h}-\mathbf{e}_k,\mathbf{h}') - \xi_t(\mathbf{h},\mathbf{h}')) \\
&+ \lambda \sum_{k=1}^{N} \mathbb{1}_{\{h_k=0\}}\frac{x_t(\mathbf{h}+\mathbf{e}_k)}{x_t(\mathbf{h})}\,(\xi_t(\mathbf{h}+\mathbf{e}_k,\mathbf{h}') - \xi_t(\mathbf{h},\mathbf{h}')) \\
&+ \beta \sum_{k=1}^{N} \mathbb{1}_{\{h_k=1\}}\,(\xi_t(\mathbf{h}-\mathbf{e}_k,\mathbf{h}') - \xi_t(\mathbf{h},\mathbf{h}')) \sum_{\mathbf{h}''\in\mathcal{H}^N : h''_k=1} \xi_t(\mathbf{h}-\mathbf{e}_k,\mathbf{h}'')\frac{x_t(\mathbf{h}-\mathbf{e}_k)x_t(\mathbf{h}'')}{x_t(\mathbf{h})} \\
&+ \sum_{k=1}^{N} \mathbb{1}_{\{h'_k=1\}}\,(\gamma+\alpha\langle \mathbf{h}'-\mathbf{e}_k,\mathbf{u}\rangle)\,\frac{x_t(\mathbf{h}'-\mathbf{e}_k)}{x_t(\mathbf{h}')}\,(\xi_t(\mathbf{h},\mathbf{h}'-\mathbf{e}_k) - \xi_t(\mathbf{h},\mathbf{h}')) \\
&+ \lambda \sum_{k=1}^{N} \mathbb{1}_{\{h'_k=0\}}\frac{x_t(\mathbf{h}'+\mathbf{e}_k)}{x_t(\mathbf{h}')}\,(\xi_t(\mathbf{h},\mathbf{h}'+\mathbf{e}_k) - \xi_t(\mathbf{h},\mathbf{h}')) \\
&+ \beta \sum_{k=1}^{N} \mathbb{1}_{\{h'_k=1\}}\,(\xi_t(\mathbf{h}'-\mathbf{e}_k,\mathbf{h}) - \xi_t(\mathbf{h},\mathbf{h}')) \sum_{\mathbf{h}''\in\mathcal{H}^N : h''_k=1} \xi_t(\mathbf{h}'-\mathbf{e}_k,\mathbf{h}'')\frac{x_t(\mathbf{h}'-\mathbf{e}_k)x_t(\mathbf{h}'')}{x_t(\mathbf{h}')}. \qquad (108)
\end{aligned}$$

We then get

$$
\begin{aligned}
\frac{d\xi_t(\mathbf{h},\mathbf{h}')}{dt} &= \rho g(\mathbf{h},\mathbf{h}')\,(1-\xi_t(\mathbf{h},\mathbf{h}')) - \rho\,\xi_t(\mathbf{h},\mathbf{h}') \\
&+ \sum_{k=1}^{N} \mathbb{1}_{\{h_k=1\}}\,(\gamma+\alpha\langle \mathbf{h}-\mathbf{e}_k,\mathbf{u}\rangle)\,\frac{x_t(\mathbf{h}-\mathbf{e}_k)}{x_t(\mathbf{h})}\,(\xi_t(\mathbf{h}-\mathbf{e}_k,\mathbf{h}')-\xi_t(\mathbf{h},\mathbf{h}')) \\
&+ \sum_{k=1}^{N} \mathbb{1}_{\{h'_k=1\}}\,(\gamma+\alpha\langle \mathbf{h}'-\mathbf{e}_k,\mathbf{u}\rangle)\,\frac{x_t(\mathbf{h}'-\mathbf{e}_k)}{x_t(\mathbf{h}')}\,(\xi_t(\mathbf{h},\mathbf{h}'-\mathbf{e}_k)-\xi_t(\mathbf{h},\mathbf{h}')) \\
&+ \lambda\sum_{k=1}^{N} \mathbb{1}_{\{h_k=0\}}\frac{x_t(\mathbf{h}+\mathbf{e}_k)}{x_t(\mathbf{h})}\,(\xi_t(\mathbf{h}+\mathbf{e}_k,\mathbf{h}')-\xi_t(\mathbf{h},\mathbf{h}')) \\
&+ \lambda\sum_{k=1}^{N} \mathbb{1}_{\{h'_k=0\}}\frac{x_t(\mathbf{h}'+\mathbf{e}_k)}{x_t(\mathbf{h}')}\,(\xi_t(\mathbf{h},\mathbf{h}'+\mathbf{e}_k)-\xi_t(\mathbf{h},\mathbf{h}')) \\
&+ \beta\sum_{k=1}^{N} \mathbb{1}_{\{h_k=1\}}\,(\xi_t(\mathbf{h}-\mathbf{e}_k,\mathbf{h}')-\xi_t(\mathbf{h},\mathbf{h}')) \sum_{\mathbf{h}''\in\mathcal{H}^N:h''_k=1} \xi_t(\mathbf{h}-\mathbf{e}_k,\mathbf{h}'')\frac{x_t(\mathbf{h}-\mathbf{e}_k)x_t(\mathbf{h}'')}{x_t(\mathbf{h})} \\
&+ \beta\sum_{k=1}^{N} \mathbb{1}_{\{h'_k=1\}}\,(\xi_t(\mathbf{h}'-\mathbf{e}_k,\mathbf{h})-\xi_t(\mathbf{h},\mathbf{h}')) \sum_{\mathbf{h}''\in\mathcal{H}^N:h''_k=1} \xi_t(\mathbf{h}'-\mathbf{e}_k,\mathbf{h}'')\frac{x_t(\mathbf{h}'-\mathbf{e}_k)x_t(\mathbf{h}'')}{x_t(\mathbf{h}')}. \qquad (109)
\end{aligned}
$$

□

Proof of Proposition 12. 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 (90) into Equation (55) gives

$$
\begin{aligned}
&\frac{d\bar{\xi}_t(s,s')}{dt} = \rho\bar{g}(s,s') - \rho(1+\bar{g}(s,s'))\bar{\xi}_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\}}\,(\gamma+\alpha\langle \mathbf{h}-\mathbf{e}_k,\mathbf{u}\rangle)\,\frac{x_t(\mathbf{h}-\mathbf{e}_k)}{x_t(\mathbf{h})}\,(\xi_t(\mathbf{h}-\mathbf{e}_k,\mathbf{h}')-\xi_t(\mathbf{h},\mathbf{h}')) \\
&+ \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\}}\,(\gamma+\alpha\langle \mathbf{h}'-\mathbf{e}_k,\mathbf{u}\rangle)\,\frac{x_t(\mathbf{h}'-\mathbf{e}_k)}{x_t(\mathbf{h}')}\,(\xi_t(\mathbf{h},\mathbf{h}'-\mathbf{e}_k)-\xi_t(\mathbf{h},\mathbf{h}')) \\
&+ \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})}\,(\xi_t(\mathbf{h}+\mathbf{e}_k,\mathbf{h}')-\xi_t(\mathbf{h},\mathbf{h}')) \\
&+ \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}')}\,(\xi_t(\mathbf{h},\mathbf{h}'+\mathbf{e}_k)-\xi_t(\mathbf{h},\mathbf{h}')) \\
&+ \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\}}\,(\xi_t(\mathbf{h}-\mathbf{e}_k,\mathbf{h}')-\xi_t(\mathbf{h},\mathbf{h}')) \sum_{\mathbf{h}''\in\mathcal{H}^N:h''_k=1} \xi_t(\mathbf{h}-\mathbf{e}_k,\mathbf{h}'')\frac{x_t(\mathbf{h}-\mathbf{e}_k)x_t(\mathbf{h}'')}{x_t(\mathbf{h})} \\
&+ \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\}}\,(\xi_t(\mathbf{h}'-\mathbf{e}_k,\mathbf{h})-\xi_t(\mathbf{h},\mathbf{h}')) \sum_{\mathbf{h}''\in\mathcal{H}^N:h''_k=1} \xi_t(\mathbf{h}'-\mathbf{e}_k,\mathbf{h}'')\frac{x_t(\mathbf{h}'-\mathbf{e}_k)x_t(\mathbf{h}'')}{x_t(\mathbf{h}')}.
\end{aligned}
\qquad (110)
$$

Note that

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

It then follows that

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

Similarly, we have that

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

Hence, we get

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

Next, we have that

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

Thus we get

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

Similarly, we have that

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

We then get

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

Moreover, we have that

$$\sum_{\mathbf{h}'\in\mathcal{H}^N:|\mathsf{S}(\mathbf{h})|=s'}\ \sum_{\mathbf{h}\in\mathcal{H}^N:|\mathsf{S}(\mathbf{h})|=s}\ \sum_{k\in\mathsf{S}(\mathbf{h})}(\xi_t(\mathbf{h}-\mathbf{e}_k,\mathbf{h}')-\xi_t(\mathbf{h},\mathbf{h}'))\,\frac{x_t(\mathbf{h}-\mathbf{e}_k)}{x_t(\mathbf{h})}\sum_{\mathbf{h}''\in\mathcal{H}^N:h''_k=1}\xi_t(\mathbf{h}-\mathbf{e}_k,\mathbf{h}'')x_t(\mathbf{h}'')$$

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

$$=\sum_{\mathbf{h}'\in\mathcal{H}^N:|\mathsf{S}(\mathbf{h})|=s'}\ \sum_{\mathbf{h}\in\mathcal{H}^N:|\mathsf{S}(\mathbf{h})|=s-1}\ \sum_{k\in\mathsf{S}(\mathbf{h}^c)}\left(\bar{\xi}_t(s-1,s')-\bar{\xi}_t(s,s')\right)\frac{\bar{x}_t(s-1)}{\bar{x}_t(s)}\sum_{s''=1}^{N}\binom{N}{s''-1}\bar{\xi}_t(s-1,s'')\bar{x}_t(s'')$$

$$=\sum_{\mathbf{h}'\in\mathcal{H}^N:|\mathsf{S}(\mathbf{h})|=s'}\ \sum_{\mathbf{h}\in\mathcal{H}^N:|\mathsf{S}(\mathbf{h})|=s-1}(N-s+1)\left(\bar{\xi}_t(s-1,s')-\bar{\xi}_t(s,s')\right)\frac{\bar{x}_t(s-1)}{\bar{x}_t(s)}\sum_{s''=1}^{N}\binom{N}{s''-1}\bar{\xi}_t(s-1,s'')\bar{x}_t(s'')$$

$$=\binom{N}{s-1}\binom{N}{s'}(N-s+1)\left(\bar{\xi}_t(s-1,s')-\bar{\xi}_t(s,s')\right)\frac{\bar{x}_t(s-1)}{\bar{x}_t(s)}\sum_{s''=1}^{N}\binom{N}{s''-1}\bar{\xi}_t(s-1,s'')\bar{x}_t(s'').$$

We then get

$$\frac{1}{\binom{N}{s}}\frac{1}{\binom{N}{s'}}\sum_{\mathbf{h}'\in\mathcal{H}^N:|\mathsf{S}(\mathbf{h})|=s'}\ \sum_{\mathbf{h}\in\mathcal{H}^N:|\mathsf{S}(\mathbf{h})|=s}\ \sum_{k\in\mathsf{S}(\mathbf{h})}(\xi_t(\mathbf{h}-\mathbf{e}_k,\mathbf{h}')-\xi_t(\mathbf{h},\mathbf{h}'))\,\frac{x_t(\mathbf{h}-\mathbf{e}_k)}{x_t(\mathbf{h})}\sum_{\mathbf{h}''\in\mathcal{H}^N:h''_k=1}\xi_t(\mathbf{h}-\mathbf{e}_k,\mathbf{h}'')x_t(\mathbf{h}'')$$

$$=s\left(\bar{\xi}_t(s-1,s')-\bar{\xi}_t(s,s')\right)\frac{\bar{x}_t(s-1)}{\bar{x}_t(s)}\sum_{s''=1}^{N}\binom{N}{s''-1}\bar{\xi}_t(s-1,s'')\bar{x}_t(s'').$$

Collecting the above terms in Equation (110) delivers

$$\begin{aligned}
&\frac{d\bar{\xi}_t(s,s')}{dt}=\rho\bar{g}(s,s')-\rho(1+\bar{g}(s,s'))\bar{\xi}_t(s,s')\\
&+s\,(\gamma+\alpha(s-1))\,\frac{\bar{x}_t(s-1)}{\bar{x}_t(s)}\left(\bar{\xi}_t(s-1,s')-\bar{\xi}_t(s,s')\right)\\
&+s'\,(\gamma+\alpha(s'-1))\,\frac{\bar{x}_t(s'-1)}{\bar{x}_t(s')}\left(\bar{\xi}_t(s'-1,s)-\bar{\xi}_t(s',s)\right)\\
&+\lambda\frac{\bar{x}_t(s+1)}{\bar{x}_t(s)}(N-s)\left(\bar{\xi}_t(s+1,s')-\bar{\xi}_t(s,s')\right)\\
&+\lambda\frac{\bar{x}_t(s'+1)}{\bar{x}_t(s')}(N-s')\left(\bar{\xi}_t(s'+1,s)-\bar{\xi}_t(s',s)\right)\\
&+\beta s\left(\bar{\xi}_t(s-1,s')-\bar{\xi}_t(s,s')\right)\frac{\bar{x}_t(s-1)}{\bar{x}_t(s)}\sum_{s''=1}^{N}\binom{N}{s''-1}\bar{\xi}_t(s-1,s'')\bar{x}_t(s'')\\
&+\beta s'\left(\bar{\xi}_t(s'-1,s)-\bar{\xi}_t(s',s)\right)\frac{\bar{x}_t(s'-1)}{\bar{x}_t(s')}\sum_{s''=1}^{N}\binom{N}{s''-1}\bar{\xi}_t(s'-1,s'')\bar{x}_t(s''). \qquad (111)
\end{aligned}$$

This can be further simplified to

$$\begin{aligned}
&\frac{d\bar{\xi}_t(s,s')}{dt} = \rho\bar{g}(s,s') - \rho(1+\bar{g}(s,s'))\bar{\xi}_t(s,s') \\
&+ \frac{\bar{x}_t(s-1)}{\bar{x}_t(s)}\left(\bar{\xi}_t(s-1,s') - \bar{\xi}_t(s,s')\right) s \left[(\gamma + \alpha(s-1)) + \beta \sum_{s''=1}^{N} \binom{N}{s''-1} \bar{\xi}_t(s-1,s'')\bar{x}_t(s'')\right] \\
&+ \frac{\bar{x}_t(s'-1)}{\bar{x}_t(s')}\left(\bar{\xi}_t(s'-1,s) - \bar{\xi}_t(s',s)\right) s' \left[(\gamma + \alpha(s'-1)) + \beta \sum_{s''=1}^{N} \binom{N}{s''-1} \bar{\xi}_t(s'-1,s'')\bar{x}_t(s'')\right] \\
&+ \lambda\frac{\bar{x}_t(s+1)}{\bar{x}_t(s)}(N-s)\left(\bar{\xi}_t(s+1,s') - \bar{\xi}_t(s,s')\right) + \lambda\frac{\bar{x}_t(s'+1)}{\bar{x}_t(s')}(N-s')\left(\bar{\xi}_t(s'+1,s) - \bar{\xi}_t(s',s)\right). \qquad (112)
\end{aligned}$$

Equations (23) and (112) provide a complete system of ODEs to describe the time evolution of $\bar{x}_t(s)$ and $\bar{\xi}_t(s,s')$. □

Proof of Lemma 2. When we start from the initial condition $h_{ik,0} = 0$ for all $i = 1,\dots,n$ and $k = 1,\dots,N$ at early times t quadratic terms in $O(y_i(t)y_j(t))$ are negligible, and we can analyze the ODE

$$\frac{dy_i(t)}{dt} = \gamma - (\lambda+\gamma)y_i(t) + \beta\sum_{j=1}^{n} a_{ij}y_j(t). \qquad (113)$$

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 (113) yields

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

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 $\mathbb{E}(|\mathsf{S}(\mathbf{h}_{it})||G) = N\mathbb{E}(h_{ik,t}|G) = Ny_i(t)$. □

Proof of Proposition 8. In the stationary state $\frac{dy_i(t)}{dt} = 0$ we obtain from Equation (50) that

$$0 = \gamma - (\lambda + \gamma)y_i + \beta \sum_{j=1}^{n} a_{ij} y_j - \beta \sum_{j=1}^{n} a_{ij} y_i y_j.$$

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 $\frac{\beta}{\lambda+\gamma} < \frac{1}{\mu_1}$, where μ_1 is the largest eigenvalue of $\mathbf{A}$, 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 [cf. Bonacich, 1987]

$$\mathbf{b}\left(G, \phi\right) = \sum_{k=1}^{n} \phi^k \mathbf{A}^k \mathbf{u},$$

for $\phi < 1/\mu_1$. The Boncaich centrality can also be written as

$$b_i\left(G, \phi\right) = 1 + \phi d_i + \phi^2 \sum_{j=1}^{n} a_{ij} d_j + O(\phi^3).$$

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

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

This gives us an upper bound on y_i given by

$$0 \leq y_i \leq 1 - \frac{\frac{\lambda}{\lambda+\gamma}}{1 + \frac{\beta}{\lambda+\gamma} d_i}.$$

□

Proof of Proposition 9. The proof follows from the fact that the graph $G \in \mathcal{G}(n, m)$ that maximizes the sum of Bonacich centralities is a nested split graph [cf. Belhaj et al., 2013; König et al., 2014]. □

Proof of Corollary 3. From Equation (51) and the symmetry implied by a k-regular graph we know that the

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

for $0 \leq k \leq n-1$, 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

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

□

Proof of Corollary 4. 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 (51) it then follows that

$$\begin{aligned} y_1 &= 1 - \frac{\frac{\lambda}{\lambda+\gamma}}{1 + \frac{\beta}{\lambda+\gamma}(n-1) y_2}, \\ y_2 &= 1 - \frac{\frac{\lambda}{\lambda+\gamma}}{1 + \frac{\beta}{\lambda+\gamma} y_1}. \end{aligned}$$

The solution is given by

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

□

Proof of Proposition 10. In order to determine welfare, we need to compute the following quantities from the endogenous variables $\bar{x}(s)$ and $\bar{\xi}(s, s')$, $s, s' \in \{0, \ldots, N\}$,

$$\begin{aligned} \bar{h} &= \sum_{s=0}^{N} s \binom{N}{s} \bar{x}(s) \\ \sigma_h^2 &= \sum_{s=0}^{N} (s - \bar{h})^2 \binom{N}{s} \bar{x}(s), \end{aligned}$$

In the case of $N = 2$ we have that

$$\begin{aligned} \bar{h} &= 2(\bar{x}(1) + \bar{x}(2)) \\ \sigma_h^2 &= \bar{h}^2(1 - 2\bar{x}(1) - \bar{x}(2)) + 2(1-\bar{h})^2\bar{x}(1) + (2-\bar{h})^2\bar{x}(2). \end{aligned}$$

From Proposition 7 we know that in the limit of $\eta \to \infty$ both cases $\theta = 1$ and $\theta = 0$ are identical, and in particular the number of links is the same, so that the welfare gain can be computed from

$$W_0(\mathbf{h}, G) - W_1(\mathbf{h}, G) = \frac{nb^2\sigma_h^2}{1 + b\bar{h}},$$

where we insert $\bar{x}(1)$ and $\bar{x}(2)$ as stated in Proposition 7 into the above expressions for $\bar{h}$ and σ_h^2 delivers Equation (54). The same observation can be made for the case of strong socks as stated in Proposition 7, where x_1 and x_2 can be obtained for the case of $\gamma = 0$. Inserting into the welfare gain gives Equation (53). $\square$