Network Formation with Local Complements and Global Substitutes: The Case of 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 analyze R&D collaboration networks in industries where firms are competitors in the product market. Firms' benefits from collaborations arise by sharing knowledge about a cost-reducing technology. By forming collaborations, however, firms also change their own competitive position in the market as well as the overall market structure. We analyze incentives of firms to form R&D collaborations with other firms and the implications of these alliance decisions for the overall network structure. We provide a general characterization of both equilibrium networks and endogenous production choices in the form of a Gibbs measure. We find that there exists a sharp transition from sparse to dense networks, and low and high output levels, respectively, with decreasing linking costs. Moreover, there exists an intermediate range of the linking cost for which multiple equilibria arise. The equilibrium selection is a path dependent process characterized by hysteresis. We also allow for firms to differ in their technological characteristics, investigate how this affects their propensity to collaborate and study the resulting network structure. We then analyze the efficient network maximizing social welfare, and find that the efficient graph is either empty, complete or shows a strong core periphery structure.

Key words: R&D networks, network formation, path dependence, efficiency
JEL: C63, D83, D85, L22

1. Introduction

R&D partnerships have become a widespread phenomenon characterizing technological dynamics, especially in industries with rapid technological development such as, for instance, the pharmaceutical, chemical and computer industries [see e.g. Hagedoorn, 2002; Powell et al., 2005; Roijakkers and Hagedoorn, 2006]. In these industries firms have become more specialized in specific domains of a technology and they tend to combine their knowledge with that of other firms that are specialized in different domains in order to jointly generate innovations that can help to reduce their production costs [Ahuja, 2000; Powell et al., 1996]. Despite the increasing importance of R&D collaborations there exists only limited research of theoretical models of these relationships which can be used for policy analysis. This paper provides the first fully tractable model of strategic R&D network formation with endogenous quantity choice, which takes into

☆I would like to thank Antonio Cabrales, Matt Jackson, Ben Golub, Yves Zenou, Fabrizio Zilibotti, Alexey Kushnir, Nick Netzer, Onur Özgür, Andrea Montanari, Sanjeev Goyal, Maria Sáez-Martí, Armin Schmutzler, Filomena Garcia, and seminar participants at the University of Zurich, the Public Economic Theory Conference in Lisbon in 2013 and Stanford University in 2012 for the helpful comments and advice. A previous version of this paper has been circulated under the title "Dynamic R&D Networks" as Working Paper No. 109 in the Department of Economics working paper series of the University of Zurich. The author acknowledges financial support from Swiss National Science Foundation through research grants PBEZP1–131169 and 100018_140266, and thanks SIEPR and the department of economics at Stanford University for their hospitality during 2010–2012.

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

account the two-way flow of influence from the market structure to the incentives to form R&D collaborations and, in turn, from the formation of collaborations to the market structure.

We study the incentives of firms to form R&D collaborations with other firms and the implications of these alliance decisions for the overall network structure. We provide a complete characterization of the stationary states of a dynamic process in which firms can adjust both, quantities produced (as well as research efforts), and the R&D collaborations between them, based on a noisy profit maximization rationale [cf. Blume, 2003; Brock and Durlauf, 2001]. Using a potential function we show that the stationary states of this process are completely characterized by a Gibbs measure [cf. Bisin et al., 2006; Grimmett, 2010]. Further, we find that for any distribution of output levels, the network can be characterized as an inhomogenous random graph with a link probability that depends on the firms' sizes and the linking cost [cf. Bollobás et al., 2007, 2001; Söderberg, 2002; Van Der Hofstad, 2009]. Moreover, we show that the stochastically stable networks (in the limit of vanishing noise) are "nested split graphs" [cf. König et al., 2013; König et al., 2011].[1] We then characterize the stationary degree distribution and compute the asymptotic network density. In particular, we find that there exists a sharp transition between sparse and dense networks with decreasing linking costs. We also compute the stationary output levels and show that there exists an intermediate range of the linking cost for which multiple equilibria arise. The equilibrium selection is a path dependent process characterized by hysteresis [cf. David, 1992, 2005]. Moreover, as in the case of the network density, there exists a sharp transition from a low output to a high output equilibrium. It is also possible to generalize our model by introducing heterogeneous marginal costs as well as heterogeneous spillovers from collaborations between firms stemming from differences in their technological characteristics. In particular, in the latter case we show that if firms' technology stocks follow a power-law distribution then the degree distribution will be power-law distributed as well [cf. Powell et al., 2005]. We then investigate the efficient network architecture and output structure that maximize social welfare, and find that the efficient graph is either empty, complete or shows a strong core periphery structure.

There exist a number of related works on R&D networks in the economics literature. Most notably, Dawid and Hellmann [2014]; Goyal and Joshi [2003]; Westbrock [2010] study the formation of R&D networks in which firms can form collaborations to reduce their production costs. In particular, Dawid and Hellmann [2014] study a perturbed best response dynamic process as we do here, and analyze the stochastically stable states. However differently to the current model, the cost reduction from a collaboration in these models is independent of the identity and the characteristics of the firms involved.[2] Our analysis also bears similarities with a number of other recent contributions in the literature which analyze a similar payoff structure. In the paper by Ballester et al. [2006] the authors derive equilibrium outcomes in a linear quadratic game where agents' efforts are local complements in an exogenously given network. Differently to Ballester et al. [2006], we make the network as well as effort choices endogenous.[3] Our ap-

[1] 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].

[2] Goyal and Moraga-Gonzalez [2001] present a more general setup which relaxes this assumption but their analysis is restricted to regular graphs and networks comprising of four firms. In this paper we take into account general equilibrium structures with an arbitrary number of firms and no ex ante restriction on the collaboration pattern between them.

[3] It is straightforward to see that the results obtained in this paper can be generalized to the payoff structure

proach is further a generalization of the endogenous network formation mechanisms proposed in Snijders [2001] and Mele [2010]. As in these papers we use a potential function to characterize the stationary states, but here both, the action choices as well as the linking decisions are fully endogenized. Similarly, Cabrales et al. [2010] allow the network to be formed endogenously, but assume that link strengths are proportional to effort levels, while we make the linking decision depending on marginal payoffs. Finally, in König et al. [2014] a similar market structure is considered, however, with an exogenous network, and the focus lies on developing optimal R&D subsidy strategies as well as characterizing key firms whose exit would have the largest impact on the output of the economy.

2. The Model

We consider a Cournot oligopoly game in which a set $\mathcal{N} = \{1, \dots, n\}$ of firms is competing in a homogeneous product market.[4] We assume that firms are not only competitors in the product market, but they can also form pairwise collaborative agreements. These pairwise links involve a commitment to share R&D results and thus lead to lower marginal cost of production of the collaborating firms. The amount of this cost reduction depends on the effort the firms invest into R&D. Given the collaboration network $G \in \mathcal{G}^n$, where $\mathcal{G}^n$ denotes the set of all graphs with n nodes, each firm sets an R&D effort level unilaterally.[5] We assume that firms can only jointly develop a cost reducing technology. Given the effort levels e_i, marginal cost c_i of firm i is given by[6,7]

$$c_i(\mathbf{e}, G) = \bar{c} - \alpha e_i - \beta \sum_{j=1}^{n} a_{ij} e_j, \tag{1}$$

where $a_{ij} = 1$ if firms i and j set up a collaboration (0 otherwise) and $a_{ii} = 0$. The parameter $\alpha \geq 0$ measures the relative cost reduction due to a firms' own R&D effort while the parameter $\beta \geq 0$ measures the relative cost reduction due to the R&D effort of its collaboration partners. In this model, firms are exposed to business stealing effects if their rivals increase their output via cost reducing R&D collaborations.[8]

Moreover, we also assume that firms incur a direct cost $\gamma \geq 0$ for their R&D efforts and a fixed cost $\zeta \geq 0$ for each R&D collaboration.[9] The profit of firm i, given the R&D network G and

introduced in Ballester et al. [2006]. See in particular the general payoff structure considered in Equation (8).

[4]Generalizations to Bertrand competition are straight forward [Westbrock, 2010].

[5]See also Kamien et al. [1992] for a similar model of competitive RJVs in which firms unilaterally choose their R&D effort levels.

[6]Note that we have neglected spillovers among non-collaborating firms.

[7]This generalizes earlier studies such as the one by D'Aspremont and Jacquemin [1988] where spillovers were assumed to take place between all firms in the industry and no distinction between collaborating and non-collaborating firms was made.

[8]In order to guarantee non-negative marginal costs we assume that $e_i \in [0, \bar{e}]$ and $\bar{c} \geq (n-1)\bar{e}$. This shows that $\bar{c}$ must be of the order of $O(n)$. Throughout the paper we shall assume parameter are constrained such that the second-order conditions hold and equilibria can be characterized in terms of first-order conditions and are interior.

[9]Observe that the direct cost ζ of collaboration is incurred by the firm initiating the collaboration. Therefore, it is the degree d_i of the firm i that appears in its profit function. We assume that R&D collaborations can only be formed if both firms agree to its establishment. One can show that marginal profit of a firm j to which firm i proposes a collaboration is given by $\pi_j(\mathbf{q}, G + ij) - \pi_j(\mathbf{q}, G) = \rho q_i q_j \geq 0$, with a constant $\rho \geq 0$. Since marginal profits are always non-negative, the firm j always accepts the proposed collaboration of i. This is a consequence of the assumption that firms are myopic and will be further discussed in Section 3.

the quantities $\mathbf{q}$ and efforts $\mathbf{e}$, is then given by

$$\pi_i(\mathbf{q}, \mathbf{e}, G) = (p_i - c_i)q_i - \gamma e_i^2 - \zeta d_i.$$

Inserting marginal cost from Equation (1) gives

$$\pi_i(\mathbf{q}, \mathbf{e}, G) = p_i q_i - \bar{c} q_i + \alpha q_i e_i + \beta q_i \sum_{j=1}^{n} a_{ij} e_j - \gamma e_i^2 - \zeta d_i.$$

The first-order condition with respect to R&D effort e_i is given by $\frac{\partial \pi_i(\mathbf{q},\mathbf{e},G)}{\partial e_i} = \alpha q_i - 2\gamma e_i = 0$. Solving for e_i and taking into account that $e_i \in [0, \bar{e}]$ delivers

$$e_i = \min\{\lambda q_i, \bar{e}\}, \tag{2}$$

where we have denoted by $\lambda = \frac{\alpha}{2\gamma}$.[10] Equation (2) can be viewed as reflecting learning-by-doing effects on R&D efforts. Various empirical studies have found that the R&D effort of a firm is proportional its output or size [Cohen and Klepper, 1996a,b]. We then can write marginal costs from Equation (1) as follows[11]

$$c_i(\mathbf{e}(\mathbf{q}), G) = \bar{c} - \lambda \alpha q_i - \lambda \beta \sum_{j=1}^{n} a_{ij} q_j. \tag{3}$$

Profits can be written as

$$\pi_i(\mathbf{q}, G) = p_i q_i - \bar{c} q_i - \lambda \alpha q_i^2 + \lambda \beta q_i \sum_{j=1}^{n} a_{ij} q_j - \lambda^2 \gamma q_i^2 - \zeta d_i. \tag{4}$$

Next we consider the demand for goods produced by firm i. A representative consumer maximizes [Singh and Vives, 1984]

$$U(I, q_1, \ldots, q_n) = I + a \sum_{i=1}^{n} q_i - \frac{1}{2} \sum_{i=1}^{n} q_i^2 - \frac{b}{2} \sum_{i=1}^{n} \sum_{j \neq i} q_i q_j, \tag{5}$$

with the budget constraint $I + \sum_{i=1}^{n} q_i \leq E$ and endowment E. The parameter a captures the total size of the market, whereas $b \in (0, 1]$, measures the degree of substitutability between products. In particular, $b = 1$ depicts a market of perfect substitutable goods, while $b \to 0$ represents the case of almost independent markets. The constraint is binding and the utility maximization of the representative consumer gives the inverse demand function for firm i

$$p_i = a - q_i - b \sum_{j \neq i} q_j. \tag{6}$$

[10] An interior solution hence requires that $q_i \leq \bar{q} \equiv \frac{2\gamma}{\alpha}\bar{e} \leq \frac{2\gamma\bar{c}}{\alpha(n-1)}$ for all $i = 1, \ldots, n$. Since $\bar{c}$ is $O(n)$ we thus require that q_i is $O(1)$. This means that quantities produced do not grow without bound as the number of firms in the industry becomes large.

[11] We assume that firms always implement the optimal R&D effort level. Since the optimal R&D effort decision only depends on a firm's own output, a firm does not face any uncertainty when implementing this strategy. In Section 3 we will, however, introduce noise in the optimal output and collaboration decisions, since these depend on the decisions of all other firms in the industry and their characteristics, which might be harder to observe.

Firms face an inverse linear demand as given in Equation (6). Firm i then sets its quantity q_i in order to maximize its profit π_i given by Equation (4). We also assume that there is a maximum production capacity $\bar{q}$ such that $q_i \leq \bar{q}$ for all $i \in \mathcal{N}$. Inserting marginal cost from Equation (3) and inverse demand from Equation (6) we can write firm i's profit as

$$\pi_i(\mathbf{q}, G) = (a - \bar{c})q_i - (1 - \lambda\alpha + \lambda^2\gamma)q_i^2 - bq_i \sum_{j \neq i} q_j + \lambda\beta \sum_{j=1}^{n} a_{ij} q_i q_j - \zeta d_i. \tag{7}$$

We assume that $a > \bar{c}$. Since $\bar{c}$ must be of the order of $O(n)$ this also implies that a is $O(n)$. In the following we will denote by $\eta = (a - \bar{c})/n$ (which is $O(1)$), $\nu = (1 - \lambda\alpha + \lambda^2\gamma)/n$ and $\rho = \lambda\beta$, so that Equation (7) becomes

$$\pi_i(\mathbf{q}, G) = \underbrace{n\eta q_i - n\nu q_i^2}_{\text{own concavity}} \quad \underbrace{-bq_i \sum_{j \neq i}^{n} q_j}_{\text{global substitutability}} \quad + \quad \underbrace{\rho q_i \sum_{j=1}^{n} a_{ij} q_j}_{\text{local complementarity}} \quad -\zeta d_i. \tag{8}$$

Firm $i \in \mathcal{N}$ sets its quantity q_i and makes profit π_i given by Equation (8). The corresponding first-order conditions are given by

$$\frac{\partial \pi_i(\mathbf{q}, G)}{\partial q_i} = n\eta - 2n\nu q_i - b \sum_{j \neq i}^{n} q_j + \rho \sum_{j=1}^{n} a_{ij} q_j = 0. \tag{9}$$

The second-order derivatives for $j \neq i$ are given by $\frac{\partial^2 \pi_i(\mathbf{q}, G)}{\partial q_j \partial q_i} = \frac{\partial^2 \pi_i}{\partial q_i \partial q_j} = -b + \rho a_{ij}$, which is positive if $b < \rho$ and $a_{ij} = 1$, and $\frac{\partial^2 \pi_i}{\partial q_i^2} = -2n\nu \leq 0$, if $\nu \geq 0$. Hence, the payoff function in Equation (8) is supermodular for linked firms expressing strategic complementarity, as allied firms' output choices are complements to each other [cf. Topkis, 1998]. From Equation (9) we can write firm i's best response quantity as

$$q_i = f_i(q_{-i}, G) \equiv \frac{\eta}{2\nu} - \frac{b}{2n\nu} \sum_{j \in \mathcal{N} \setminus \{i\}} q_j + \frac{\rho}{2n\nu} \sum_{j \in \mathcal{N}_i} q_j = \frac{\eta}{2\nu} - \frac{b}{2n\nu} \sum_{j \notin (N_i \cup \{i\})} q_j + \frac{\rho - b}{2n\nu} \sum_{j \in \mathcal{N}_i} q_j, \tag{10}$$

with the constraint that $0 \leq q_i \leq \bar{q}$. Equation (10) shows that output of i is decreasing in the output of the firms j not connected to i. Moreover, if $\rho < b$, then firm i's output is also decreasing in its neighbors' output. However, if $\rho > b$, i's output is increasing in its neighbors' output.

3. Equilibrium Characterization

In the following we provide a complete equilibrium analysis of the R&D collaboration game. The profit function introduced in Equation (8) admits a potential game with a corresponding potential function [cf. Monderer and Shapley, 1996], which not only accounts for quantity adjustments but also for the linking strategies.

Proposition 1. *Assume that both, quantities and links can be changed according to a myopic profit maximizing rationale of firms. Then the profit function of Equation (8) admits a potential game with potential function* $\Phi\colon \mathbb{R}_+^n \times \mathcal{G}_n \to \mathbb{R}$ *given by*

$$\Phi(\mathbf{q}, G) = \sum_{i=1}^{n} (n\eta q_i - \nu n q_i^2) - \frac{b}{2} \sum_{i=1}^{n} \sum_{j \neq i}^{n} q_i q_j + \frac{\rho}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} q_i q_j - \zeta m = \phi(\mathbf{q}, G) - \zeta m. \tag{11}$$

where m is the number of links in G.

We allow the network to be formed endogenously, based on the profit maximizing decisions of firms with whom to collaborate, and share knowledge about a cost reducing technology. The precise definition of the dynamics of quantity adjustment and network evolution is given in the following:

Definition 1. *The evolution of the population of firms is characterized by a sequence of states $(\boldsymbol{\omega}_t)_{t\in\mathbb{R}_+}$, $\omega_t \in \Omega$, where each state $\boldsymbol{\omega}_t = (\mathbf{q}_t, G_t)$ consists of a vector of firms' output levels $\mathbf{q}_t \in \mathcal{Q}^n$ and a network of collaborations $G_t \in \mathcal{G}^n$. We assume that firms choose quantities from an arbitrarily fine discretization $\mathcal{Q} = \{0, \Delta, 2\Delta, \ldots, \bar{q}\}$ of the interval $[0, \bar{q}]$ with $|\mathcal{Q}| = s$. In a short time interval $[t, t + \Delta t)$, $t \in \mathbb{R}_+$, one of the following events happens:*

Output adjustment *At rate $\chi s > 0$ a firm $i \in \mathcal{N}$ is selected at random and given a revision opportunity of its current output level q_{it}. When firm i receives such a revision opportunity, it draws a new output level q'_i from the set $\mathcal{Q}$ uniformly at random (with probability $1/s$) and evaluates its marginal payoffs from changing its output level from q_{it} to q'_i. The computation of marginal payoffs is perturbed by an additive i.i.d. shock ε_{it}, so that the probability that we observe a switch from output level q_{it} to q'_i is given by*

$$\begin{aligned}\mathbb{P}\left(\boldsymbol{\omega}_{t+\Delta t} = (q'_i, \mathbf{q}_{-it}, G_t) | \boldsymbol{\omega}_t = (q_{it}, \mathbf{q}_{-it}, G_t)\right) \\ = \nu \mathbb{P}\left(\pi_i(q'_i, \mathbf{q}_{-it}, G_t) - \pi_i(q_{it}, \mathbf{q}_{-it}, G_t) + \varepsilon_{it} > 0\right) \Delta t + o(\Delta t) \\ = \chi \mathbb{P}\left(\Phi(q'_i, \mathbf{q}_{-it}, G_t) - \Phi(q_{it}, \mathbf{q}_{-it}, G_t) + \varepsilon_{it} > 0\right) \Delta t + o(\Delta t).\end{aligned}$$

where ϑ is a scale parameter measuring the extent of noise relative to payoff maximization, and we have used the fact that $\pi_i(q'_i, \mathbf{q}_{-it}, G_t) - \pi_i(q_{it}, \mathbf{q}_{-it}, G_t) = \Phi(q'_i, \mathbf{q}_{-it}, G_t) - \Phi(q_{it}, \mathbf{q}_{-it}, G_t)$.

Link formation *With rate $\lambda > 0$ a pair of firms ij which is not already connected receives an opportunity to form a link. The formation of a link depends on the marginal payoff the firms receive from the link plus an additive pairwise i.i.d. error term $\varepsilon_{ij,t}$. The probability that link ij is created is then given by*

$$\begin{aligned}\mathbb{P}\left(\boldsymbol{\omega}_{t+\Delta t} = (\mathbf{q}_t, G_t + ij) | \boldsymbol{\omega}_{t-1} = (\mathbf{q}, G_t)\right) = \lambda\, \mathbb{P}\left(\{\pi_i(\mathbf{q}_t, G_t + ij) - \pi_i(\mathbf{q}_t, G_t) + \varepsilon_{ij,t} > 0\}\right. \\ \left.\cap \{\pi_j(\mathbf{q}_t, G_t + ij) - \pi_j(\mathbf{q}_t, G_t) + \varepsilon_{ij,t} > 0\}\right) \Delta t + o(\Delta t) \\ = \lambda\, \mathbb{P}\left(\Phi(\mathbf{q}_t, G_t + ij) - \Phi(\mathbf{q}_t, G_t) + \varepsilon_{ij,t} > 0\right) \Delta t + o(\Delta t),\end{aligned}$$

where we have used the fact that $\pi_i(\mathbf{q}_t, G_t + ij) - \pi_i(\mathbf{q}_t, G_t) = \pi_j(\mathbf{q}_t, G_t + ij) - \pi_j(\mathbf{q}_t, G_t) = \Phi(\mathbf{q}_t, G_t + ij) - \Phi(\mathbf{q}_t, G_t)$.

Link removal *With rate $\xi > 0$ a pair of connected firms ij receives an opportunity to terminate their connection. The link is removed if at least one firm finds this profitable. The marginal payoffs from removing the link ij are perturbed by an additive pairwise i.i.d. error term $\varepsilon_{ij,t}$. The probability that the link ij is removed is then given by*

$$\begin{aligned}\mathbb{P}\left(\boldsymbol{\omega}_{t+\Delta t} = (\mathbf{q}_t, G_t - ij) | \boldsymbol{\omega}_t = (\mathbf{q}, G_t)\right) = \xi\, \mathbb{P}\left(\{\pi_i(\mathbf{q}_t, G_t - ij) - \pi_i(\mathbf{q}_t, G_t) + \varepsilon_{ij,t} > 0\}\right. \\ \left.\cup \{\pi_j(\mathbf{q}_t, G_t - ij) - \pi_j(\mathbf{q}_t, G_t) + \varepsilon_{ij,t} > 0\}\right) \Delta t + o(\Delta t) \\ = \xi\, \mathbb{P}\left(\Phi(\mathbf{q}_t, G_t - ij) - \Phi(\mathbf{q}_t, G_t) + \varepsilon_{ij,t} > 0\right) \Delta t + o(\Delta t),\end{aligned}$$

where we have used the fact that $\pi_i(\mathbf{q}_t, G_t - ij) - \pi_i(\mathbf{q}_t, G_t) = \pi_j(\mathbf{q}_t, G_t - ij) - \pi_j(\mathbf{q}_t, G_t) = \Phi(\mathbf{q}_t, G_t - ij) - \Phi(\mathbf{q}_t, G_t)$.

Note that we can numerically implement 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]. We will use this method throughout the paper to illustrate our theoretical predictions for various network statistics.

In the following we make a specific assumption on the distribution of the random shocks. In particular, we assume that these shocks are independent and identically exponentially distributed with parameter $\vartheta \geq 0$. We then can write[12]

$$\begin{aligned}\mathbb{P}\left(\omega_{t+\Delta t} = (q_i', \mathbf{q}_{-it}, G_t) | \omega_t = (q_i, \mathbf{q}_{-it}, G_t)\right) &= \chi \, \mathbb{P}\left(-\varepsilon_{it} < \Phi(q_i', \mathbf{q}_{-it}, G_t) - \Phi(q_{it}, \mathbf{q}_{-it}, G_t)\right) \Delta t + o(\Delta t) \\ &= \chi \frac{e^{\vartheta \Phi(q_i', \mathbf{q}_{-i}, G_t)}}{e^{\vartheta \Phi(q_i', \mathbf{q}_{-it}, G_t)} + e^{\vartheta \Phi(q_i, \mathbf{q}_{-it}, G_t)}} \Delta t + o(\Delta t),\end{aligned}$$

and similarly we obtain for the creation of the link ij

$$\begin{aligned}\mathbb{P}\left(\omega_{t+\Delta t} = (\mathbf{q}_t, G_t + ij) | \omega_t = (\mathbf{q}_t, G_t)\right) &= \lambda \, \mathbb{P}\left(-\varepsilon_{ij,t} < \Phi(\mathbf{q}_t, G_t + ij) - \Phi(\mathbf{q}_t, G_t)\right) \Delta t + o(\Delta t) \\ &= \lambda \frac{e^{\vartheta \Phi(\mathbf{q}_t, G_t + ij)}}{e^{\vartheta \Phi(\mathbf{q}_t, G_t + ij)} + e^{\vartheta \Phi(\mathbf{q}_t, G_t)}} \Delta t + o(\Delta t), \qquad (12)\end{aligned}$$

and the removal of the link ij

$$\begin{aligned}\mathbb{P}\left(\omega_{t+\Delta t} = (\mathbf{q}_t, G_t - ij) | \omega_t = (\mathbf{q}_t, G_t)\right) &= \xi \, \mathbb{P}\left(-\varepsilon_{ij,t} < \Phi(\mathbf{q}_t, G_t - ij) - \Phi(\mathbf{q}_t, G_t)\right) \Delta t + o(\Delta t) \\ &= \xi \frac{e^{\vartheta \Phi(\mathbf{q}_t, G_t - ij)}}{e^{\vartheta \Phi(\mathbf{q}_t, G_t - ij)} + e^{\vartheta \Phi(\mathbf{q}_t, G_t)}} \Delta t + o(\Delta t).\end{aligned}$$

Let $\mathcal{F}$ denote the smallest σ-algebra generated by $\boldsymbol{\sigma}(\omega_t : t \in \mathbb{R}_+)$. The filtration is the non-decreasing family of sub-σ-fields $\{\mathcal{F}_t\}_{t \in \mathbb{R}_+}$ on the measure space $(\Omega, \mathcal{F})$, with the property that $\mathcal{F}_0 \subseteq \mathcal{F}_1 \subseteq \cdots \subseteq \mathcal{F}_t \subseteq \cdots \subseteq \mathcal{F}$. The probability space is given by the triple $(\Omega, \mathcal{F}, \mathbb{P})$, where $\mathbb{P} \colon \mathcal{F} \to [0,1]$ is the probability measure satisfying $\int_\Omega \mathbb{P}(\omega) d\mu(\omega) = 1$. The sequence of states $(\omega_t)_{t \in \mathbb{R}_+}$, $\omega_t \in \Omega$ induces an irreducible and aperiodic (i.e. ergodic) Markov chain. The one step transition probability $P : \Omega^2 \to [0,1]$ from a state $\omega \in \Omega$ to a state $\omega' \in \Omega$ is given by $\mathbb{P}(\omega_{t+\Delta t} = \omega' | \mathcal{F}_t = \sigma(\omega_0, \omega_1, \ldots, \omega_t = \omega)) = \mathbb{P}(\omega_{t+\Delta t} = \omega' | \omega_t = \omega) = p(\omega' | \omega)\Delta$, where $p(\omega'|\omega)$ is the transition rate from state ω to state ω'. Observe that in the continuous limit when $s \to \infty$ and $\mathbf{q}_t \in \mathbb{R}^n$ any function $f : \Omega \to \mathbb{R}$ of the state variables $\omega \in \Omega$ is a Carathéodory function since $f(\mathbf{q}, \cdot)$ is continuous for each $\mathbf{q} \in [0, \bar{q}]^n$ and $f(\cdot, G)$ is $(\mathcal{G}_n, \mathcal{B}_G)$ measurable [Aliprantis and Border, 2006].

In vector-matrix notation we can write $\Phi(\mathbf{q}, G) = \phi(\mathbf{q}, G) - \frac{\zeta}{2} \mathbf{u}^\top \mathbf{q}\mathbf{u}$. With the potential function $\Phi(\mathbf{q}, G)$ we then can state the following proposition.

Proposition 2. *The dynamic process $(\omega_t)_{t \in \mathbb{R}_+}$ induces an irreducible and aperiodic Markov chain with a unique stationary distribution $\mu^\vartheta \colon \mathcal{Q}^n \times \mathcal{G}^n \to [0,1]$ such that $\lim_{t \to \infty} \mathbb{P}(\omega_t = (\mathbf{q}, G) | \omega_0 = (\mathbf{q}_0, G_0)) = \mu^\vartheta(\mathbf{q}, G)$. The probability measure μ^ϑ is given by*

$$\mu^\vartheta(\mathbf{q}, G) = \frac{e^{\vartheta\left(\Phi(\mathbf{q}, G) - m \ln\left(\frac{\xi}{\lambda}\right)\right)}}{\sum_{G' \in \mathcal{G}^n} \sum_{\mathbf{q}' \in \mathcal{Q}^n} e^{\vartheta\left(\Phi(\mathbf{q}', G') - m' \ln\left(\frac{\xi}{\lambda}\right)\right)}}.$$

[12] Let z be i.i. logistically distributed with mean 0 and scale parameter ϑ, i.e. $F_z(x) = \frac{e^{\vartheta x}}{1+e^{\vartheta x}}$. Consider the random variable $\varepsilon = g(z) = -z$. Since g is monotonic decreasing, and z is a continuous random variable, the distribution of ε is given by $F_\varepsilon(y) = 1 - F_z(g^{-1}(y)) = \frac{e^{\vartheta y}}{1+e^{\vartheta y}}$.

In the limit of vanishing noise $\vartheta \to \infty$, the (stochastically stable) states in the support of μ^ϑ are given by [Kandori et al., 1993]

$$\lim_{\vartheta \to \infty} \mu^\vartheta(\mathbf{q}, G) \begin{cases} > 0, & \text{if } \Phi(\mathbf{q}, G) \geq \Phi(\mathbf{q}', G'), \ \ \forall \mathbf{q}' \in [0, \bar{q}]^n, \ \ G' \in \mathcal{G}_n, \\ = 0, & \text{otherwise.} \end{cases} \tag{13}$$

Note that we could also allow quantity adjustments of Definition 1 to follow a noisy directional learning process as in Anderson et al. [1998, 2002, 2004].[13] Quantity adjustments then follow a logit dynamics such that

$$\mathbb{P}\left(\boldsymbol{\omega}_{t+\Delta t} = (q_i', \mathbf{q}_{-it}, G_t) | \boldsymbol{\omega}_t = (q_i, \mathbf{q}_{-it}, G_t)\right) = \chi \frac{e^{\vartheta \pi_i (q_i', \mathbf{q}_{-it}, G_t)}}{\int_{[0,\bar{q}]^n} e^{\vartheta \pi_i (q, \mathbf{q}_{-it}, G_t)} dq^n} \Delta t.$$

However, this alternative definition would give rise to the same stationary distribution μ^ϑ as in Proposition 2 in the continuous limit when $s \to \infty$.

In the following we will set $\lambda = \xi$. The stationary distribution $\mu^\vartheta(\mathbf{q}, G)$ can then be further analyzed by computing the partition function [14]

$$\mathscr{Z}_\vartheta = \sum_{G \in \mathcal{G}_n} \sum_{q \in \mathcal{Q}^n} e^{\vartheta \Phi(\mathbf{q}, G)}, \tag{14}$$

so that we can write $\mu^\vartheta(\mathbf{q}, G) = e^{\vartheta \Phi(\mathbf{q}, G)} / \mathscr{Z}_\vartheta$. This allows us to compute the marginal distribution as stated in the following proposition.

Proposition 3. *The marginal distribution for the firms' output levels is given by*

$$\mu^\vartheta(\mathbf{q}) = \frac{1}{\mathscr{Z}_\vartheta} \sum_{G \in \mathcal{G}_n} e^{\vartheta \Phi(\mathbf{q}, G)} = \frac{1}{\mathscr{Z}_n^\vartheta} \prod_{i=1}^n e^{\vartheta\left(\eta n - \nu n q_i - \frac{b}{2} \sum_{j \neq i} q_j\right) q_i} \prod_{i<j}^n \left(1 + e^{\vartheta(\rho q_i q_j - \zeta)}\right).$$

Moreover, we can compute the probability of observing a network G given a specified output distribution $\mathbf{q}$.[15,16]

Proposition 4. *The probability of observing a network $G \in \mathcal{G}_n$, given an output distribution $\mathbf{q} \in \mathcal{Q}^n$ is determined by conditional distribution*

$$\mu^\vartheta(G|\mathbf{q}) = \prod_{i<j}^n \frac{e^{\vartheta a_{ij}(\rho q_i q_j - \zeta)}}{1 + e^{\vartheta(\rho q_i q_j - \zeta)}}, \tag{15}$$

[13] Mattsson and Weibull [2002] provide a motivation from boundedly rational choices with implementation costs.

[14] See Park and Newman [2004] for an excellent discussion in the context of exponential random graphs.

[15] For a discussion of inhomogeneous random graphs see Bollobás et al. [2001]; Van Der Hofstad [2009] and the "hidden variables" model studied in Boguñá and Pastor-Satorras [2003]. Observe that the complementary problem of determining the distribution of random variables that depend only on their neighbors in a given network G is associated with a *Markov random field*, whose distribution is a *Gibbs measure* (by the Hammersley-Clifford theorem), and can be decomposed into a sum over all cliques in G (see Besag [1974] and Kolaczyk [2009, Chap. 8] as well as Rue and Held [2005]).

[16] Proposition 4 has important implications. Numerous empirical studies have shown that the distribution of output levels among firms tends to follow a power-law distribution (Zipf's law) [Axtell, 2001; Gabaix, 1999; Growiec et al., 2008; Stanley et al., 1996]. If the output levels q_i and q_j in the link probability of Equation (16) are distributed according to a power-law, then we obtain the so called *fitness model* analyzed in Boguñá and Pastor-Satorras [2003]; Caldarelli et al. [2002]. This models also refer to random threshold graphs [Diaconis et al., 2008; Ide and Konno, 2007; Ide et al., 2010].

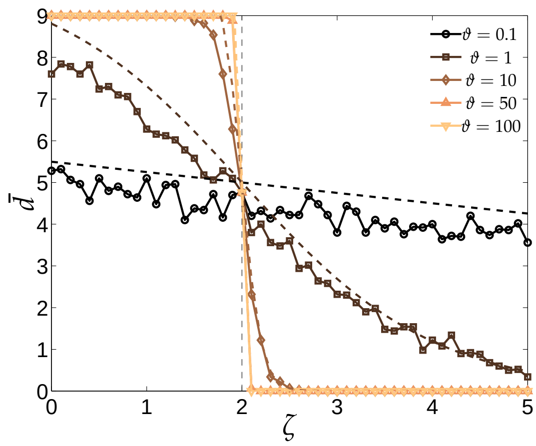

Figure 1: The average degree $\bar{d}$ as a function of the linking cost ζ for a fixed, homogeneous output distribution $q_i = 1$ for all $i = 1, \ldots, n$ with $\chi = 0$, $b = 0$, $\rho = 2$ and $n = 10$. The critical linking cost is $\zeta^* = \rho q_0^2 = 2$ is indicated with a vertical dashed line. Dashed lines indicate the theoretical prediction of Proposition 7.

which is equivalent to the probability of observing an inhomogeneous random graph with link probability

$$p^{\vartheta}(q_i, q_j) = \frac{e^{\vartheta(\rho q_i q_j - \zeta)}}{1 + e^{\vartheta(\rho q_i q_j - \zeta)}}. \tag{16}$$

A special case is one in which all firms produce at fixed output levels $q_i = q_0 \leq \bar{q}$ for all $i = 1, \ldots, n$ (letting $\chi \to 0$)[17] and there are no substitutability effects, $b = 0$. One can then show that the stochastically stable network (letting $\vartheta \to \infty$) is a nested split graph, which is characterized by the fact that the neighborhood of every node is contained in the neighborhoods of the nodes with higher degrees [cf. König et al., 2013; Mahadev and Peled, 1995]:

Proposition 5. *Consider the case where there are no substitutability effects, that is, setting $b = 0$, and assume that there are no output adjustments, i.e. $\chi = 0$.*

(i) If firms produce at the fixed output levels $q_i \in \mathcal{Q}$ for all $i = 1, \ldots, n$, then the stochastically stable network is given by a nested split graph with adjacency matrix $\mathbf{A} = (a_{ij})_{1 \leq i,j \leq n}$ whose a_{ij} elements are given by

$$a_{ij} = \begin{cases} 1, & \text{if } \rho q_i q_j > \zeta, \\ 0, & \text{if } \rho q_i q_j < \zeta. \end{cases}$$

(ii) If all firms produce at the same output level given by $q_i = q_0$ with $q_0 \in \mathcal{Q}$ for all $i = 1, \ldots, n$, then the stochastically stable network is given by the complete graph K_n if $\rho q_0^2 > \zeta$ and it is given by the empty graph $\overline{K}_n$ if $\rho q_0^2 < \zeta$.

The transition from the empty to the complete graph that occurs at $\zeta^* = \rho q_0^2$ in part (ii) in Proposition 5 is shown in Figure 1 for different values of ϑ and $n = 10$ nodes. Note that Proposition 5 can be generalized to endogenous quantity levels (when $\chi > 0$), as we can always condition any outcome on a distribution of quantities $\mathbf{q}$, and then sum over all $\mathbf{q} \in \mathcal{Q}^n$ weighted with the marginal probability measure $\mu^{\vartheta}(\mathbf{q})$. Since for any such $\mathbf{q}$ the stochastically stable network will be a nested split graph, the stochastically stable network will be a nested split graph with probability one. Further note that nested split graphs are paramount examples of core-periphery

[17] Appendix A provides an equilibrium characterization in the case of an exogenously given network. This corresponds to setting $\lambda = \xi = 0$ in Definition 1.

networks. The core-periphery structure of R&D alliance networks has also been documented empirically in Kitsak et al. [2010] and Rosenkopf and Schilling [2007]. Our model thus provides a theoretical explanation for why real-world R&D networks exhibit such a core-periphery structure.

We next compute the marginal and conditional asymptotic probabilities. Using a Laplace expansion around the equilibrium values $\mathbf{q}^*$ (i.e. the potential maximizers) we can write for large ϑ [cf. Wong, 2001]

$$\mu^{\vartheta}(G) = \frac{1}{\mathscr{Z}_{\vartheta}} \sum_{\mathbf{q} \in \mathcal{Q}^n} e^{\vartheta \Phi(G,\mathbf{q})} \approx \frac{1}{\mathscr{Z}_{\vartheta}} \left(\frac{\vartheta}{2\pi}\right)^{\frac{n}{2}} \left| \left(\frac{\partial^2 \Phi(G,\mathbf{q})}{\partial q_i \partial q_j}\right)_{\mathbf{q}=\mathbf{q}^*} \right|^{-\frac{1}{2}} e^{\vartheta \Phi(G,\mathbf{q}^*)},$$

and the conditional distribution is given by

$$\mu^{\vartheta}(\mathbf{q}|G) = \frac{\mu^{\vartheta}(G,\mathbf{q})}{\mu^{\vartheta}(G)} \approx \left(\frac{\vartheta}{2\pi}\right)^{\frac{n}{2}} \left| \left(\frac{\partial^2 \Phi(G,\mathbf{q})}{\partial q_i \partial q_j}\right)_{\mathbf{q}=\mathbf{q}^*} \right|^{-\frac{1}{2}} e^{\vartheta(\Phi(G,\mathbf{q}) - \Phi(G,\mathbf{q}^*))}.$$

The above expressions allow us to compute the stationary output levels in the limit of vanishing noise.

Proposition 6. *For large ϑ (in the stochastically stable equilibrium), we have that the stationary output levels are the roots of the equation*

$$(b+2\nu)q - \eta = \frac{\rho}{2}\left(1 + \tanh\left(\frac{\vartheta}{2}\left(\rho q^2 - \zeta\right)\right)\right) q, \tag{17}$$

from which it follows that for $\vartheta \to \infty$

$$q = \begin{cases} \frac{\eta}{b+2\nu-\rho}, & \text{if } \zeta < \frac{\rho\eta^2}{(b+2\nu)^2}, \\ \left\{\frac{\eta}{b+2\nu-\rho}, \frac{\eta}{\rho}\right\}, & \text{if } \frac{\rho\eta^2}{(b+2\nu)^2} < \zeta < \frac{\rho\eta^2}{(b+2\nu-\rho)^2}, \\ \frac{\eta}{b+2\nu}, & \text{if } \frac{\rho\eta^2}{(b+2\nu-\rho)^2} < \zeta. \end{cases} \tag{18}$$

Note that the stationary output levels in Proposition 6 are increasing in ρ and η, and decreasing in ζ and b (cf. Figure 2).

An illustration with the average output level from numerical simulations starting with different initial conditions can be seen in Figure 3. The next proposition determines the expected number of links in the limit of large n.

Proposition 7. *In the limit of large n, the expected number of links is given by*

$$\mathbb{E}^{\vartheta}(m) = \frac{n(n-1)}{2}\left(1 + \tanh\left(\frac{\vartheta}{2}\left(\rho q^2 - \zeta\right)\right)\right) + O(n),$$

where q derives from Equation (17) in the proof of Proposition 6 in Appendix C.

Proposition 7 shows that the expected number of links is increasing in ρ, q and η, and decreasing in ζ and b (by reducing the equilibrium quantity q). The left panel in Figure 4 shows the convergence of the average degree to its stationary value from Proposition 7 for varying values of ϑ.

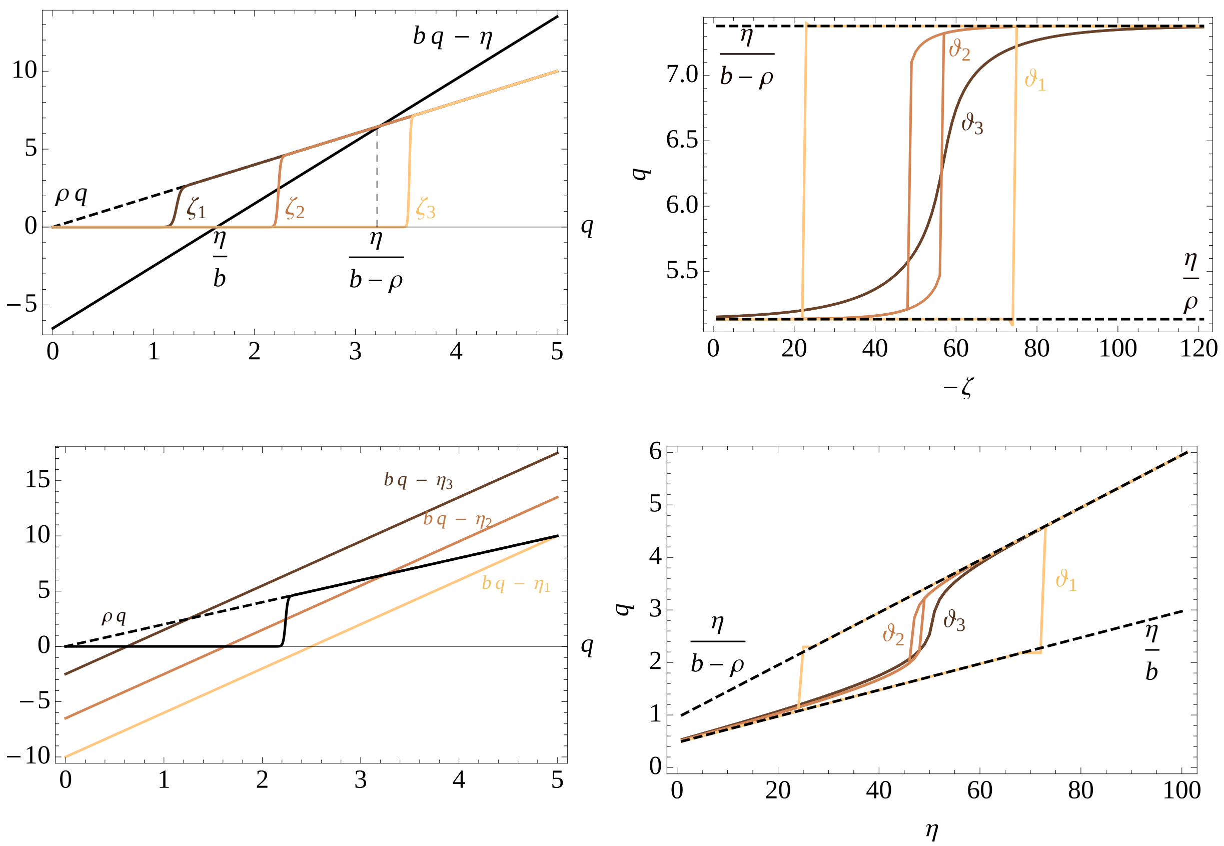

Figure 2: (Top left panel) The right hand side of Equation (17) for different values of $\zeta_1 = 25$, $\zeta_2 = 10$, $\zeta_3 = 3$ and $b = 4$, $\rho = 2$, $\eta = 6.5$, $\nu = 0$ and $\vartheta = 10$. (Top right panel) The values of q solving Equation (17) for different values of ζ with $b = 1.48$, $\rho = 0.45$ and $\vartheta_1 = 49.5$, $\vartheta_2 = 0.495$, $\vartheta_3 = 0.2475$. (Bottom left panel) The right hand side of Equation (17) for different values of $\eta_1 = 2.5$, $\eta_2 = 6.5$, $\eta_3 = 10$ and $b = 4$, $\rho = 2$, $\zeta = 10$ and $\vartheta = 10$. (Bottom right panel) The values of q solving Equation (17) for different values of η with $b = 4$, $\rho = 2$ and $\vartheta_1 = 10$, $\vartheta_2 = 0.26$, $\vartheta_3 = 0.2$

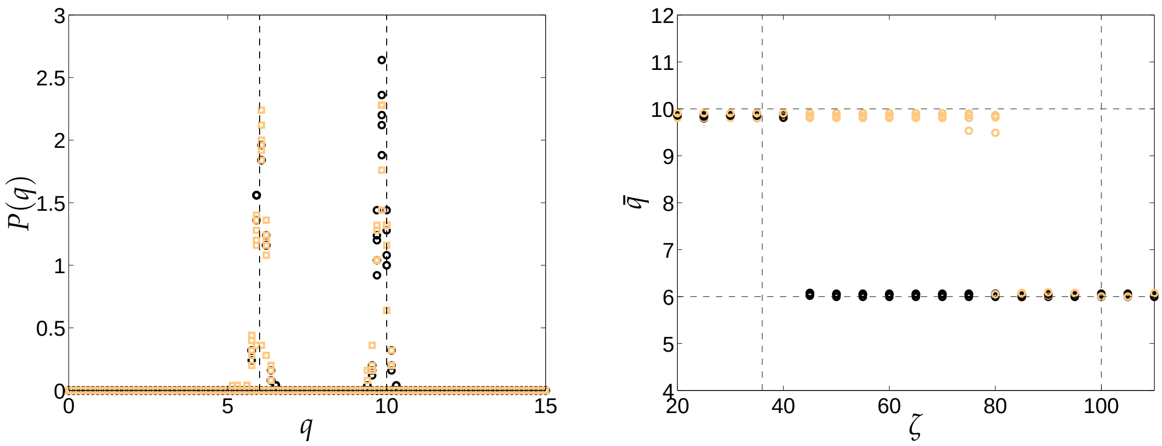

Figure 3: (Left panel) The stationary output distribution. The vertical dashed lines indicate the theoretical predictions from Equation (18). (Right panel) The average output level from numerical simulations with $\vartheta = 1$ starting with different initial conditions (indicated with different colors). The horizontal dashed lines indicate the equilibrium quantities and the vertical dashed lines the threshold cost levels from Equation (18). In the region of the cost ζ between the lower and upper thresholds two equilibria exist.

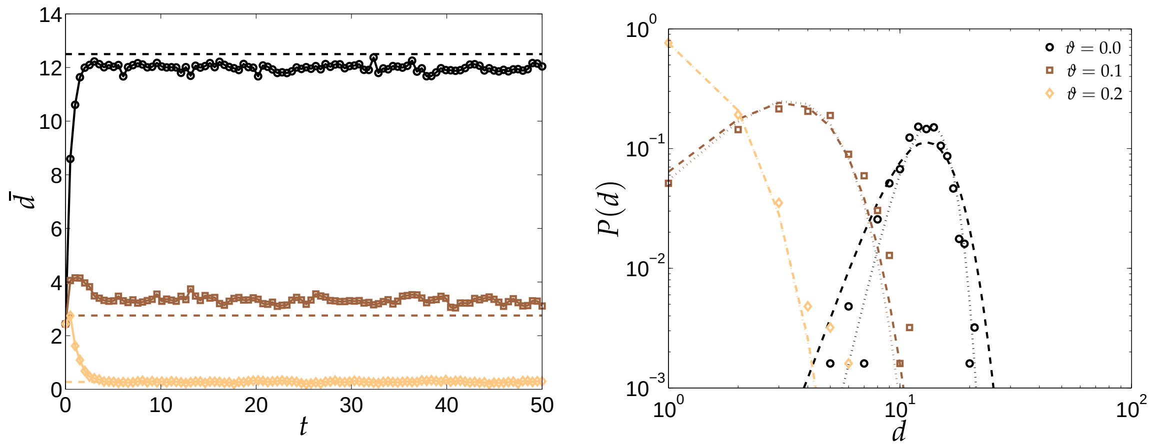

Figure 4: (Left panel) The time evolution of the average degree $\bar{d}$ with the dashed horizontal lines indicating the stationary solution from Proposition 7 for varying values of $\vartheta \in \{0, 0.1, 0.2\}$ with $\lambda = \xi = \chi = 1$, $\eta = 15$, $\nu = 1$ $b = 0.5$ $\rho = 1$ and $n = 25$ firms. (Right panel) The stationary degree distribution $P(k)$ for the same parameter values. The dashed lines indicate the solution from Proposition 8. The dashed lines indicate a Poisson, the dotted lines a binomial distribution.

With the partition function $\mathscr{Z}_\theta$ in Equation (52) in the proof of Proposition 6 in Appendix C we are able to compute the marginal distribution

$$\mu^\vartheta(\mathbf{q}) = \sum_{G \in \mathcal{G}_n} \mu^\vartheta(G, \mathbf{q}) = \frac{1}{\mathscr{Z}_\theta} \prod_{i=1}^n e^{\vartheta\left(n\eta - n\nu q_i - \frac{b}{2}\sum_{j \neq i} q_j\right) q_i} \prod_{i<j}^n \left(1 + e^{\vartheta(\rho q_i q_j - \zeta)}\right),$$

and the joint distribution can then be written as $\mu^\vartheta(G, \mathbf{q}) = \mu^\vartheta(G|\mathbf{q})\mu^\vartheta(\mathbf{q})$, where the conditional distribution $\mu^\vartheta(G|\mathbf{q})$ is given in Equation (15). The marginal distribution is given by

$$\mu^\vartheta(q) = \sum_{q_1 \in \mathcal{Q}} \sum_{q_2 \in \mathcal{Q}} \cdots \sum_{q_{i-1} \in \mathcal{Q}} \sum_{q_{i+1} \in \mathcal{Q}} \cdots \sum_{q_n \in \mathcal{Q}} \mu^\vartheta(\mathbf{q}), \tag{19}$$

which is independent of i due to symmetry. The next proposition characterizes the degree distribution.

Proposition 8. *In the limit of large linking costs c, the degree distribution is Poissonian with*

$$P(k) = \frac{1}{k!} \sum_{q \in \mathcal{Q}^n} \mu^\vartheta(q) e^{-\bar{d}(q)} \bar{d}(q)^k,$$

where the average degree of a firm with output q is given by $\bar{d}(q) = n \sum_{q' \in \mathcal{Q}^n} \mu^\vartheta(q') p(q, q')$, while $\mu^\vartheta(q)$ is given by Equation (19) and $p^\vartheta(q, q')$ is given by Equation (16).

When the distribution μ^ϑ is concentrated on the output level q^* (cf. Proposition 6) then one can show that the degree distribution is binomial with $P(k) = \binom{n}{k} p^\vartheta(q^*, q^*)^k (1 - p^\vartheta(q^*, q^*))^{n-k}$, where $p^\vartheta(q^*, q^*) = \frac{1}{1+e^{-\vartheta(\rho(q^*)^2 - \zeta)}}$. The right panel in Figure 4 shows the degree distribution for varying values of ϑ.

4. Extensions

The model presented so far can be extended in a number of directions which are described in Appendix B. First, in Appendix B.1 we allow firms to differ in their technologies, which in turn affect the spillovers generated from collaborations [cf. Cohen and Levinthal, 1990; Griffith et al.,

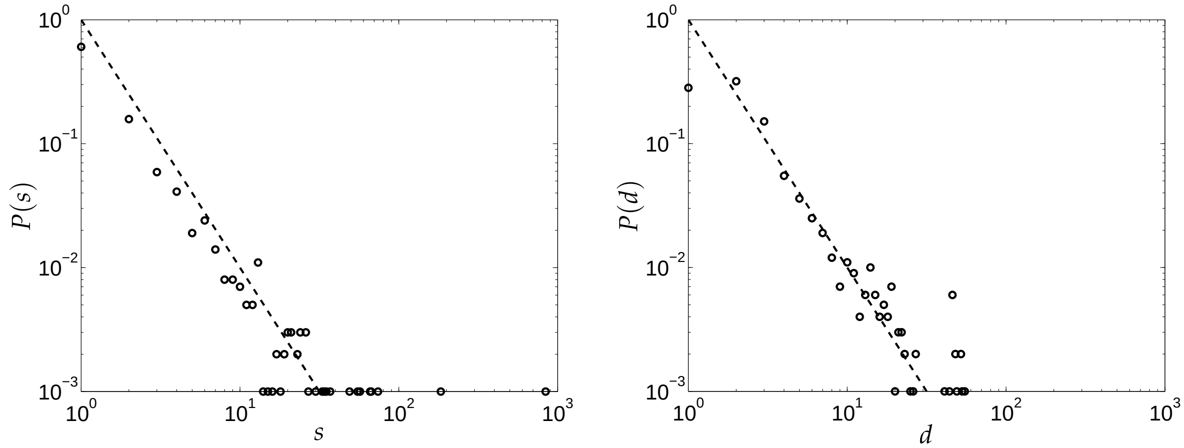

Figure 5: (Left panel) The distribution of the stocks of knowledge following a power-law with exponent $\gamma = 2$ across $n = 100$ firms. (Right panel) The stationary degree distribution with power-law distributed knowledge stocks with coefficient $\gamma = 2$ for $n = 100$, $\eta = 15$, $b = 0.5$, $\nu = 1$, and $\rho = 1$. The dashed line indicates a power-law with the same exponent.

2003]. One can show that a similar equilibrium characterization using a Gibbs measure as in the previous section is possible. Moreover, in the special case of firms' technology stocks being power-law distributed, and assuming that the spillovers from collaborations exhibit technological complementarities, one can show that the degree distribution also follows a power-law, confirming previous empirical studies of R&D networks [e.g. Gay and Dousset, 2005; Powell et al., 2005]. The result is stated in the following proposition.[18]

Proposition 9. *Assume that the spillovers from collaboration between firms i and j are proportional to their knowledge stocks, s_i, s_j, and further assume that the knowledge stocks s are distributed as a power-law $P(s) \sim s^{-\gamma}$ with exponent γ. Then the asymptotic degree distribution is also power-law distributed, $P(k) \sim k^{-\frac{\gamma}{\gamma-1}}$, with exponent $\frac{\gamma}{\gamma-1}$.*

An example can be seen in Figure 5 in the case of the stocks of knowledge following a power-law $P(s) \sim s^{-\gamma}$ with exponent $\gamma = 2$, and the degree distribution being power-law distributed $P(k) \sim k^{-\frac{\gamma}{\gamma-1}}$ with the same exponent $\frac{\gamma}{\gamma-1} = 2$.

A second extension outlined in Appendix B.2 considers ex ante heterogeneity among firms in the variable cost $\bar{c}_i \geq 0$ for $i = 1, \ldots, n$ [see also Banerjee and Duflo, 2005], expressing their different technological and organizational capabilities. Similarly to above, we can characterize the equilibrium states using a Gibbs measure. Moreover, the equilibrium networks are nested split graphs in which firms with lower marginal costs are more central.

5. Efficiency

For a given network G, social welfare $W(G)$ is given by the sum of consumer surplus and firms' profits. When firms compete in a homogeneous product oligopoly then social welfare is given

[18] For the proof see Appendix B.1.

G^*	q_1^*	q_2^*
empty graph	$q_1^* = q_2^* = 0$	$q_1^* = q_2^* = 0$
dominant group	$q_1^* > 0$	$q_2^* = 0$
dominant group	$q_1^* > 0$	$q_2^* > 0$
complete graph	$q_1^* = q_2^* > 0$	$q_1^* = q_2^* > 0$

Table 1: Summary of efficient networks and quantities. The optimal quantities $\mathbf{q}^* = (q_1^*, \ldots, q_2^*, \ldots)^\top$ are given by Equation (55) subject to $0 \leq q_1^*, q_2^* \leq \bar{q}$, where the optimal size n_1 of the dominant group maximizes Equation (56) and $n_2 = n - n_1$ in the proof of the proposition in Appendix C.

by[19]

$$\begin{aligned} W(\mathbf{q}, G) &= \frac{1}{2}\left(\sum_{i=1}^{n} q_i\right)^2 + \sum_{i=1}^{n} \pi_i(\mathbf{q}, G) \\ &= \frac{1}{2}\left(\sum_{i=1}^{n} q_i\right)^2 + \sum_{i=1}^{n}\left(\eta n q_i - \nu n q_i^2 - b\sum_{j\neq i} q_i q_j + \rho q_i \sum_{j=1}^{n} a_{ij} q_i q_j\right) - 2\zeta m. \end{aligned} \tag{20}$$

Note that welfare $W(\mathbf{q}, G)$ is related to the potential $\Phi(\mathbf{q}, G)$ as follows

$$W(\mathbf{q}, G) = \frac{1}{2}\left(\sum_{i=1}^{n} q_i\right)^2 - \sum_{i=1}^{n} q_i n(\eta - \nu q_i) + 2\Phi(\mathbf{q}, G). \tag{21}$$

Hence, the states maximizing the potential $\Phi(\mathbf{q}, G)$ are not necessarily identical to the ones maximizing welfare $W(\mathbf{q}, G)$. The latter are determined in the following proposition.[20]

Proposition 10. *The efficient network G^* maximizing welfare is either (i) the empty network, (ii) the complete network, or (iii) has the dominant group architecture. The optimal quantities $\mathbf{q}^* = (q_1^*, \ldots, q_2^*, \ldots)$ are given by Equation (55) in Appendix C subject to $0 \leq q_1^*, q_2^* \leq \bar{q}$, where the optimal size n_1 of the dominant group maximizes Equation (56) in Appendix C and $n_2 = n - n_1$.*

The efficient graphs and quantities are also summarized in Table 1. As we might expect, with increasing cost, the efficient network becomes more sparse.

6. Future Work

Three important avenues are left for future work. First, it would be interesting to study entry and exit dynamics in the current framework. It has been argued that entry and exit play an important role in shaping the distribution of firm sizes [Acemoglu and Cao, 2010; Luttmer, 2007].[21] A promising approach seems to be a union of the model proposed in this paper and

[19] In the empirical paper by König et al. [2014] this welfare analysis is extended to account for R&D subsidies. Moreover, the authors characterize the firms that are most critical in terms of their contribution to the aggregate productivity of the economy.

[20] Proposition 10 characterizes the the efficient outcome in the first best solution where the social planner can set both, the production levels as well as the network of collaborations between them. A characterization of the second best solution, in which the planner chooses the network, but output levels are chosen in a decentralized manner by profit maximizing firms is studied in König et al. [2014]. Moreover, an equilibrium characterization with an exogenously given network is provided in Appendix A.

[21] As Boguñá and Pastor-Satorras [2003] show, entry and exit can be incorporated by introducing a time dependent firm characteristic (more precisely, inversely related to time) that impacts the connection probability as in Appendix B.1.

the one in Garlaschelli et al. [2007]. Second, it would be interesting to analyze the dynamics of technological change and convergence and their relation with firm and network dynamics in the current model. Such an extension could shed light on the coevolution of R&D networks and the knowledge portfolios of firms [cf. König, 2011]. Finally, an empirical application of the model to real-world R&D networks could help to shed light on the often significant differences between sectors and, in particular, why the biotech sector has witnessed a steady increase in the number of collaborations while other sectors have experienced a less sustained development [cf. Schilling, 2009].

References

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

Ahuja, G. (2000). Collaboration networks, structural holes, and innovation: A longitudinal study. *Administrative Science Quarterly*, 45:425–455.

Aliprantis, C. and Border, K. (2006). *Infinite dimensional analysis: a hitchhiker's guide*. Springer Verlag.

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, S., Goeree, J., and Holt, C. (1998). Rent seeking with bounded rationality: An analysis of the all-pay auction. *Journal of Political Economy*, 106(4):828–853.

Anderson, S., Goeree, J., and Holt, C. (2002). The logit equilibrium: A perspective on intuitive behavioral anomalies. *Southern Economic Journal*, pages 21–47.

Anderson, S., Goeree, J., and Holt, C. (2004). Noisy directional learning and the logit equilibrium. *Scandinavian Journal of Economics*, 106(3):581–602.

Axtell, R. (2001). Zipf distribution of US firm sizes. *Science*, 293(5536):1818.

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

Banerjee, A. and Duflo, E. (2005). Growth theory through the lens of development economics. *Handbook of Economic growth*, 1:473–552.

Baum, J., Cowan, R., and Jonard, N. (2009). Network-independent partner selection and the evolution of innovation networks. *UNU-MERIT working paper series; 2009-022*.

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.

Besag, J. (1974). Spatial interaction and the statistical analysis of lattice systems. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 192–236.

Bisin, A., Horst, U., and Özgür, O. (2006). Rational expectations equilibria of economies with local interactions. *Journal of Economic Theory*, 127(1):74–116.

Bloom, N., Schankerman, M., and Van Reenen, J. (2007). Identifying technology spillovers and product market rivalry. Technical report, NBER Working Paper No. 13060.

Blume, L. (2003). How noise matters. *Games and Economic Behavior*, 44(2):251–271.

Blundell, R., Griffith, R., and Van Reenen, J. (1995). Dynamic count data models of technological innovation. *The Economic Journal*, 105(429):333–344.

Boguñá, M. and Pastor-Satorras, R. (2003). Class of correlated random networks with hidden variables. *Physical Review E*, 68(3):036112.

Boguñá, M., Pastor-Satorras, R., Diaz-Guilera, A., and Arenas, A. (2004). Models of social networks based on social distance attachment. *Physical Review E*, (70):056122.

Bollobás, B., Janson, S., and Riordan, O. (2007). The phase transition in inhomogeneous random graphs. *Random Structures & Algorithms*, 31(1):3–122.

Bollobás, B., Riordan, O., Spencer, J., and Tusnády, G. (2001). The degree sequence of a scale-free random graph process. *Random Structures and Algorithms*, 18(3):279–290.

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

Boyd, S. and Vandenberghe, L. (2004). *Convex Optimization*. Cambridge University Press.

Bramoullé, Y., Kranton, R., and D'Amours, M. (2010). Strategic interaction and networks. Cahiers de recherche 10–18, CIRPEE.

Brock, W. and Durlauf, S. (2001). Discrete choice with social interactions. *Review of Economic Studies*, 68(2):235–260.

Cabrales, A., Calvó-Armengol, A., and Zenou, Y. (2010). Social interactions and spillovers. *Games and Economic Behavior*.

Caldarelli, G., Capocci, A., De Los Rios, P., and Muñoz, M. A. (2002). Scale-free networks from varying vertex intrinsic fitness. *Physical Review Letters*, 89(25):258702.

Calvó-Armengol, A., Patacchini, E., and Zenou, Y. (2009). Peer effects and social networks in education. *Review of Economic Studies*, 76:1239–1267.

Cohen, W. and Klepper, S. (1996a). A reprise of size and R & D. *The Economic Journal*, 106(437):925–951.

Cohen, W. and Klepper, S. (1996b). Firm size and the nature of innovation within industries: the case of process and product R&D. *The Review of Economics and Statistics*, 78(2):232–243.

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

Corchon, L. C. and Mas-Colell, A. (1996). On the stability of best reply and gradient systems with applications to imperfectly competitive models. *Economics Letters*, 51(1):59 – 65.
D'Aspremont, C. and Jacquemin, A. (1988). Cooperative and noncooperative R&D in duopoly with spillovers. *The American Economic Review*, 78(5):1133–1137.
David, P. A. (1992). Heroes, herds and hysteresis in technological history: Thomas edison and the battle of the systems reconsidered. *Industrial and Corporate Change*, 1(1):129–180.
David, P. A. (2005). Path dependence in economic processes: implications for policy analysis in dynamical systems contexts. *The evolutionary foundations of economics*, pages 151–194.
Dawid, H. and Hellmann, T. (2014). The evolution of R&D networks. *Journal of Economic Behavior & Organization*, 105:158–172.
Diaconis, P., Holmes, S., and Janson, S. (2008). Threshold graph limits and random threshold graphs. *Internet Mathematics*, 5(3):267–320.
Gabaix, X. (1999). Zipf's Law For Cities: An Explanation. *Quarterly Journal of Economics*, 114(3):739–767.
Gabaix, X. (2009). Power laws in economics and finance. *Annual Review of Economics*, 1(1):255–294.
Garlaschelli, D., Capocci, A., and Caldarelli, G. (2007). Self-organised network evolution coupled to extremal dynamics. *Nature Physics*, 3:813–817.
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.
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.
Goyal, S. and Joshi, S. (2003). Networks of collaboration in oligopoly. *Games and Economic Behavior*, 43(1):57–85.
Goyal, S. and Moraga-Gonzalez, J. L. (2001). R&D networks. *RAND Journal of Economics*, 32(4):686–707.
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.
Grimmett, G. (2010). *Probability on graphs*. Cambridge University Press.
Growiec, J., Pammolli, F., Riccaboni, M., and Stanley, H. (2008). On the size distribution of business firms. *Economics Letters*, 98(2):207–212.
Hagedoorn, J. (2002). Inter-firm R&D partnerships: an overview of major trends and patterns since 1960. *Research Policy*, 31(4):477–492.
Horn, R. A. and Johnson, C. R. (1990). *Matrix Analysis*. Cambridge University Press.
Ide, Y. and Konno, N. (2007). Limit theorems for some statistics of a generalized threshold network model. *Theory of Biomathematics and its Applications III*, 1551:81–86.
Ide, Y., Konno, N., and Masuda, N. (2010). Statistical properties of a generalized threshold network model. *Methodology and Computing in Applied Probability*, 12(3):361–377.
Ide, Y., Konno, N., and Obata, N. (2009). Spectral properties of the threshold network model. *Internet Mathematics*, 6(2):173–187.
Jaffe, A. B. (1989). Characterizing the technological position of firms, with application to quantifying technological opportunity and research spillovers. *Research Policy*, 18(2):87 – 97.
Kamien, M. I., Muller, E., and Zang, I. (1992). Research joint ventures and R&D cartels. *The American Economic Review*, 82(5):1293–1306.
Kandori, M., Mailath, G., and Rob, R. (1993). Learning, mutation, and long run equilibria in games. *Econometrica*, pages 29–56.
Kitsak, M., Riccaboni, M., Havlin, S., Pammolli, F., and Stanley, H. (2010). Scale-free models for the structure of business firm networks. *Physical Review E*, 81:036117.
Kolaczyk, E. (2009). *Statistical Analysis of Network Data: Methods and Models*. Springer.
König, M., Tessone, C., and Zenou, Y. (2013). Nestedness in networks: A theoretical model and some applications. *Forthcoming in Theoretical Economics*.
König, M. D. (2011). Technology cycles in dynamic R&D networks. *Mimeo, University of Zurich*.
König, M. D., Battiston, S., Napoletano, M., and Schweitzer, F. (2011). The efficiency and stability of R&D networks. *Games and Economic Behavior*, 75:694–713.
König, M. D., Liu, X., and Zenou, Y. (2014). R&D networks: Theory, empirics and policy implications. *CEPR Discussion Paper No. 9872*.
König, M. D., Lorenz, J., and Zilibotti, F. (2012). Innovation vs. imitation and the evolution of productivity distributions. *CEPR Discussion Paper No. 8843*.
Lee, G., Tam, N., and Yen, N. (2005). *Quadratic programming and affine variational inequalities: a qualitative study*. Springer Verlag.
Liben-Nowell, D. and Kleinberg, J. (2007). The link-prediction problem for social networks. *Journal of the American society for information science and technology*, 58(7):1019–1031.
Luttmer, E. (2007). Selection, Growth, and the Size Distribution of Firms. *The Quarterly Journal of Economics*, 122(3):1103–1144.
Mahadev, N. and Peled, U. (1995). *Threshold Graphs and Related Topics*. North Holland.
Mattsson, L. and Weibull, J. (2002). Probabilistic choice and procedurally bounded rationality. *Games and Economic Behavior*, 41(1):61–78.
Mele, A. (2010). Segregation in Social Networks: Theory, Estimation and Policy. NET Institute Working Paper 10-16.
Melitz, M., Ottaviano, G., and Maggiore, S. (2008). Market Size, Trade, and Productivity. *Review of Economic Studies*, 75(1):295–316.
Mitzenmacher, M. (2004). A Brief History of Generative Models for Power Law and Lognormal Distributions. *Internet Mathematics*, 1(2):226–251.
Monderer, D. and Shapley, L. (1996). Potential Games. *Games and Economic Behavior*, 14(1):124–143.
Nooteboom, B., Van Haverbeke, W., Duysters, G., Gilsing, V., and Van Den Oord, A. (2007). Optimal cognitive distance and absorptive capacity. *Research Policy*, 36(7):1016–1034.
Park, J. and Newman, M. (2004). Statistical mechanics of networks. *Physical Review E*, 70(6):66117.
Powell, W., Koput, K. W., and Smith-Doerr, L. (1996). Interorganizational collaboration and the locus of innovation:

Networks of learning in biotechnology. *Administrative Science Quarterly*, 41(1):116–145.
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.
Roijakkers, N. and Hagedoorn, J. (2006). Inter-firm R&D partnering in pharmaceutical biotechnology since 1975: Trends, patterns, and networks. *Research Policy*, 35(3):431–446.
Rosenkopf, L. and Schilling, M. (2007). Comparing alliance network structure across industries: Observations and explanations. *Strategic Entrepreneurship Journal*, 1:191–209.
Rue, H. and Held, L. (2005). *Gaussian Markov random fields: theory and applications*, volume 104. Chapman & Hall.
Sandholm, W. (2010). *Population games and evolutionary dynamics*. MIT Press.
Schilling, M. (2009). Understanding the alliance data. *Strategic Management Journal*, 30(3):233–260.
Singh, N. and Vives, X. (1984). Price and quantity competition in a differentiated duopoly. *The RAND Journal of Economics*, 15(4):546–554.
Snijders, T. (2001). The Statistical Evaluation of Social Network Dynamics. *Sociological Methodology*, 31(1):361–395.
Söderberg, B. (2002). General formalism for inhomogeneous random graphs. *Physical Review E*, 66(6):066121.
Stanley, M., Amaral, L., Buldyrev, S., Havlin, S., Leschhorn, H., Maass, P., Salinger, M., and Stanley, H. (1996). Scaling behaviour in the growth of companies. *Nature*, 379(6568):804–806.
Topkis, D. M. (1998). *Supermodularity and complementarity*. Princeton University Press.
Van Der Hofstad, R. (2009). Random graphs and complex networks.
Weibull, J. (1997). *Evolutionary game theory*. The MIT press.
Westbrock, B. (2010). Natural concentration in industrial research collaboration. *The RAND Journal of Economics*, 41(2):351–371.
Wong, R. (2001). *Asymptotic approximations of integrals*, volume 34. Society for Industrial Mathematics.

Appendix

A. Equilibrium Analysis for Exogenous Networks

In this section we consider the case of $\lambda = \xi = 0$, when the network is exogenously given. Let us first assume that $\zeta = 0$, such that firms do not incur a fixed cost for an R&D collaboration. Then the best response of firm i can be written as follows (see also Equation (10))

$$q_i = \min\left(\bar{q}, \max\left(0, f_i(q_{-i}, G)\right)\right). \tag{22}$$

Observe that if $f_i(q_{-i}, G) < 0$, then the business stealing effects become so large, that firm i's best response is to leave the market ($q_i = 0$). If $\zeta > 0$ then firm i's best response is

$$q_i = \begin{cases} \min\left(\bar{q}, f_i(q_{-i}, G)\right), & \text{if } \pi_i(f_i(q_{-i}, G), G) > 0, \\ 0, & \text{otherwise.} \end{cases} \tag{23}$$

The set of networks G for which $q_i > 0$ is starkly reduced if $\zeta > 0$, and if ζ becomes large enough, no firm will operate at positive quantity.

We now analyze the best response dynamics of quantities in a fixed network G where firms adjust their output levels optimally, given the output levels of all other firms in the industry [Corchon and Mas-Colell, 1996; Weibull, 1997]. We also assume that an interior equilibrium exists. This dynamics is given by

$$\frac{dq_i}{dt} = f_i(q_{-i}, G) - q_i = \frac{\eta}{2\nu} - \frac{b}{2n\nu}\sum_{j \neq i}^{n} q_j + \frac{\rho}{2n\nu}\sum_{j=1}^{n} a_{ij} q_j - q_i, \tag{24}$$

with some appropriate initial conditions $\mathbf{q}(0) \geq 0$. The equilibrium quantities for a given network G can be obtained as the fixed points of the best response dynamics. In vector-matrix notation the dynamics can be written as

$$\frac{d\mathbf{q}}{dt} = \frac{\eta}{2\nu}\mathbf{u} - \frac{1}{2n\nu}\left((2n\nu - b)\mathbf{I}_n + b\mathbf{u}\mathbf{u}^\top - \rho\mathbf{q}\right)\mathbf{q}. \tag{25}$$

This is an inhomogeneous linear first-order ordinary differential equation with constant coefficients. Let us denote by $\mathbf{U} = \mathbf{u}\mathbf{u}^\top$ and introduce the matrix

$$\mathbf{Q} \equiv \mathbf{I}_n + \frac{b}{2n\nu - b}\mathbf{U} - \frac{\rho}{2n\nu - b}\mathbf{q}. \tag{26}$$

The solution of Equation (25) is stable if an only if all eigenvalues of $\mathbf{Q}$ have a positive real part. If a stable solution exists and if $\mathbf{Q}$ is invertible, then the steady state $\mathbf{q}^* = \lim_{t\to\infty} \mathbf{q}(t)$ is given by

$$\mathbf{q}^* = \frac{n\eta}{2n\nu - b}\mathbf{Q}^{-1}\mathbf{u}, \tag{27}$$

and the solution trajectory is given by

$$\mathbf{q}(t) = \mathbf{q}^* + e^{-\mathbf{Q}t}(\mathbf{q}(0) - \mathbf{q}^*). \tag{28}$$

If $\mathbf{q}(0) = 0$ then we can write

$$\mathbf{q}(t) = \frac{n\eta}{2n\nu - b}\left(1 - e^{-\mathbf{Q}t}\right)\mathbf{Q}^{-1}\mathbf{u}. \tag{29}$$

We have that $\mathbf{Q} = \mathbf{I}_n - \frac{\rho}{2n\nu - b}\mathbf{q}$, and $\lambda_i(\mathbf{Q}) = 1 - \frac{\rho}{2n\nu - b}\lambda_i(\mathbf{A})$. This implies that the stability condition $\lambda_{\min}(\mathbf{Q}) < 0$ is equivalent to $\lambda_{\max}(\mathbf{A}) > \frac{2n\nu - b}{\rho}$.

Observe that the profit function introduced in Equation (8) admits a potential game with a corresponding potential function [cf. Monderer and Shapley, 1996]. This is stated in the following proposition.

Proposition 11. *For a given network $G \in \mathcal{G}_n$, the profit function of Equation (8) admits a potential game with potential function $\phi(q, G)\colon \mathbb{R}^n_+ \times \mathcal{G}_n \to \mathbb{R}$ given by*

$$\phi(\mathbf{q}, G) = \sum_{i=1}^{n} n(\eta q_i - \nu q_i^2) - \frac{b}{2}\sum_{i=1}^{n}\sum_{j\neq i} q_i q_j + \frac{\rho}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij} q_i q_j. \tag{30}$$

The existence of a potential function given in Equation (47) allows us to state a condition for a Nash equilibrium in a static network. For a given network G and no fixed linking costs, $\zeta = 0$, a Nash equilibrium of our game solves the following constrained optimization problem [cf. Sandholm, 2010, Sec. 3.1.4]

$$\max_{\mathbf{q}\in\mathbb{R}^n_+} \ \phi(\mathbf{q}, G) \tag{31}$$

$$\text{s.t.} \quad \forall i = 1, \ldots, n$$

$$\frac{\partial \phi}{\partial q_i} = 0 \text{ and } q_i > 0, \text{ or} \tag{32}$$

$$\frac{\partial \phi}{\partial q_i} \leq 0 \text{ and } q_i = 0. \tag{33}$$

We can write the potential function $\phi(\mathbf{q}, G)$ in vector-matrix notation as follows

$$\phi(\mathbf{q}, G) = n\eta \mathbf{u}^\top \mathbf{q} - \frac{1}{2}\mathbf{q}^\top \underbrace{\left((2n\nu - b)\mathbf{I}_n + b\mathbf{U} - \rho\mathbf{q}\right)}_{=(2n\nu - b)\mathbf{Q}} \mathbf{q}. \tag{34}$$

The Hessian of the potential is given by $\Delta\phi(\mathbf{q}, G) = \left(\frac{\partial^2 \phi(\mathbf{q}, G)}{\partial q_i \partial q_j}\right)_{i,j\in\mathcal{N}} = -(2\nu - b)\mathbf{Q}$. If $2n\nu > b$ and the matrix $\mathbf{Q}$ is positive definite then $\Delta\phi(\mathbf{q}, G) < 0$ is negative definite.[22] The matrix $\mathbf{Q}$ is positive definite if and only if the matrix $\mathbf{B} = \mathbf{I}_n - \frac{\rho}{2n\nu - b}\mathbf{A}$ is positive definite. If $\mathbf{B}$ is positive definite then its inverse $\mathbf{B}^{-1}$ exists and is positive definite. $\mathbf{B}^{-1}$ exists if and only if the following eigenvalue condition is satisfied

$$\frac{\rho}{2n\nu - b} < \frac{1}{\lambda_{\max}(\mathbf{A})}, \tag{35}$$

where $\lambda_{\max}(\mathbf{A})$ is the largest (real) eigenvalue of the (real and symmetric) adjacency matrix $\mathbf{A}$. If the inequality in (35) is satisfied, then the maximization of $\phi(\mathbf{q}, G)$ as stated in (33), for a given G, is a linear-quadratic programming problem [Boyd and Vandenberghe, 2004; Lee et al., 2005], where $\phi(\mathbf{q}, G)$ is a concave function of $\mathbf{q}$, and this optimization problem has a unique solution.

In the following we assume that the inequality in (35) is satisfied. Then we can obtain firms' equilibrium quantities and profits as follows:

Proposition 12. *Denote by $\varphi = \frac{\rho}{2n\nu - b} = \frac{\alpha\beta}{2(2-b)\gamma - \alpha^2}$ and consider a network $G \in \mathcal{G}_n$ satisfying $\varphi < 1/\lambda_{\max}(\mathbf{q})$.*

(i) If $\zeta = 0$, equilibrium output and profit are given by

$$q_i = \frac{n\eta}{2\nu + b(\|\mathbf{b}(G, \varphi)\| - 1)} b_i(G, \varphi) \tag{36}$$

and

$$\pi_i = \frac{\nu n^2 \eta^2}{(2n\nu + b(\|\mathbf{b}(G, \varphi) - 1\|))^2} b_i^2(G, \varphi), \tag{37}$$

for all $i = 1, \ldots, n$.

[22] The $n \times n$ matrix $\mathbf{Q}$ is positive definite if and only if for all $q \in \mathbb{R}^n_+$ we have that $\mathbf{q}^\top Q \mathbf{q} > 0$. If $\mathbf{Q}$ is positive definite, then all its eigenvalues are positive.

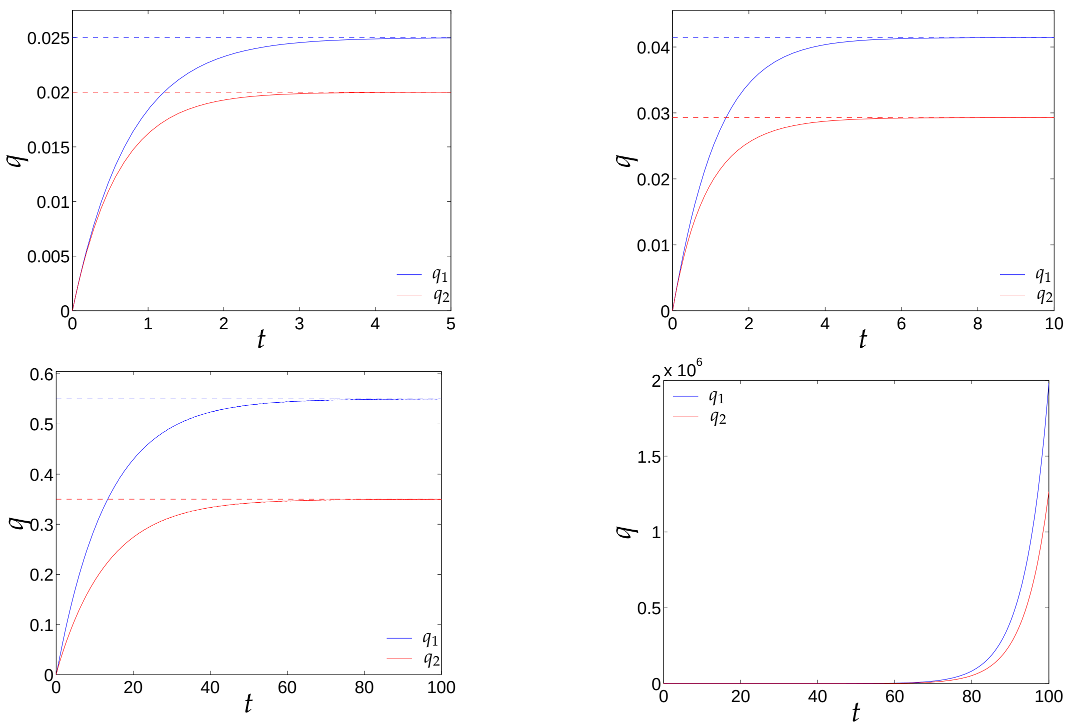

Figure 6: A star network with $n = 3$, $\lambda_{\max}(A) = \sqrt{2} = 1.41$ and no linking costs, i.e. $\zeta = 0$. The top left panel shows the evolution $\mathbf{q}(t)$ for $\varphi = 1/3 < \lambda_{\max}(A)^{-1} = 0.70$, the top right panel for $\varphi = \lambda_{\max}(A)^{-1} = 0.70$, the bottom left panel for $\varphi = 4/3 > \lambda_{\max}(A)^{-1} = 0.70$, and the bottom right panel for $\varphi = 3/2 > \lambda_{\max}(A)^{-1} = 0.70$. The dashed lines indicate the solutions from Equation (36).

(ii) If $\zeta > 0$ and $\pi_i > \zeta d_i$ for all $i = 1, \ldots, n$ in Equation (37) then equilibrium quantities are given by Equation (36) and equilibrium profits are given by Equation (37) less the cost of collaboration ζd_i.

Observe that, in the limit $\varphi \uparrow \lambda_{\max}^{-1}$, the normalized Bonacich centrality converges to the eigenvector centrality $\mathbf{v}$, where $\mathbf{qv} = \lambda_{\max}\mathbf{v}$. This implies that

$$\lim_{\varphi \uparrow \lambda_{\max}^{-1}} q_i = \frac{n\eta}{b} v_i \tag{38}$$

and

$$\lim_{\varphi \uparrow \lambda_{\max}^{-1}} \pi_i = \frac{n^2\eta^2\nu}{b^2} v_i^2 - \zeta d_i, \tag{39}$$

for all $i = 1, \ldots, n$.

Note that if the eigenvalue condition (35) is not satisfied, then corner solutions must be considered.[23]

In Figure 6 we give an example of the evolution of $\mathbf{q}(t)$ for the star network $K_{1,n-1}$ with $n = 3$. The stationary state $\mathbf{q}^*$ for values of $\varphi \leq \lambda_{\max}(\mathbf{q})^{-1}$ is correctly described by Equation (36). Interestingly, this solution is also correct for values of $\varphi > \lambda_{\max}(\mathbf{q})^{-1}$ (see bottom left panel in Figure 6), unless the equilibrium quantities from Equation (36) grow without bound (see bottom right panel in Figure 6), and $q_i(t)$ grows to its capacity constraint $\bar{q}$.

[23]See Bramoullé et al. [2010] for the case of strategic substitutes and Cabrales et al. [2010] for the case of strategic complements and linking strengths proportional to socialization effort.

B. Extensions

B.1. Endogenous Networks with Heterogeneous Spillovers among Firms

In this section we allow for heterogeneity among firms in terms of their technological abilities. We assume that the knowledge embodied in a firm $i \in \mathcal{I} = \{1, \ldots, n\}$ can be represented as an N-dimensional vector $\mathbf{h}_i$ in the knowledge space $\mathcal{H}^N = \{0,1\}^N$, which consists of all binary sequences with elements in $\{0,1\}$ of length N. The number of such sequences is 2^N. The knowledge vector $\mathbf{h}_i$, with components $h_{ik} \in \{0,1\}$, indicates whether firm i knows idea $k \in \{1, \ldots, N\}$ or not. We introduce a spillover function $f : \mathcal{H}^N \times \mathcal{H}^N \to \mathbb{R}$ capturing the potential technology transfer between any pairs of firms. A simple choice for the function f could be $f(\mathbf{h}_i, \mathbf{h}_j) = a|\mathbf{h}_i \cap \mathbf{h}_j|$, where $a \in \mathbb{R}_+$ and $|\mathbf{h}_i \cap \mathbf{h}_j| = \mathbf{h}_i^\top \mathbf{h}_j = \sum_{k=1}^N h_{ik}h_{jk}$ denotes the common knowledge of i and j. Or a "gravity function" of the form $f(\mathbf{h}_i, \mathbf{h}_j) = \mathbb{1}_{\{|S(\mathbf{h}_i)| \cdot |S(\mathbf{h}_j)| > \tau\}}$ where $|S(\mathbf{h}_i)|$ counts the number of technologies known to i and $\tau > 0$ is a threshold. Alternative specifications for similarity can be found in Liben-Nowell and Kleinberg [2007] and Bloom et al. [2007]; Jaffe [1989]. Alternatively, following Berliant and Fujita [2008, 2009], a possible parametric specification for f would be $f(\mathbf{h}_i, \mathbf{h}_j) = |\mathbf{h}_i \cap \mathbf{h}_j|^\kappa d(\mathbf{h}_i, \mathbf{h}_j)^{\frac{1-\kappa}{2}}$ for some $\kappa \in (0,1)$. The distance is the product of the total number of ideas known by agent i but not by j times the total number of ideas known by j but not by i, i.e. $d(\mathbf{h}_i, \mathbf{h}_j) = |\mathbf{h}_i \backslash \mathbf{h}_j| \times |\mathbf{h}_j \backslash \mathbf{h}_i| = |\mathbf{h}_i \cap \mathbf{h}_j^c| \times |\mathbf{h}_i^c \cap \mathbf{h}_j| = \sum_{k=1}^N h_{ik}(1 - h_{jk}) \sum_{k=1}^N (1 - h_{ik})h_{jk}$, where $\mathbf{u} = (1, \ldots, 1)^\top$ and $\mathbf{h}_i^c = \mathbf{u} - \mathbf{h}_i$.[24]

Given the spillover function $f(\mathbf{h}_i, \mathbf{h}_j)$, the marginal cost of production of a firm i becomes

$$c_i = \bar{c} - \alpha e_i - \beta \sum_{j=1}^n a_{ij} f(\mathbf{h}_i, \mathbf{h}_j) e_j$$

and profits of firm i are given by

$$\pi_i = (a - \bar{c})q_i - q_i^2 - bq_i \sum_{j \neq i} q_j + \alpha q_i e_i + \beta q_i \sum_{j=1}^n a_{ij} f(\mathbf{h}_i, \mathbf{h}_j) e_j - \gamma e_i^2 - \zeta d_i.$$

The FOC with respect to effort e_i is given by

$$\frac{\partial \pi_i}{\partial e_i} = \alpha q_i - 2\gamma e_i = 0,$$

from which it follows that

$$e_i = \frac{\alpha}{2\gamma} q_i = \lambda q_i.$$

Inserting into profits yields

$$\begin{aligned}
\pi_i &= (a - \bar{c})q_i - (1 - \lambda\alpha + \lambda^2\gamma)q_i^2 - bq_i \sum_{j \neq i} q_j + \lambda\beta q_i \sum_{j=1}^n a_{ij} f(\mathbf{h}_i, \mathbf{h}_j) q_j - \zeta d_i \\
&= \eta n q_i - \nu n q_i^2 - bq_i \sum_{j \neq i} q_j + \rho q_i \sum_{j=1}^n a_{ij} f(\mathbf{h}_i, \mathbf{h}_j) q_j - \zeta d_i.
\end{aligned}$$

We can then obtain a potential function given by

$$\Phi(\mathbf{q}, G, \mathbf{h}) = \sum_{i=1}^n ((a - \bar{c})q_i - \nu q_i^2) - \frac{b}{2} \sum_{i=1}^n q_i \sum_{j \neq i} q_j + \sum_{i=1}^n q_i \sum_{j=1}^n a_{ij} f(\mathbf{h}_i, \mathbf{h}_j) q_j - \zeta m.$$

[24] Other functional forms have been suggested in the literature [see e.g. Baum et al., 2009; Nooteboom et al., 2007], such as $f(\mathbf{h}_i, \mathbf{h}_j) = a_1|\mathbf{h}_i \cap \mathbf{h}_j| - a_2|\mathbf{h}_i \cap \mathbf{h}_j|^2$, with constants $a_1, a_2 \geq 0$.

The stationary distribution is given by

$$\mu^{\vartheta}(\mathbf{q}, G, \mathbf{h}) = \frac{e^{\vartheta\Phi(\mathbf{q},G,\mathbf{h})}}{\sum_{\mathbf{h}'\in\mathcal{H}^N}\sum_{H\in\mathcal{G}_n}\int_{[0,\bar{q}]^n} e^{\vartheta\Phi(\mathbf{s},H,\mathbf{h}')}d\mathbf{s}}.$$

The probability of observing a network $G \in \mathcal{G}_n$, given an output distribution $\mathbf{q} \in [0, \bar{q}]^n$ and technology portfolios $\mathbf{h} \in \mathcal{H}^N$ is determined by conditional distribution

$$\mu^{\vartheta}(G|\mathbf{q}, \mathbf{h}) = \prod_{i<j} \frac{e^{\vartheta a_{ij}(\rho f(\mathbf{h}_i,\mathbf{h}_j)q_i q_j - \zeta)}}{1 + e^{\vartheta(\rho f(\mathbf{h}_i,\mathbf{h}_j)q_i q_j - \zeta)}}, \tag{40}$$

which is equivalent to the probability of observing an inhomogeneous random graph with link probability

$$p_{ij} = p(\mathbf{h}_i, q_i, \mathbf{h}_j, q_j) = \frac{e^{\vartheta(\rho f(\mathbf{h}_i,\mathbf{h}_j)q_i q_j - \zeta)}}{1 + e^{\vartheta(\rho f(\mathbf{h}_i,\mathbf{h}_j)q_i q_j - \zeta)}}. \tag{41}$$

Note that an inhomogeneous random graph with a link probability similar to the one in Equation (41) has been analyzed in Boguñá et al. [2004]. The authors show that if the technology levels are drawn from a multivariate uniform distribution a number of network characteristics can be computed which closely reproduce the empirically observed patterns of R&D networks.

Let the degree distribution be given by $P(k)$ for $k = 0, \ldots, n-1$. Further, let $g(k|\mathbf{h}, q)$ be the conditional probability that a firm with technology vector $\mathbf{h}$ and output q has k links. Then the degree distribution can be written as follows [cf. Boguñá and Pastor-Satorras, 2003; Söderberg, 2002]

$$P(k) = \sum_{\mathbf{h}\in\mathcal{H}^N}\sum_{q\in\mathcal{Q}} g(k|\mathbf{h}, q) f(\mathbf{h}) \mu^{\vartheta}(q),$$

where $f(\mathbf{h})$ is the probability distribution over firms with technology $\mathbf{h} \in \mathcal{H}^N$. The average degree of a firm with technology $\mathbf{h}$ and output q is then given by

$$\bar{d}(\mathbf{h}, q) = \sum_{k=0}^{n-1} k g(k|\mathbf{h}, q),$$

and the average degree is given by

$$\bar{d} = \sum_{k=0}^{n-1} kP(k) = \sum_{\mathbf{h}\in\mathcal{H}^N}\sum_{q\in\mathcal{Q}} f(\mathbf{h})\mu^{\vartheta}(q)\bar{d}(\mathbf{h}, q).$$

The probability that a firm with degree k has technology $\mathbf{h}$ and output q is given by Bayes' rule as

$$g(\mathbf{h}, q|k) = \frac{f(\mathbf{h})\mu^{\vartheta}(q) g(k|\mathbf{h}, q)}{P(k)}.$$

Let $X_{q,q'}^{\mathbf{h},\mathbf{h}'}(\omega)$ count the number of links between firms with technology $\mathbf{h}$ and output q, and technology $\mathbf{h}'$ and output q' in a state $\omega \in \Omega$. Then we have that

$$X_{q,q'}^{\mathbf{h},\mathbf{h}'}(\omega) = \sum_{i=1}^{n}\sum_{j=i+1}^{n} \mathbb{1}_{\{q_i(\omega)=q\}}\mathbb{1}_{\{\mathbf{h}_i(\omega)=\mathbf{h}\}}\mathbb{1}_{\{q_j(\omega)=q'\}}\mathbb{1}_{\{\mathbf{h}_j(\omega)=\mathbf{h}'\}}\mathbb{1}_{\{a_{ij}(\omega)=1\}}.$$

If $X_q^{\mathbf{h}}$ counts the number of links of a firm with technology $\mathbf{h}$ and output q then we can write $X_q^{\mathbf{h}} = \sum_{\mathbf{h}'\in\mathcal{H}^N}\sum_{q'\in\mathcal{Q}} X_{q,q'}^{\mathbf{h},\mathbf{h}'}$, and we have that $\mathbb{P}(X_q^{\mathbf{h}} = k) = g(k|\mathbf{h}, q)$. Observe that the random variables $X_{q,q'}^{\mathbf{h},\mathbf{h}'}$ are independent, and binomially distributed with

$$\mathbb{P}(X_{q,q'}^{\mathbf{h},\mathbf{h}'} = k) = \binom{n(\mathbf{h}', q')}{k} p(\mathbf{h}, q, \mathbf{h}', q')^k (1 - p(\mathbf{h}, q, \mathbf{h}', q'))^{n(\mathbf{h}',q')-k},$$

where $n(\mathbf{h}', q') = f(\mathbf{h}')\mu^{\vartheta}(q')$, because the links between firms with technologies $\mathbf{h}$ and $\mathbf{h}'$ and output levels q and q' are independently drawn with probability $p(\mathbf{h}, q, \mathbf{h}', q')$. The generating function of the random variable $X_{q,q'}^{\mathbf{h},\mathbf{h}'}$ is given by [25]

$$\hat{g}(z|\mathbf{h}, q, \mathbf{h}', q') = (1 - (1 - z)p(\mathbf{h}, q, \mathbf{h}', q'))^{n(\mathbf{h}', q')}.$$

where $p(\mathbf{h}, q, \mathbf{h}', q')$ is given by Equation (41), and can be written as

$$p(\mathbf{h}, q, \mathbf{h}', q') = \frac{e^{\vartheta(\rho f(\mathbf{h},\mathbf{h}')qq'-\zeta)}}{1 + e^{\vartheta(\rho f(\mathbf{h},\mathbf{h}')qq'-\zeta)}}.$$

Since the random variable $X_q^{\mathbf{h}}$ is a sum of independent binomial random variables $X_{q,q'}^{\mathbf{h},\mathbf{h}'}$, we have that the generating function $\hat{g}(z|\mathbf{h}, q)$ of $X_q^{\mathbf{h}}$ is the product of the generating functions $\hat{g}(z|\mathbf{h}, q, \mathbf{h}', q')$ of $X_{q,q'}^{\mathbf{h},\mathbf{h}'}$, so that after taking logs we get

$$\ln \hat{g}(z|\mathbf{h}, q) = n \sum_{\mathbf{h}' \in \mathcal{H}^N} \sum_{q' \in \mathcal{Q}} f(\mathbf{h}')\mu^{\vartheta}(q') \ln \left(1 - (1 - z)p(\mathbf{h}, q, \mathbf{h}', q')\right). \tag{42}$$

We then can solve this equation to obtain $\hat{g}(z|\mathbf{h}, q)$, and from inverting the generating function

$$g(k|\mathbf{h}, q) = \frac{1}{k!} \frac{d^k \hat{g}(0|\mathbf{h}, q)}{dz^k}\bigg|_{z=0},$$

the degree distribution follows as

$$P(k) = \sum_{\mathbf{h} \in \mathcal{H}^N} \sum_{q \in \mathcal{Q}} f(\mathbf{h})\mu^{\vartheta}(q)g(k|\mathbf{h}, q).$$

When the linking cost c is high, and the connection probability $p(\mathbf{h}, q, \mathbf{h}', q')$ is small, we can expanding Equation (42) as

$$\hat{g}(z|\mathbf{h}, q) \approx e^{n \sum_{\mathbf{h}' \in \mathcal{H}^N} \sum_{q' \in \mathcal{Q}} f(\mathbf{h}')\mu^{\vartheta}(q')p(\mathbf{h},q,\mathbf{h}',q')(z-1)}.$$

This is the generating function of a Poisson distribution with mean

$$\bar{d}(\mathbf{h}, q) = n \sum_{\mathbf{h}' \in \mathcal{H}^N} \sum_{q' \in \mathcal{Q}} f(\mathbf{h}')\mu^{\vartheta}(q')p(\mathbf{h}, q, \mathbf{h}', q'),$$

which means that

$$g(k|\mathbf{h}, q) = \frac{1}{k!} e^{-\bar{d}(\mathbf{h},q)} \bar{d}(\mathbf{h}, q)^k,$$

and the degree distribution is given by

$$P(k) = \frac{1}{k!} \sum_{\mathbf{h} \in \mathcal{H}^N} \sum_{q \in \mathcal{Q}} f(\mathbf{h})\mu^{\vartheta}(q)e^{-\bar{d}(\mathbf{h},q)}\bar{d}(\mathbf{h}, q)^k,$$

while the average degree of a firm with technology $\mathbf{h}$ and output q is given by

$$\bar{d}(\mathbf{h}, q) = n \sum_{\mathbf{h}' \in \mathcal{H}^N} \sum_{q' \in \mathcal{Q}} f(\mathbf{h}')\mu^{\vartheta}(q')p(\mathbf{h}, q, \mathbf{h}', q').$$

In the following we consider a special case in which the link function takes the form $f(\mathbf{h}_i, \mathbf{h}_j) = |\mathsf{S}(\mathbf{h}_i)| \cdot |\mathsf{S}(\mathbf{h}_j)|$, where $|\mathsf{S}(\mathbf{h}_i)|$ counts the number of technologies known to i. This functional form expresses complementarity effects between the stocks of knowledge between firms i and j. As-

[25] The probability generating function of a binomial random variable, with the number of successes in n trials, and probability p of success in each trial, is given by $g(z) = (1 - p(1 - z))^n$.

sume that ideas are distributed to firms following a stochastic urn process, where the probability that a firm i obtains a new idea k is proportional to the number of ideas $|\mathcal{S}(\mathbf{h}_i)|$ it already has. This process generates a power law distribution over the number of ideas firms possess [Gabaix, 2009; Mitzenmacher, 2004].[26] Let $s = 0, \ldots,$ denote the stock of knowledge of a firm, and let $f(s) = \mathbb{P}(|\mathcal{S}(\mathbf{h}_i)| = s)$ be the distribution of the knowledge stocks. Then the degree distribution can be written as

$$P(k) = \frac{1}{k!} \int ds \sum_{q \in \mathcal{Q}} f(s) \mu^{\vartheta}(q) e^{-\bar{d}(s,q)} \bar{d}(s,q)^k,$$

while the average degree of a firm with knowledge stock s and output q is given by

$$\bar{d}(s,q) = n \int ds' \sum_{q' \in \mathcal{Q}} f(s') \mu^{\vartheta}(q') p(s,q,s',q').$$

Assuming that the knowledge stocks are power-law distributed [cf. e.g. König et al., 2012; Melitz et al., 2008], $f(s) \sim s^{-\gamma}$, we can write

$$\ln \hat{g}(z|s,q) = n \sum_{q' \in \mathcal{Q}} \mu^{\vartheta}(q') \int ds'(s')^{-\gamma} \ln \left(1 - (1-z) \frac{e^{\vartheta(\rho s s' q q' - \zeta)}}{1 + e^{\vartheta(\rho s s' q q' - \zeta)}} \right).$$

In the limit of $\vartheta \to \infty$ we have that

$$\lim_{\vartheta \to \infty} \frac{e^{\vartheta(\rho s s' q q' - \zeta)}}{1 + e^{\vartheta(\rho s s' q q' - \zeta)}} = \mathbb{1}_{\{\rho s s' q q' > \zeta\}},$$

so that we can write

$$\ln \hat{g}(z|s,q) = n \int_{\frac{\zeta}{\rho s q q^*}}^{\infty} ds s^{-\gamma} \ln z = n \frac{\left(\frac{\zeta}{\rho s q q^*}\right)^{1-\gamma}}{\gamma - 1} \ln z,$$

and we get

$$\hat{g}(z|s,q) = z^{\frac{n}{\gamma-1}\left(\frac{\zeta}{\rho s q q^*}\right)^{1-\gamma}}.$$

It then follows that

$$g(k|s,q) = \delta \left(k - \frac{n}{\gamma - 1} \left(\frac{\zeta}{\rho s q q^*} \right)^{1-\gamma} \right).$$

The degree distribution is then given by

$$\begin{aligned}
P(k) &= \int ds g(k|s,q^*) s^{-\gamma} \\
&= \int ds \delta \left(k - \frac{n}{\gamma - 1} \left(\frac{\zeta}{\rho s (q^*)^2} \right)^{1-\gamma} \right) s^{-\gamma} \\
&= \zeta^{-\gamma} \rho^{\gamma} (q^*)^{2\gamma} \left(\frac{(\gamma - 1)k}{n} \right)^{\frac{\gamma}{1-\gamma}} \\
&\sim k^{-\frac{\gamma}{\gamma-1}}.
\end{aligned}$$

Hence, we obtain a power-law degree distribution with parameter $\frac{\gamma}{\gamma-1}$, confirming previous empirical studies [e.g. Powell et al., 2005]. An illustration can be seen in Figure 5 for the case of $\gamma = 2$.

[26] Alternatively, if each technology category $k = 1, \ldots, N$ is drawn independently with probability $p \in [0,1]$, then $|\mathcal{S}(\mathbf{h}_i)|$ has a Binomial distribution with success probability p, that is, $\mathbb{P}(|\mathcal{S}(\mathbf{h}_i)| = s) = \binom{N}{k} p^s (1-p)^{N-s}$.

B.2. Endogenous Networks with Heterogeneous Marginal Costs

We consider ex ante heterogeneity among firms in the variable cost $\bar{c}_i \geq 0$ [see also Banerjee and Duflo, 2005], expressing their different technological and organizational capabilities.[27] The marginal cost of production of firm $i \in \mathcal{N}$ is then given by

$$c_i(\mathbf{e}, G) = \bar{c}_i - \alpha e_i - \beta \sum_{j=1}^{n} a_{ij} e_j, \tag{43}$$

where $e_i \in [0, \bar{e}_i]$ and $\alpha, \beta \in [0,1]$. Requiring that $c_i \geq 0$ we must have that $\bar{c}_i \geq \sum_{j=1}^{n} e_j = n\bar{e}$ for all $i = 1, \ldots, n$. Hence, $\bar{c}_i$ is $O(n)$. Similarly, as in the previous sections, the first-order conditions for efforts imply that $e_i = \max\{\frac{\alpha}{2\gamma} q_i, \bar{e}_i\}$. The non-negativity of marginal cost in the case of an interior equilibrium then requires that $q_i \leq \frac{2\gamma}{\alpha} \bar{e}_i$ for all $i = 1, \ldots, n$. We further assume that $a > \max_{1 \leq i \leq n}\{\bar{c}_i\}$. We denote by $\lambda = \frac{\alpha}{2\gamma}$. Profits of firm i from Equation (7) then become

$$\pi_i = \underbrace{(a - \bar{c}_i)}_{n\eta_i} q_i - b \sum_{j \neq i} q_i q_j + \alpha q_i e_i + \beta \sum_{j=1}^{n} a_{ij} q_i e_j - \gamma e_i^2 - \zeta d_i. \tag{44}$$

Using the fact that in the interior equilibrium $e_i = \lambda q_i$, we can write firm i's profit as

$$\pi_i = n\eta_i q_i - n\nu q_i^2 - b q_i \sum_{j \neq i} q_j + \rho q_i \sum_{j=1}^{n} a_{ij} q_j - \zeta d_i. \tag{45}$$

We first compute the equilibrium quantities for a given network G. The FOC can be written as

$$\frac{\partial \pi_i}{\partial q_i} = n\eta_i - 2n\nu q_i - b \sum_{j \neq i} q_j + \rho \sum_{j=1}^{n} a_{ij} q_j = n\eta_i - (2\nu - b) q_i - b\|\mathbf{q}\| + \rho \sum_{j=1}^{n} a_{ij} q_j = 0.$$

Let $\boldsymbol{\eta} \equiv (\eta_1, \ldots, \eta_n)^\top$. Then in vector-matrix notation this is

$$n\boldsymbol{\eta} = ((2n\nu - b)\mathbf{I}_n - \rho\mathbf{q})\mathbf{q} + b\mathbf{u}\|q\|.$$

Denoting by $\boldsymbol{\mu} \equiv \frac{n}{2n\nu - b}\boldsymbol{\eta}$, $\phi \equiv \frac{\rho}{2n\nu - b}$ and $\kappa \equiv \frac{b}{2n\nu - b}$ this is

$$\mathbf{q} = (\mathbf{I} - \phi\mathbf{q})^{-1}(\boldsymbol{\mu} - \kappa\|\mathbf{q}\|)\mathbf{u}.$$

Following Calvó-Armengol et al. [2009] we define the $\boldsymbol{\mu}$-weighted Bonacich centrality as

$$\mathbf{b}_{\boldsymbol{\mu}}(G, \phi) = (\mathbf{I}_n - \phi\mathbf{q})^{-1}\boldsymbol{\mu} = \sum_{k=0}^{\infty} \phi^k \mathbf{q}^k \boldsymbol{\mu}, \tag{46}$$

where suitable conditions have to be imposed on the vector $\boldsymbol{\mu}$, the parameter ϕ and the eigenvalue $\lambda_{\mathrm{PF}}(G)$. The Bonacich centrality is then simply given by $\mathbf{b}(G, \phi) = \mathbf{b}_{\mathbf{u}}(G, \phi)$. Then we can write

$$\mathbf{q} = \mathbf{b}_{\boldsymbol{\mu}}(G, \phi) - \kappa\|\mathbf{q}\|\mathbf{b}_{\mathbf{u}}(G, \phi).$$

Multiplying from the left with $\mathbf{u}^\top$ gives

$$\mathbf{u}^\top\mathbf{q} = \|\mathbf{q}\| = \|\mathbf{b}_{\boldsymbol{\mu}}(G, \phi)\| - \nu\|\mathbf{q}\|\|\mathbf{b}_{\mathbf{u}}(G, \phi)\|.$$

[27] Blundell et al. [1995] argued that because the main source of unobserved heterogeneity in models of innovation lies in the different knowledge stocks with which firms enter a sample, a variable that approximates the build-up of firm knowledge at the time of entering the sample is a particularly good control for unobserved heterogeneity.

from which we get

$$\|\mathbf{q}\| = \frac{\|\mathbf{b}_{\boldsymbol{\mu}}(G,\phi)\|}{1+\kappa\|\mathbf{b}_{\mathbf{u}}(G,\phi)\|}.$$

It follows that equilibrium quantity can be written as

$$\mathbf{q} = \mathbf{b}_{\boldsymbol{\mu}}(G,\phi) - \frac{\kappa\|\mathbf{b}_{\boldsymbol{\mu}}(G,\phi)\|}{1+\kappa\|\mathbf{b}_{\mathbf{u}}(G,\phi)\|}\mathbf{b}_{\mathbf{u}}(G,\phi).$$

Note that the profit function of Equation (45) admits a potential function $\Phi\colon \mathbb{R}_+^n \times \mathcal{G}_n \to \mathbb{R}$ given by

$$\Phi(\mathbf{q},G|\boldsymbol{\eta}) = \sum_{i=1}^n n(\eta_i q_i - \nu q_i^2) - \frac{b}{2}\sum_{i=1}^n\sum_{j\neq i} q_i q_j + \frac{\rho}{2}\sum_{i=1}^n\sum_{j=1}^n a_{ij}q_iq_j - \zeta m, \tag{47}$$

where m is the number of links in G. A similar equilibrium characterization using a Gibbs measure as in the previous sections can thus be obtained. Note further that the Hamiltonian in the partition function $\mathscr{Z}_\vartheta^{\boldsymbol{\eta}} = \sum_{\mathbf{q}\in\mathcal{Q}^n} e^{\vartheta\mathscr{H}(\mathbf{q}|\boldsymbol{\eta})}$ in the case of heterogeneous marginal costs is given by

$$\mathscr{H}(\mathbf{q}|\boldsymbol{\eta}) = \sum_{i=1}^n\left(n\eta_i q_i - n\nu q_i^2 + \sum_{j>i}\left(\frac{1}{\vartheta}\ln\left(1+e^{\vartheta(\rho q_i q_j-\zeta)}\right) - bq_iq_j\right)\right).$$

When $\vartheta\to\infty$ we then can write

$$\lim_{\vartheta\to\infty}\mathscr{H}(\mathbf{q}|\boldsymbol{\eta}) = \sum_{i=1}^n\left(n\eta_iq_i - n\nu q_i^2 + \sum_{j>i}\left(\rho q_iq_j-\zeta\right)\mathbb{1}_{\{\rho q_iq_j>\zeta\}} - bq_iq_j\right).$$

From the maximization of this expression we find that if the capacity constraints $\bar{q}_i$ are binding, then the stochastically stable state will be a threshold graph (nested split graph) in which a link ij is present if and only if $\bar{q}_{(i)}\bar{q}_{(j)} > \frac{\zeta}{\rho}$ and quantities are given by the ordered vector $(\bar{q}_{(1)},\bar{q}_{(2)},\ldots,\bar{q}_{(k)},0,\ldots,0)$ where $k=\max\{1\leq j\leq n : \bar{q}_{(1)}\bar{q}_{(j)} > \frac{\zeta}{\rho}\}$. In the case of finite ϑ we obtain a generalized threshold graph as they have been studied in Boguñá and Pastor-Satorras [2003]; Diaconis et al. [2008]; Ide et al. [2010, 2009]; Söderberg [2002].

C. Proofs

PROOF OF PROPOSITION 1. The potential $\Phi(\mathbf{q},G)$ has the property that

$$\Phi(\mathbf{q},G+ij) - \Phi(\mathbf{q},G) = \rho q_iq_j - \zeta = \pi_i(\mathbf{q},G+ij) - \pi_i(\mathbf{q},G). \tag{48}$$

From the properties of $\pi_i(\mathbf{q},G)$ it also follows that $\Phi(q_i',\mathbf{q}_{-i},G) - \Phi(q_i,\mathbf{q}_{-i},G) = \phi(q_i',\mathbf{q}_{-i},G) - \phi(q_i,\mathbf{q}_{-i},G) = \pi_i(q_i',\mathbf{q}_{-i},G) - \pi_i(q_i,\mathbf{q}_{-i},G)$. □

PROOF OF PROPOSITION 2. In the following we show that the stationary distribution $\mu^\vartheta(\boldsymbol{\omega})$ satisfies the detailed balance condition

$$\mu^\vartheta(\boldsymbol{\omega})p(\boldsymbol{\omega}'|\boldsymbol{\omega}) = \mu^\vartheta(\boldsymbol{\omega}')p(\boldsymbol{\omega}|\boldsymbol{\omega}') \tag{49}$$

where $p(\boldsymbol{\omega}'|\boldsymbol{\omega})$ denotes the transition rate of the Markov chain from state $\boldsymbol{\omega}$ to $\boldsymbol{\omega}'$. Observe that the detailed balance condition is trivially satisfied if $\boldsymbol{\omega}'$ and $\boldsymbol{\omega}$ differ in more than one link or more than one quantity level. Hence, we consider only the case of link creation $G'=G+ij$ (and removal $G'=G-ij$) or an adjustment in quantity $q_i'\neq q_i$ for some $i\in\mathcal{N}$. For the case of link creation with a transition from $\boldsymbol{\omega}=(\mathbf{q},G)$ to $\boldsymbol{\omega}'=(\mathbf{q},G+ij)$ we can write the detailed balance condition as follows

$$e^{\vartheta\left(\Phi(\mathbf{q},G)-m\ln\left(\frac{\xi}{\lambda}\right)\right)}\frac{e^{\vartheta\Phi(\mathbf{q},G+ij)}}{e^{\vartheta\Phi(\mathbf{q},G+ij)}+e^{\vartheta\Phi(\mathbf{q},G)}}\lambda = e^{\vartheta\left(\Phi(\mathbf{q},G+ij)-(m+1)\ln\left(\frac{\xi}{\lambda}\right)\right)}\frac{e^{\vartheta\Phi(\mathbf{q},G)}}{e^{\vartheta\Phi(\mathbf{q},G)}+e^{\vartheta\Phi(\mathbf{q},G+ij)}}\xi.$$

This equality is trivially satisfied. A similar argument holds for the removal of a link with a transition from $\boldsymbol{\omega} = (\mathbf{q}, G)$ to $\boldsymbol{\omega} = (\mathbf{q}, G - ij)$ where the detailed balance condition reads

$$e^{\vartheta(\Phi(\mathbf{q},G)-m\ln\left(\frac{\xi}{\lambda}\right))}\frac{e^{\vartheta\Phi(\mathbf{q},G-ij)}}{e^{\vartheta\Phi(\mathbf{q},G-ij)}+e^{\vartheta\Phi(\mathbf{q},G)}}\xi = e^{\vartheta(\Phi(\mathbf{q},G+ij)-(m-1)\ln\left(\frac{\xi}{\lambda}\right))}\frac{e^{\vartheta\Phi(\mathbf{q},G)}}{e^{\vartheta\Phi(\mathbf{q},G)}+e^{\vartheta\Phi(\mathbf{q},G-ij)}}\lambda.$$

For a change in the output level with a transition from $\boldsymbol{\omega} = (q_i, \mathbf{q}_{-i}, G)$ to $\boldsymbol{\omega}' = (q_i', \mathbf{q}_{-i}, G)$ we get for the detailed balance condition

$$e^{\vartheta(\Phi(q_i,\mathbf{q}_{-i},G)-m\ln\left(\frac{\xi}{\lambda}\right))}\frac{e^{\vartheta\Phi(q_i',\mathbf{q}_{-i},G)}}{e^{\vartheta\Phi(q_i',\mathbf{q}_{-i},G)}+e^{\vartheta\Phi(q_i,\mathbf{q}_{-i},G)}}\chi = e^{\vartheta(\Phi(q_i',\mathbf{q}_{-i},G)-(m+1)\ln\left(\frac{\xi}{\lambda}\right))}\frac{e^{\vartheta\Phi(q_i,\mathbf{q}_{-i},G)}}{e^{\vartheta\Phi(q_i,\mathbf{q}_{-i},G)}+e^{\vartheta\Phi(q_i',\mathbf{q}_{-i},G)}}\chi,$$

Hence, the probability measure $\mu^{\vartheta}(\boldsymbol{\omega})$ satisfies a detailed balance condition and therefore is the stationary distribution of the Markov chain with transition rates $p(\boldsymbol{\omega}'|\boldsymbol{\omega})$. □

PROOF OF PROPOSITION 3. Observe that the potential can be written as

$$\Phi(\mathbf{q}, G) = \underbrace{\sum_{i=1}^{n}\left(n\eta - n\nu q_i - \frac{b}{2}\sum_{j\neq i} q_j\right) q_i}_{\psi(\mathbf{q})} + \sum_{i=1}^{n}\sum_{j=i+1}^{n} a_{ij}\underbrace{(\rho q_i q_j - \zeta)}_{\sigma_{ij}}.$$

We then have that

$$e^{\vartheta\Phi(\mathbf{q},G)} = e^{\vartheta\psi(\mathbf{q})}e^{\vartheta\sum_{i<j}^{n} a_{ij}\sigma_{ij}}.$$

Observe that only the second factor in the above expression is network dependent. We then can use the fact that

$$\sum_{G\in\mathcal{G}_n} e^{\vartheta\sum_{i<j}^{n} a_{ij}\sigma_{ij}} = \prod_{i<j}\left(1 + e^{\vartheta\sigma_{ij}}\right)$$

to obtain

$$\sum_{G\in\mathcal{G}_n} e^{\vartheta\Phi(\mathbf{q},G)} = e^{\vartheta\psi(\mathbf{q})}\prod_{i<j}\left(1 + e^{\vartheta\sigma_{ij}}\right) \tag{50}$$

$$= \prod_{i=1}^{n} e^{\vartheta\left(n\eta - n\nu q_i - \frac{b}{2}\sum_{j\neq i} q_j\right)q_i}\prod_{i<j}\left(1 + e^{\vartheta(\rho q_i q_j - \zeta)}\right). \tag{51}$$

We can use this expression to compute the marginal distribution

$$\mu^{\vartheta}(\mathbf{q}) = \frac{1}{\mathscr{Z}_{\vartheta}}\sum_{G\in\mathcal{G}_n} e^{\vartheta\Phi(\mathbf{q},G)} = \frac{1}{\mathscr{Z}_n^{\vartheta}}\prod_{i=1}^{n} e^{\vartheta\left(n\eta - n\nu q_i - \frac{b}{2}\sum_{j\neq i} q_j\right)q_i}\prod_{i<j}\left(1 + e^{\vartheta(\rho q_i q_j - \zeta)}\right).$$

□

PROOF OF PROPOSITION 4. The conditional distribution is given by

$$\begin{aligned}
\mu^{\vartheta}(G|\mathbf{q}) = \frac{\mu^{\vartheta}(\mathbf{q}, G)}{\mu^{\vartheta}(\mathbf{q})} &= \frac{e^{\vartheta\Phi(\mathbf{q},G)}}{\prod_{i=1}^{n} e^{\vartheta\left(a - \bar{c} - \nu q_i - \frac{b}{2}\sum_{j\neq i} q_j\right)q_i}\prod_{i<j}\left(1 + e^{\vartheta(\rho q_i q_j - \zeta)}\right)} \\
&= \frac{e^{\vartheta\sum_{i<j}^{n} a_{ij}(\rho q_i q_j - \zeta)}}{\prod_{i<j}\left(1 + e^{\vartheta(\rho q_i q_j - \zeta)}\right)} \\
&= \prod_{i<j}\frac{e^{\vartheta a_{ij}(\rho q_i q_j - \zeta)}}{1 + e^{\vartheta(\rho q_i q_j - \zeta)}} \\
&= \prod_{i<j}\left(\frac{e^{\vartheta(\rho q_i q_j - \zeta)}}{1 + e^{\vartheta(\rho q_i q_j - \zeta)}}\right)^{a_{ij}}\left(1 - \frac{e^{\vartheta(\rho q_i q_j - \zeta)}}{1 + e^{\vartheta(\rho q_i q_j - \zeta)}}\right)^{1-a_{ij}} \\
&= \prod_{i<j} p_{ij}^{a_{ij}}\left(1 - p_{ij}\right)^{1-a_{ij}}.
\end{aligned}$$

□

PROOF OF PROPOSITION 5. We first give a proof of part (ii) of the proposition. When $b = 0$ then we can write the potential function as

$$\Phi(\mathbf{q}, G) = \sum_{i=1}^{n} (n\eta q_i - \nu q_i^2) + \frac{\rho}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} q_i q_j - \zeta m.$$

Using the fact that the number of links in G can be written as $m = \frac{1}{2}\sum_{i=1}^{n} d_i = \frac{1}{2}\mathbf{u}^\top \mathbf{A}\mathbf{u}$, for $\mathbf{q} = q_0 \mathbf{u}$, we can write the potential as

$$\Phi(\mathbf{q}, G) = n^2 \eta q_0 - n\nu n^2 q_0^2 + \frac{1}{2}(\rho q_0^2 - \zeta) m.$$

From this expression we see that $\Phi(\mathbf{q}, G)$ is maximized for $G = K_n$ if $\rho q_0^2 > \zeta$ and $G = \bar{K}_n$ if $\rho q_0^2 < \zeta$. The phase transition from the empty to the complete graph that occurs at $\zeta^* = \rho q_0^2$ is shown in Figure 1 for different values of ϑ for $n = 10$ nodes.

Next, we prove part (i) of the proposition. Let $b = 0$ and assume that the output levels are fixed and given by $q_i \in [0, \bar{q}]$ for all $i = 1, \ldots, n$. Then we can write the potential as

$$\Phi(\mathbf{q}, G) = \sum_{i=1}^{n} n(\eta q_i - \nu q_i^2) + \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} (\rho q_i q_j - \zeta).$$

The second term in the above expression for $\Phi(\mathbf{q}, G)$ is a sum over positive terms with $a_{ij} = 1$ if $\rho q_i q_j > \zeta$ and negative otherwise. Hence, $\Phi(\mathbf{q}, G)$ is maximized if $a_{ij} = 1$ for all $\rho q_i q_j > \zeta$ and $a_{ij} = 0$ for all $\rho q_i q_j < \zeta$. □

PROOF OF PROPOSITION 6. We first analyze the partition sum $\mathscr{Z}_\vartheta$ in more detail. Note that

$$\begin{aligned}
\mathscr{Z}_\vartheta &= \int_{[0,\bar{q}]^n} \prod_{i=1}^{n} e^{\vartheta\left(n\eta - n\nu q_i - \frac{b}{2}\sum_{j\neq i} q_j\right) q_i} \prod_{i<j} \left(1 + e^{\vartheta(\rho q_i q_j - \zeta)}\right) d\mathbf{q} \\
&= \int_{[0,\bar{q}]^n} e^{\vartheta \sum_{i=1}^{n}\left(n\eta - n\nu q_i - \frac{b}{2}\sum_{j\neq i} q_j\right) q_i} e^{\sum_{i<j} \ln\left(1 + e^{\vartheta(\rho q_i q_j - \zeta)}\right)} d\mathbf{q}.
\end{aligned}$$

Next, we introduce the Hamiltonian defined by

$$\mathscr{H}(\mathbf{q}) \equiv \sum_{i=1}^{n} \left(n\eta q_i - n\nu q_i^2 + \sum_{j>i} \left(\frac{1}{\vartheta} \ln\left(1 + e^{\vartheta(\rho q_i q_j - \zeta)}\right) - b q_i q_j \right) \right),$$

so that $\sum_{G \in \mathcal{G}_n} e^{\Phi(\mathbf{q}, G)} = e^{\mathscr{H}(\mathbf{q})}$ (cf. Equation (51)). Then we can write the partition function as

$$\mathscr{Z}_\vartheta = \int_{[0,\bar{q}]^n} e^{\vartheta \mathscr{H}(\mathbf{q})} d\mathbf{q}.$$

In the following we make the Laplace approximation [cf. Wong, 2001]

$$\mathscr{Z}_\vartheta \approx \left(\frac{2\pi}{\vartheta}\right)^{\frac{n}{2}} \left| \left(\frac{\partial^2 \mathscr{H}}{\partial q_i \partial q_j}\right)_{q_i = q^*} \right|^{-\frac{1}{2}} e^{\mathscr{H}(\mathbf{q}^*)}, \tag{52}$$

where $\mathbf{q}^* = \text{argmax}_{\mathbf{q} \in [0,\bar{q}]^n} \mathscr{H}(\mathbf{q})$, and the Hessian is given by $\frac{\partial^2 \mathscr{H}}{\partial q_i \partial q_j}$ for $1 \leq i, j \leq n$. We have that

$$\frac{\partial \mathscr{H}(\mathbf{q})}{\partial q_i} = n\eta - 2n\nu q_i + \sum_{j \neq i} \left(\frac{\rho}{2}\left(1 + \tanh\left(\frac{\vartheta}{2}(\rho q_i q_j - \zeta)\right)\right) - b \right) q_j.$$

The first order conditions imply that

$$\eta = \frac{1}{n} \sum_{j \neq i} \left(b + 2\nu - \frac{\rho}{2}\left(1 + \tanh\left(\frac{\vartheta}{2}(\rho q_i q_j - \zeta)\right)\right) \right) q_j.$$

This system of equations has a symmetric solution, $q_i = q$ for all $i = 1, \dots, n$, where

$$(b + 2\nu)q - \eta = \frac{\rho}{2}\left(1 + \tanh\left(\frac{\vartheta}{2}\left(\rho q^2 - \zeta\right)\right)\right) q.$$

In the limit of $\vartheta \to \infty$ we obtain from the FOC of Equation (17) that

$$(b + 2\nu)q - \eta = \begin{cases} \rho q, & \text{if } \zeta < \rho q^2, \\ 0, & \text{if } \rho q^2 < \zeta. \end{cases}$$

This shows that the right hand side of Equation (17) has a point of discontinuity at $\sqrt{\frac{\zeta}{\rho}}$ (cf. Figure 2). It then follows that, in the limit of $\vartheta \to \infty$ (for the stochastically stable equilibrium), we have

$$q = \begin{cases} \frac{\eta}{b+2\nu-\rho}, & \text{if } \zeta < \frac{\rho\eta^2}{(b+2\nu)^2}, \\ \left\{\frac{\eta}{b+2\nu-\rho}, \frac{\eta}{\rho}\right\}, & \text{if } \frac{\rho\eta^2}{(b+2\nu)^2} < \zeta < \frac{\rho\eta^2}{(b+2\nu-\rho)^2}, \\ \frac{\eta}{b+2\nu}, & \text{if } \frac{\rho\eta^2}{(b+2\nu-\rho)^2} < \zeta, \end{cases} \tag{53}$$

which is increasing in ρ and η, and decreasing in ζ and b (cf. Figure 2). □

PROOF OF PROPOSITION 7. For notational simplicity, in the following we set $\nu = 0$. Observe that

$$\frac{\partial^2 \mathscr{H}}{\partial q_i \partial q_j} = \begin{cases} \frac{\vartheta\rho^2}{4}\sum_{j\neq i}\left(1 - \tanh\left(\frac{\vartheta}{2}\left(\rho q_i q_j - \zeta\right)\right)^2\right) q_j^2, & \text{if } i = j, \\ \frac{\rho}{2}\left(1 + \tanh\left(\frac{\vartheta}{2}\left(\rho q_i q_j - \zeta\right)\right)\right)\left(1 + \vartheta q_i q_j \frac{\rho}{2}\tanh\left(\frac{\vartheta}{2}\left(\rho q_i q_j - \zeta\right)\right)\right) - b, & \text{if } i \neq j. \end{cases}$$

In the symmetric equilibrium $q_i = q$ for all $i = 1, \dots, n$ this is

$$\frac{\partial^2 \mathscr{H}}{\partial q_i \partial q_j} = \begin{cases} \vartheta(n-1)(bq - \eta)(\eta - (b-\rho)q), & \text{if } i = j, \\ \vartheta(bq - \eta)(\eta - (b-\rho)q) - \frac{\eta}{q}, & \text{if } i \neq j. \end{cases}$$

Using the fact that

$$\begin{vmatrix} a & b & b & \dots \\ b & a & b & \dots \\ b & b & a & \\ \vdots & \vdots & & \ddots \end{vmatrix} = (a-b)^{n-1}(a + (n-1)b),$$

which is a special case of a circulant matrix and the determinant follows from the general formula [Horn and Johnson, 1990], we obtain

$$\left|\left(\frac{\partial^2 \mathscr{H}}{\partial q_i \partial q_j}\right)_{q_i = q}\right| = \frac{1}{q^n}\left(\vartheta(n-2)(bq - \eta)(\eta - (b-\rho)q)q + \eta\right)^{n-1}$$
$$\times\left(2\vartheta(n-1)(bq - \eta)(\eta - (b-\rho)q)q - (n-1)\eta\right).$$

In the symmetric case $q_i = q$ for all $i = 1, \dots, n$ the Laplace approximation of Equation (52) can be written as

$$\mathscr{Z}_\vartheta \approx \left(\frac{2\pi}{\vartheta}\right)^{\frac{n}{2}}\left|\left(\frac{\partial^2 \mathscr{H}}{\partial q_i \partial q_j}\right)_{q_i = q^*}\right|^{-\frac{1}{2}} e^{\vartheta n \mathscr{H}(q^*)},$$

where

$$\begin{aligned} \mathscr{H}(q) &\equiv n\eta q + (n-1)\left(\frac{1}{\vartheta}\ln\left(1 + e^{\vartheta(\rho q^2 - \zeta)}\right) - bq^2\right) \\ &\approx n^2\left(\eta q + \frac{1}{\vartheta}\ln\left(1 + e^{\vartheta(\rho q^2 - \zeta)}\right) - bq^2\right), \end{aligned}$$

for large n. We also introduce the free energy $\mathscr{F}_\vartheta \equiv -\ln \mathscr{Z}_\vartheta$, which allows us to write the expected number of links as follows

$$\mathbb{E}_\vartheta(m) = \sum_{G\in\mathcal{G}_n}\int_{[0,\bar{q}]^n} m\mu^\vartheta(\mathbf{q},G)d\mathbf{q} = \frac{1}{\mathscr{Z}_\vartheta}\sum_{G\in\mathcal{G}_n}\int_{[0,\bar{q}]^n} \underbrace{me^{\vartheta\Phi(\mathbf{q},G)}}_{-\frac{1}{\vartheta}\frac{\partial}{\partial\zeta}e^{\vartheta\Phi(\mathbf{q},G)}}\, d\mathbf{q}$$

$$= -\frac{1}{\vartheta}\frac{1}{\mathscr{Z}_\vartheta}\frac{\partial \mathscr{Z}_\vartheta}{\partial\zeta} = \frac{1}{\vartheta}\frac{\partial \mathscr{F}_\vartheta}{\partial\zeta}.$$

From the Laplace approximation of the partition function we find that

$$\frac{\partial \mathscr{F}_\vartheta}{\partial\zeta} \approx -\vartheta\frac{\partial\mathscr{H}(\mathbf{q})}{\partial\zeta} + \frac{1}{2}\mathrm{tr}\left(\left(\frac{\partial^2\mathscr{H}}{\partial q_i\partial q_j}\right)^{-1}\frac{\partial}{\partial\zeta}\left(\frac{\partial^2\mathscr{H}}{\partial q_i\partial q_j}\right)\right).$$

We further have that

$$\frac{\partial\mathscr{H}}{\partial\zeta} = -\frac{1}{2}\sum_{i=1}^n\sum_{j>i}^n\left(1+\tanh\left(\frac{\vartheta}{2}\left(\rho q_iq_j-\zeta\right)\right)\right),$$

and in the symmetric equilibrium this is

$$\left(\frac{\partial\mathscr{H}}{\partial\zeta}\right)_{q_i=q} = -\frac{n(n-1)}{2}\left(1+\tanh\left(\frac{\vartheta}{2}\left(\rho q^2-\zeta\right)\right)\right).$$

Moreover we find that

$$\frac{\partial^2\mathscr{H}}{\partial q_i^2} = \frac{\vartheta^2\rho^2}{4}\sum_{j\neq i}^n \tanh\left(\frac{\vartheta}{2}\left(\rho q_iq_j-\zeta\right)\right)\left(1-\tanh\left(\frac{\vartheta}{2}\left(\rho q_iq_j-\zeta\right)\right)^2\right)q_j^2,$$

and

$$\frac{\partial^2\mathscr{H}}{\partial q_i\partial q_j} = \frac{\vartheta\rho}{4}\left(1-\tanh\left(\frac{\vartheta}{2}\left(\rho q_iq_j-\zeta\right)\right)^2\right)\left(\vartheta\rho q_iq_j\tanh\left(\frac{\vartheta}{2}\left(\rho q_iq_j-\zeta\right)\right)-1\right).$$

Assuming symmetry this is

$$\left(\frac{\partial^2\mathscr{H}}{\partial q_i^2}\right)_{q_i=q} = \frac{\vartheta^2(n-1)}{\rho q}\left(2(bq-\eta)-\rho q\right)(bq-\eta)(\eta-(b-\rho)q),$$

and

$$\left(\frac{\partial^2\mathscr{H}}{\partial q_i\partial q_j}\right)_{q_i=q} = \frac{\vartheta}{\rho q^2}(bq-\eta)(\eta-(b-\rho)q)\left(2\vartheta q(bq-\eta)-\vartheta\rho q^2-1\right).$$

After some simplifications we then can write the expected number of links as follows

$$\mathbb{E}_\vartheta(m) \approx n(n-1)\frac{bq-\eta}{\rho q} + \frac{n}{2}\vartheta\frac{(2b-\rho)q-2\eta}{\rho q} = \frac{n(n-1)}{2}\left(1+\tanh\left(\frac{\vartheta}{2}\left(\rho q^2-\zeta\right)\right)\right)+O(n),$$

where q derives from Equation (17). Hence, the expected number of links is increasing in ρ, q and η, and decreasing in ζ and b (by reducing the equilibrium quantity q). Note that the above expression becomes exact as n becomes large. □

PROOF OF PROPOSITION 8. Let the degree distribution be given by $P(k)$ for $k = 0,\ldots,n-1$. Further, let $g(k|q)$ be the conditional probability that a firm with output level q has k links. Then the degree distribution can be written as follows [cf. Boguñá and Pastor-Satorras, 2003; Söderberg, 2002]

$$P(k) = \sum_{q\in\mathcal{Q}}\mu^\vartheta(q)g(k|q),$$

where $\mu^\vartheta(q)$ is the marginal probability distribution of firms with output levels $q \in \mathcal{Q}$. The average degree of a firm with output q is then given by

$$\bar{d}(q) = \sum_{k=0}^{n-1} k g(k|q),$$

and the average degree is given by

$$\bar{d} = \sum_{k=0}^{n-1} kP(k) = \sum_{q \in \mathcal{Q}} \mu^{\vartheta}(q)\bar{d}(q).$$

The probability that a firm with degree k has output q is given by Bayes' rule as

$$g(q|k) = \frac{\mu^{\vartheta}(q)g(k|q)}{P(k)}.$$

Let $X_{q,q'}(\omega)$ count the number of links between firms with output q and q' in a in a state $\omega \in \Omega$. We have that

$$X_{q,q'}(\omega) = \sum_{i=1}^{n}\sum_{j=i+1}^{n} \mathbb{1}_{\{q_i(\omega)=q\}}\mathbb{1}_{\{q_j(\omega)=q'\}}\mathbb{1}_{\{a_{ij}(\omega)=1\}}.$$

If X_q counts the number of links of a firm with output q then we can write $X_q = \sum_{q' \in \mathcal{Q}} X_{q,q'}$, and we have that $\mathbb{P}(X_q = k) = g(k|q)$. Observe that the random variables $X_{q,q'}$ are independent, and binomially distributed with

$$\mathbb{P}(X_{q,q'} = k) = \binom{n(q')}{k} p(q,q')^k (1 - p(q,q'))^{n(q')-k},$$

because the links between firms with output levels q and q' are independently drawn with probability $p(q,q')$. The generating function of the random variable $X_{q,q'}$ is given by[28]

$$\hat{g}(z|q,q') = (1 - (1-z)p(q,q'))^{n(q')}.$$

where $p(q,q')$ is given by Equation (41), and can be written as

$$p(q,q') = \frac{e^{\vartheta(\rho q q' - \zeta)}}{1 + e^{\vartheta(\rho q q' - \zeta)}}.$$

Since the random variable X_q is a sum of independent binomial random variables $X_{q,q'}$, we have that the generating function $\hat{g}(z|q)$ of X_q is the product of the generating functions $\hat{g}(z|q,q')$ of $X_{q,q'}$, so that after taking logs we get

$$\ln \hat{g}(z|q) = n \sum_{q' \in \mathcal{Q}} \mu^{\vartheta}(q') \ln\left(1 - (1-z)p(q,q')\right). \tag{54}$$

We then can solve this equation to obtain $\hat{g}(z|q)$, and from inverting the generating function

$$g(k|q) = \frac{1}{k!}\frac{d^k \hat{g}(0|q)}{dz^k}\bigg|_{z=0},$$

the degree distribution follows as

$$P(k) = \sum_{q \in \mathcal{Q}} g(k|q)\mu^{\vartheta}(q).$$

When the linking cost c is high, and the connection probability $p(q,q')$ is small, we can expanding Equation (54) as

$$\hat{g}(z|q) \approx e^{n\sum_{q' \in \mathcal{Q}} \mu^{\vartheta}(q')p(q,q')(z-1)}.$$

This is the generating function of a Poisson distribution with mean $\bar{d}(q) = n\sum_{q' \in \mathcal{Q}} \mu^{\vartheta}(q')p(q,q')$, which

[28]The probability generating function of a binomial random variable, with the number of successes in n trials, and probability p of success in each trial, is given by $g(z) = (1 - p(1-z))^n$.

means that

$$g(k|q) = \frac{1}{k!} e^{-\bar{d}(q)} \bar{d}(q)^k,$$

and the degree distribution is given by

$$P(k) = \frac{1}{k!} \sum_{q \in \mathcal{Q}} \mu^{\vartheta}(q) e^{-\bar{d}(q)} \bar{d}(q)^k,$$

while the average degree of a firm with output q is given by $\bar{d}(q) = n \sum_{q' \in \mathcal{Q}} \mu^{\vartheta}(q') p(q, q')$. This follows from

$$\begin{aligned}
\bar{d}(q) &= \left. \frac{d\hat{g}(z|q)}{dz} \right|_{z=1} \\
&= e^{n \sum_{q' \in \mathcal{Q}} \mu^{\vartheta}(q') \ln(1 - p(q,q')(1-z))} n \sum_{q'' \in \mathcal{Q}} \mu^{\vartheta}(q'') \left. \frac{p(q, q'')}{1 - p(q, q'')(1-z)} \right|_{z=1} \\
&= n \sum_{q' \in \mathcal{Q}} \mu^{\vartheta}(q') p(q, q').
\end{aligned}$$

Moreover, when the distribution $\mu^{\vartheta}(q) = \delta(q - q^*)$ is concentrated at an output level q^* then we can write the generating function of Equation (54) as

$$\hat{g}(z|q) = e^{n \ln(1 - p(q,q^*)(1-z))} = (1 - p(q, q^*)(1-z))^n,$$

which is the generating function of a binomially distributed random variable with success probability $p(q, q^*)$ in n trials. □

Proof of Proposition 10. The only network dependent part in $W(\mathbf{q}, G)$ is the potential function $\Phi(\mathbf{q}, G)$. For a given vector of outputs $\mathbf{q}$ the network that maximizes the potential is the threshold graph G where each link $ij \in G$ if and only if $\rho q_i q_j > \zeta$. Hence, we can write welfare reduced to this class of networks as follows

$$W(\mathbf{q}) = \eta \sum_{i=1}^{n} q_i + \frac{1-2b}{2} \left(\sum_{i=1}^{n} q_i \right)^2 - (n\nu - b) \sum_{i=1}^{n} q_i^2 + \sum_{i=1}^{n} \sum_{j \neq i} (\rho q_i q_j - \zeta) \mathbb{1}_{\{\rho q_i q_j > \zeta\}}.$$

From the form of $W(\mathbf{q})$ we see that we can distinguish two types of firms: those that are connected, and those that are not. Let $\mathcal{N}_1$ denote the set of the first, and $\mathcal{N}_2$ the set of the latter. The FOC for $i \in \mathcal{N}_1$ is given by

$$\frac{\partial W(\mathbf{q})}{\partial q_i} = \eta + (1-2b)(n_1 q_1 + n_2 q_2) - 2(n\nu - b) q_1 + 2\rho(n_1 - 1) q_1 = 0,$$

while the FOC for $i \in \mathcal{N}_2$ is

$$\frac{\partial W(\mathbf{q})}{\partial q_i} = \eta + (1-2b)(n_1 q_1 + n_2 q_2) - 2(n\nu - b) q_2 = 0.$$

The FOCs for all firms in the same set are identical, so their quantities must be identical too. We denote by q_1 the optimal quantity level of the firms in $\mathcal{N}_1$ and by q_2 the optimal quantity level of the firms in $\mathcal{N}_2$. Moreover, let $n_1 = |\mathcal{N}_1|$ and $n_2 = |\mathcal{N}_2| = n - n_1$. For $0 \leq q_1, q_2 \leq \bar{q}$ we then we have that

$$\begin{aligned}
q_1(n_1, n_2) &= \frac{\eta(b - n\nu)}{(b - n\nu)(2(n\nu - b) + n(2b-1)) + (n_1 - 1)(2(n\nu - b) + (2b-1)n_2)\rho}, \\
q_2(n_1, n_2) &= \frac{\eta(b - n\nu + (n_1 - 1)\rho)}{(b - n\nu)(2(n\nu - b) + n(2b-1)) + (n_1 - 1)(2(n\nu - b) + (2b-1)n_2)\rho},
\end{aligned} \tag{55}$$

and welfare can be written as as a function of $0 \leq n_1 \leq n$ (since $n_2 = n - n_1$) as follows

$$\begin{aligned}
W(n_1, n_2(n_1)) = \eta(n_1 q_1 + n_2 q_2) + \frac{1-2b}{2}(n_1 q_1 + n_2 q_2)^2 \\
- (n\nu - b)(n_1 q_1^2 + n_2 q_2^2) + n_1(n_1 - 1)(\rho q_1^2 - \zeta).
\end{aligned} \tag{56}$$

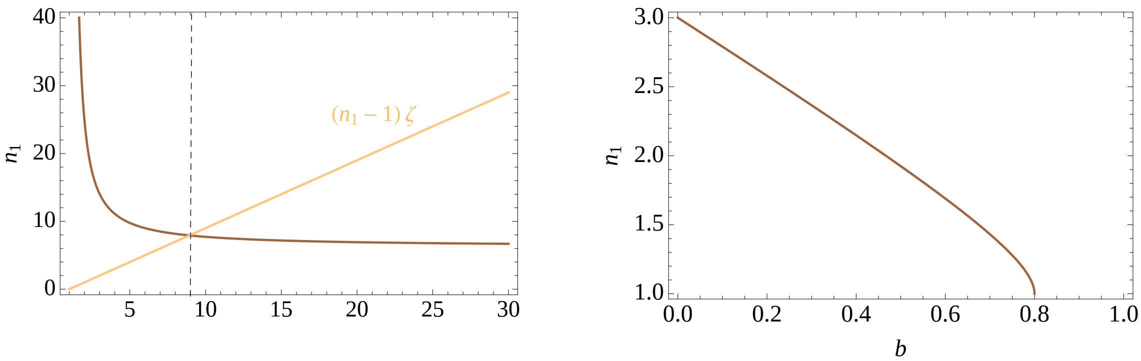

Figure 7: (Left panel) The optimal size n_1 solving Equation (57) when $q_2 = 0$ with $\rho = 0.3$ and $n\nu = 0.8$. (Right panel) The optimal size n_1 for $\zeta = 0$ from Equation (58) with $\rho = 0.4$, $n\nu = 0.5$ and $\eta = \sqrt{10}n$.

The above discussion can be summarized in the following proposition (see also Table 1).

In the following we discuss two special cases. First, for $q_2 = 0$ we find that

$$W(q_1) = \eta n_1 q_1 + \frac{1-2b}{2} n_1^2 q_1^2 - (n\nu - b) n_1^2 q_1^2 + n_1(n_1-1)(\rho q_1^2 - \zeta),$$

and from the FOC $\frac{\partial W(q_1)}{\partial n_1} = 0$ we obtain

$$q_1 = \frac{\eta}{(2(n\nu-\rho)-1)n_1 + 2\rho}.$$

Inserting into welfare gives

$$W(n_1) = \frac{n_1}{2}\left(\frac{\eta^2}{2\rho + (2(n\nu-\rho)-1)n_1} - 2(n_1-1)\zeta\right).$$

We take n_1 as a continuous variable, so that the FOC of $W(n_1)$ with respect to n_1 leads us to the condition

$$\frac{\eta^2\rho}{(2\rho + (2(n\nu-\rho)-1)n_1)^2} = (n_1-1)\zeta. \tag{57}$$

The optimal size n_1 solving this equality is illustrated in Figure 7. Next, for $\zeta = 0$ we obtain

$$W(n_1) = \frac{\eta^2(((b-n\nu)n + (n-n_1)(n_1-1)\rho}{2(b-n\nu)(2b(n-1)-n+2n\nu) + 2(n_1-1)(2b(n-n_1-1)-n-n_1+2n\nu)\rho}.$$

Note that in this case $q_2 = 0$. From the FOC $\frac{\partial W(q_1)}{\partial n_1} = 0$ we get

$$n_1 = \frac{1}{\rho}\left(\rho + n\nu - b + \sqrt{(n\nu-b)(n\nu-b+\rho)}\right). \tag{58}$$

The optimal size n_1 for $\zeta = 0$ is illustrated in Figure 7. □

Proof of Proposition 11. The potential $\phi(\mathbf{q}, G)$ of Equation (47) has the property that for any $q_i' \neq q_i \in [0, \bar{q}]$ we have that

$$\begin{aligned}\phi(q_i', \mathbf{q}_{-i}, G) - \phi(q_i, \mathbf{q}_{-i}, G) &= n\eta(q_i' - q_i) - n\nu(q_i'^2 - q_i^2) - b(q_i' - q_i)\sum_{j \neq i} q_j + \rho(q_i' - q_i)\sum_{j \in \mathcal{N}_i} q_j \\ &= \pi_i(q_i', \mathbf{q}_{-i}, G) - \pi_i(q_i, \mathbf{q}_{-i}, G).\end{aligned}$$

□

Proof of Proposition 12. Equation (17) can be written as

$$q_i - \frac{\rho}{2n\nu - b}\sum_{j=1}^{n} a_{ij}q_j = \frac{n\eta}{2n\nu - b} - \frac{b}{2n\nu - b}\|q\|$$

where $\|q\| = \mathbf{u}^\top \mathbf{q}$. Let us denote by $\varphi = \frac{\rho}{2n\nu-b}$, $A = \frac{n\eta}{2n\nu-b}$ and $B = \frac{b}{2n\nu-b}$. Then, in vector-matrix notation, the above equation can then be written as

$$(\mathbf{I}_n - \varphi\mathbf{q})\,\mathbf{q} = (A - B\|q\|)\mathbf{u}.$$

If $\varphi < \frac{1}{\lambda_{\max}}(\mathbf{q})$ then the matrix $\mathbf{I}_n - \varphi\mathbf{q}$ is invertible, and we obtain

$$\mathbf{q} = (A - B\|q\|)\,(\mathbf{I}_n - \varphi\mathbf{q})^{-1}\,\mathbf{u}.$$

Noting that

$$(\mathbf{I}_n - \varphi\mathbf{q})^{-1}\,\mathbf{u} = \mathbf{b}(G,\varphi),$$

where $\mathbf{b}(G,\varphi)$ is the vector of Bonacich centralities with parameter φ [Bonacich, 1987], we obtain

$$\mathbf{q} = (A - B\|\mathbf{q}\|)\mathbf{b}(G,\varphi).$$

With

$$\|\mathbf{q}\| = (A - B\|\mathbf{q}\|)\|b(G,\varphi)\|$$

we obtain

$$\|\mathbf{q}\| = \frac{A\|b(G,\varphi)\|}{1 + B\|b(G,\varphi)\|)}$$

and it follows that

$$\mathbf{q} = \frac{A}{1 + B\|\mathbf{b}(G,\varphi)\|)}\mathbf{b}(G,\varphi) = \frac{n\eta}{2\nu + b(\|\mathbf{b}(G,\varphi)\| - 1)}\mathbf{b}(G,\varphi).$$

Next, we compute equilibrium profits. Let us denote by $C = \frac{n\eta}{2n\nu + b(\|\mathbf{b}(G,\varphi)\| - 1)}$, so that $q_i = Ab_i(G,\varphi)$. Profit of firm i from Equation (8) can then be written as

$$\pi_i \;=\; n\eta C b_i(G,\varphi) - (n\nu - b)C^2 b_i(G,\varphi)^2 - bC^2\|\mathbf{b}(G,\varphi)\|b_i(G,\varphi) + \rho C^2 b_i(G,\varphi)\sum_{j=1}^{n} a_{ij}b_j(G,\varphi) - \zeta d_i.$$

Using the fact that

$$b_i(G,\varphi) = 1 + \frac{\rho}{2n\nu - b}\sum_{j=1}^{n} a_{ij}b_j(G,\varphi)$$

we obtain

$$\begin{aligned}\pi_i &= n\eta C b_i(G,\varphi) - (n\nu - b)C^2 b_i(G,\varphi)^2 - bC^2\|\mathbf{b}(G,\varphi)\|b_i(G,\varphi)\\ &\quad + \rho C^2(2n\nu - b)b_i(G,\varphi)(b_i(G,\varphi) - 1) - \zeta d_i.\\ &= n\nu C^2 b_i(G,\varphi)^2 - \zeta d_i.\end{aligned}$$

which gives

$$\pi_i = \frac{n^2\eta^2\nu}{(2n\nu + b(\|\mathbf{b}(G,\varphi)\| - 1))^2} b_i^2(G,\varphi) - \zeta d_i.$$

$\square$