R&D Networks: Theory, Empirics and Policy Implications[☆]

Michael D. König[a], Xiaodong Liu[b], Yves Zenou[c,d]

[a]*Department of Economics, University of Zurich, Schönberggasse 1, CH-8001 Zurich, Switzerland.*
[b]*Department of Economics, University of Colorado Boulder, Boulder, Colorado 80309–0256, United States.*
[c]*Department of Economics, Stockholm University, 106 91 Stockholm, Sweden.*
[d]*Research Institute of Industrial Economics (IFN), Box 55665, Stockholm, Sweden.*

Abstract

We study a structural model of R&D alliance networks where firms jointly form R&D collaborations to lower their production costs while competing on the product market. We derive the Nash equilibrium of this game, provide a welfare analysis and determine the optimal R&D subsidy program that maximizes total welfare. We also identify the key firms, i.e. the firms whose exit would reduce welfare the most. We then structurally estimate our model using a panel dataset of R&D collaborations and annual company reports. We use our estimates to identify the key firms and analyze the impact of R&D subsidy programs. Moreover, we analyze temporal changes in the rankings of key firms and how these changes affect the optimal R&D policy.

Key words: R&D networks, key firms, optimal subsidies
JEL: D85, L24, O33

[☆]We would like to thank Philippe Aghion, Nick Bloom, Chad Jones, Greg Crawford, Guido Cozzi, Stefan Bühler, Christian Helmers, Coralio Ballester, Matt Jackson, Michelle Sovinsky, Art Owen, Hang Hong, Marcel Fafchamps, Adam Szeidl, Bastian Westbrock, Fabrizio Zilibotti, Andrew F. Daughety, Jennifer Reiganum, Francis Bloch, Nikolas Tsakas, Ufuk Akcigit, Alfonso Gambardella and seminar participants at the University of Zurich, University of St.Gallen, Utrecht University, Stanford University, University College London, University of Washington, the PEPA/cemmap workshop on Microeconomic Applications of Social Networks Analysis, the Public Economic Theory Conference, the IZA Workshop on Social Networks in Bonn and the CEPR Workshop on Moving to the Innovation Frontier in Vienna for their helpful comments. Michael D. König 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. Yves Zenou acknowledges financial support from the Swedish Research Council (Vetenskaprådet) through research grant 421–2010–1310.

Email addresses: `michael.koenig@econ.uzh.ch` (Michael D. König), `xiaodong.liu@colorado.edu` (Xiaodong Liu), `yves.zenou@ne.su.se` (Yves Zenou)

1. Introduction

R&D partnerships have become a widespread phenomenon characterizing technological dynamics, especially in industries with a rapid technological development such as, for instance, the pharmaceutical, chemical and computer industries [cf. Ahuja, 2000; Hagedoorn, 2002; Powell et al., 2005; Riccaboni and Pammolli, 2002; Roijakkers and Hagedoorn, 2006]. In those industries, firms have become more specialized in specific domains of a technology and they tend to combine their knowledge with the knowledge of other firms that are specialized in different technological domains [Powell et al., 1996; Weitzman, 1998]. The increasing importance of R&D collaborations has spurred research for theoretical models studying these relationships, and for empirical tests of these models.

In this paper, we consider a general model of competition à la Cournot, where firms choose both R&D expenditures and output levels. Firms can reduce their costs of production by investing in R&D as well as by establishing R&D collaborations with other firms. An important – and realistic – innovation of our framework is to study the equilibrium outcomes in which firms can establish R&D collaborations with both competing firms in their own sector and firms in other sectors. In this model, R&D collaborations can be represented by a network. This allows us to write the profit function of each firm as a function of two matrices, $\mathbf{A}$ and $\mathbf{B}$, where $\mathbf{A}$ is the adjacency matrix of the network capturing all direct R&D collaborations, while $\mathbf{B}$ is a competition matrix that keeps track of which firm is in competition with which other firm in the same product market. Due to these two matrices and thus, these two opposing effects of technology spillovers and competition, all firms indirectly interact with all other firms. To illustrate this point, consider, for example, the car manufacturing sector. The price of a car is determined by the demand for cars and the competition between other car producing firms. However, when these firms do not only have R&D collaborations with other car manufacturing firms but also with firms from other sectors, the price of cars will also be indirectly influenced by firms from other industries.

We characterize the Nash equilibrium of our model for any type of R&D collaboration network (i.e. any matrix $\mathbf{A}$) as well as for any type of competition structure between firms (i.e. any matrix $\mathbf{B}$). We show that there exists a key trade off faced by firms between the *technology (or knowledge) spillover effect* of R&D and the *product rivalry effect* of competition. The former effect captures the *positive* impact of R&D collaborations on output and profits (through the matrix $\mathbf{A}$) while the latter captures the *negative* impact of competition and market stealing effects (through the matrix $\mathbf{B}$).

We show that the Nash equilibrium can be characterized by the fact that firms produce their goods proportionally to their Katz-Bonacich centrality, a well-known measure in the sociology literature that determines how central each firm is in the network, and also the degree of competition in the product market. In particular, a very central firm in the network will not always produce the highest output because the optimal output choice will also depend on the competition intensity the firm faces in the product market.

We also provide a welfare analysis with an explicit expression for total welfare as a function of the fundamental parameters of the model. We further provide a lower and an upper bound on the welfare function with bounds that depend on the parameters as well as the topology of the network. Moreover, we study the problem of optimal *network design* where we show which network is the most efficient (i.e. the one maximizing welfare determined by producer and consumer surplus among all possible networks). Then, we study two important policies. First, our equilibrium characterization allows us to define the *key firms*, i.e. the firms whose exit would reduce welfare the most. These are the systemically relevant firms for industry productivity and performance. Second, we study subsidy policies where the planner can subsidize the R&D effort of each firm. In both cases, we are able to derive an exact formula for any type of network and competition structure that determines which the key firm is and the amount of subsidy that should be given to each of them.

Then, we bring the model to the data by using a panel of R&D collaborations and annual company reports over different sectors and years. We estimate the first-order conditions of the model by testing the trade-off for firms between the *technology (or knowledge) spillover effect* of R&D and the *product rivalry effect* of competition mentioned above. In terms of identification strategy, we use firm and time fixed effects (as we have a panel of firms), an IV strategy and an estimation of a network formation model. As predicted by the theoretical model, we find that the spillover effect has a *positive* and significant impact on output and profit while the competition effect has a *negative* and significant impact. We also show that the net effect of collaboration is positive.

Following our theoretical results, we empirically rank the *key firms* in terms of their contribution to welfare across different sectors and countries. In particular, in our analysis of the key firms, we quantify theoretically and empirically the highest welfare loss incurred due to the exit of a firm. Observe that we determine the key firms for each year during the whole period (1950-2006). As a result, key firms can differ from year to year and our key-player policy is mainly a *short-run policy analysis* in which the network does not change during the given year the key firms are calculated. We believe that our results could thus help guide policy makers in evaluating how much it would be worth bailing out a particular firm in a given stage of the industry evolution.

We also perform the same analysis for R&D subsidies. We further analyze the temporal changes of the rankings of key firms and subsidized firms. In particular, we show that the key firms are not always the most central ones by any conventional measure. In other words, the key firms are not always those with the largest number of R&D collaborations, nor the highest eigenvector, betweenness or closeness centrality. More importantly, we also show that the key firms are not those with the highest market share in their industry. For example, we find that *General Motors* is a key firm, but it does not have the highest market share in its sector, since it "only" has 12.14 % of the market share while, for example, *Hitachi*, *Altria* or *Pepsico* have a much higher share (up to 50 %), but are not the top key firms. This means that it is not straightforward to determine which firm should be "targeted" in the network by only observing its market share, size or even its position in the network.

Interestingly, in late 2008, *General Motors* (along with *Chrysler*) told America that it was in danger of folding. George W. Bush agreed to a temporary bailout, but handed the auto company's long-term future over to his successor, President-Elect Barack Obama. Obama then shepherded a comprehensive bailout of *General Motors* that allowed it to stay in business but imposed numerous conditions that, it was hoped, would secure its viability and allow the companies to eventually return to profitability. Even though it was very controversial, it turns out that the bailout of *General Motors* was a success since, with profits of 3.2 billion U.S. dollars, the first quarter of 2011 was *General Motors*'s best performance in ten years and its fifth-consecutive profitable quarter. In our ranking of key firms, *General Motors* ranks first in 1990 and fourth in 2005. If *General Motors* had gone bankrupt in 1990 and exited the market, the loss of total welfare for both firms and consumers would have been as high as 8.14 %. Our results indicate that *General Motors* is a key firm, not only because it is a large company but also because it connects other firms to each other in an important way, and our analysis thus brings conditional support to the recent government intervention program in the automobile manufacturing sector in the United States.

The rest of the paper is organized as follows. In the next section, we compare our contribution to the existing literature. In Section 3, we develop a model of firms competing in the product market with technology sharing R&D collaborations that allow them to reduce their production costs. We characterize the Nash equilibrium of this game and show under which conditions it exists, is unique and interior. Section 4 determines welfare and investigates the optimal network structure of R&D collaborations. Section 5 introduces the definition and computation of the key firms while Section 6 discusses optimal R&D subsidies. Section 7 describes the data. Section 8 is divided into two parts. In Section 8.1, we define the econometric specification of our model while, in Section 8.2, we highlight our identification strategy. The empirical results are given in Section 9. The policy results of our empirical analysis are given in Section 10 where the key player analysis can be found in Section 10.1 while that of the subsidy analysis appears in Section 10.2. Finally, Section 11 concludes the paper. The network definitions and characterizations used throughout the paper are given in Appendix A, an analysis in terms of Bertrand competition is performed in Appendix B and some additional results for welfare are to be found in the Appendix C. In Appendix D, we provide a theoretical model of intra and interindustry collaborations. All proofs can be found in Appendix F.

2. Related Literature

Our paper lies at the intersection of different strands of the literature. We would like to expose them in order to highlight our contribution.

Network Theory Our theoretical model analyzes a game with strategic complementarities where firms decide about output and R&D effort by taking the network as given. Thus, it belongs to the class of games known as *games on networks* [cf. Jackson and Zenou, 2015].[1] Compared to this literature, where a prominent paper is that of Ballester et al. [2006],[2] we re-interpret their model in terms of R&D networks and extend their framework to account for competition between firms not only within the same product market but also between different product markets (see our Proposition 1). This yields very general results that can encompass any possible network of collaborations and any possible interaction structure of competition between firms. We also provide an explicit welfare characterization, provide lower and upper bounds and determine which network that maximizes total welfare (see Propositions 2, 3 and 4). To the best of our knowledge, this is one of the first papers to provide such an analysis.[3] We also provide two policy analyses. The first consists of subsidizing firms' R&D efforts. We are able to determine the optimal subsidy level both when it is homogenous (Proposition 6) and when it is targeted to firms (Proposition 7). We are not aware of any other studies of subsidy policies in the context of networks.[4] Finally, we extend the key player analysis proposed by Ballester et al. [2006]. In their paper, they define a key player in the context of crime where the removal of the key player generates the highest reduction of crime in the network. In our context of R&D networks, we define a key firm as the one that would reduce total welfare the most if it were removed. It is a different notion since the key firms are those whose disappearance from the market would result in a dramatic total welfare loss. Thus, we generalize the inter-centrality formula proposed in Ballester et al. [2006] by having both network and competition effects defining the key player (see Proposition 5).

Theoretical Studies of R&D Collaboration Networks In the industrial organization literature, there is a long tradition of models that analyze product and price competition with R&D collaborations, first pioneered by Arrow [1962] and then pursued by Spence [1984]. One of their main insights is that the incentives to invest in R&D are reduced by the presence of such technology spillovers. This raised the interest in R&D cooperation as a means of internalizing spillovers. More recently, the seminal works by D'Aspremont and Jacquemin [1988] and Suzumura [1992], Kamien et al. [1992] focus on the direct links between firms in the R&D collaboration process.

In all of this literature, however, there is no explicit network of R&D collaborations.

[1]The economics of networks is a growing field. For overviews of the literature, see Vega-Redondo [2007], Goyal [2007], Jackson [2008], De Martí and Zenou [2011], Jackson and Zenou [2013, 2015], Zenou [2015a].

[2]See also Bramoullé et al. [2014].

[3]In a recent paper, Belhaj et al. [2013] study network design in a game on networks with strategic complements, but without competition effects.

[4]There are some papers that look at subsidies in industries with R&D collaborations but the network is not explicitly modeled. See e.g. Acemoglu et al. [2012]; Bagwell and Staiger [1994]; Bloom et al. [2002]; Hinloopen [2001]; Impullitti [2010]; Leahy and Neary [1997]; Qiu and Tao [1998]; Song and Vannetelbosch [2007]; Spencer and Brander [1983].

The first paper that provides an explicit analysis of R&D networks is that by Goyal and Moraga-Gonzalez [2001].[5] The authors introduce a strategic Cournot oligopoly game in the presence of externalities induced by a network of R&D collaborations. Benefits arise in these collaborations from sharing knowledge about a cost-reducing technology. However, by forming collaborations, firms also change their own competitive position in the market as well as the overall market structure. Thus, there exists a 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. Westbrock [2010] extends their framework to analyze welfare and inequality in R&D collaboration networks, but abstracts from R&D investment decisions.

However, these papers typically provide results only for a small number of firms or specific networks, such as regular networks (i.e. all firms have the same number of R&D collaborations), star-shaped or complete networks, networks that we typically do not observe in the real-world. Compared to these papers, we provide results for all possible networks with an arbitrary number of firms and a complete characterization of equilibrium output and R&D effort choices in multiple interdependent markets. We also determine policies related to network design, the identification of the key player and optimal R&D subsidies.

Econometrics of Networks There has recently been a significant progress in the literature on identification and estimation of social network models (see Blume et al. [2011], for a recent survey). In his seminal work, Manski [1993] introduces a linear-in-means social interaction model with endogenous effects, contextual effects, and correlated effects. Manski shows that the linear-in-means specification suffers from the "reflection problem" and the different social interaction effects cannot be separately identified. Bramoullé et al. [2009] generalize Manski's linear-in-means model to a general local-average social network model, whereas the endogenous effect is represented by the average outcome of the peers. They provide some general conditions for the identification of the local-average model using the characteristics of an indirect connection as an instrument for the endogenous effect assuming that the network (and its adjacency matrix) is exogenous. However, if the adjacency matrix is endogenous, i.e., if there exists some unobservable factor that could affect both the link formation and the outcome, then the above identification strategy will fail. Here, as we have a panel data where the network changes over time (whereas in many applications, the network is observed at one point in time; [see e.g. Bramoullé et al., 2009; Calvó-Armengol et al., 2009]), we adopt a similar identification strategy using instruments but with both *firm and time fixed effects* to attenuate the potential endogeneity of the adjacency matrix. Then, we go even further by explicitly modeling the network formation process of R&D collaborations. Indeed, we add a first stage, where we explain an R&D collaboration between two firms i and j at time t by whether these two firms had an R&D collaboration in the past, whether they are

[5] See also Dawid and Hellmann [2012] and Goyal and Joshi [2003].

technologically close in terms of patents and whether they are in the same industry. Then, we carry out our instrumental variable (IV) estimation strategy described above using the *predicted adjacency matrix* derived from the first stage and compare our results to the ones with the *observed adjacency matrix*.

Empirical Studies of Technology Spillovers and R&D Collaborations There is a large empirical literature on technology spillovers [see e.g. Bloom et al., 2013; Einiö, 2013; Griffith et al. , 2006; Jones, 1998]. Besides, there is also a large number of empirical papers on R&D collaboration networks, which are mostly descriptive [see e.g. Fleming, 2007; Hanaki et al., 2010; Powell et al., 2005; Rosenkopf and Schilling, 2007]. Compared to these two literatures, we explicitly model the network of R&D collaborations, structurally estimate our model and derive policy implications.

To illustrate our contribution, let us consider a prominent paper within the first strand of literature, namely that by Bloom et al. [2013]. This paper highlights the key trade-off faced by firms between the *technology (or knowledge) spillover effect* of R&D and the *product rivalry effect*. The former effect captures the *positive* impact of R&D collaborations on output and profit while the latter captures the *negative* impact of competition. The authors first provide different "distance" measures between firms to capture technology spillovers and then test the impact of these two effects on output and profits of firms. They show that the net effect of R&D is positive so that the former dominates the latter effect. In our analysis, we can directly measure the technological spillovers between two firms through the presence of an R&D collaboration between them. Within this framework, we further provide a theoretical model of R&D collaboration networks that incorporates the trade off between the knowledge spillover effect and the product rivalry effect. We structurally estimate our theoretical model using the CATI alliance database and Compustat data and show that the net effect of R&D collaborations is positive. Using our estimates, we empirically apply our model to analyze subsidy and key player policies and provide a ranking of the top 25 firms. We believe that this is the first empirical paper to provide such a ranking based on these two types of R&D policies.

However, it should be clear that R&D spillovers do not only take place through R&D collaborations. For example, in Bloom et al. [2013], none of this is through R&D collaborations, and this paper and others in this literature point to important spillovers through building on the shoulders of giants or technological neighborhood. In particular, R&D spillovers could extend to (i) reading other firms' patents; (ii) imitating their products; and (iii) hiring their employees. As a result, using our dataset and that of Bloom et al. [2013], we extend both our theoretical and empirical analysis by allowing for direct (R&D collaborations between firms) and indirect weighted technology spillovers (between non-collaborating firms), where weights characterize alternative channels for technology spillovers than R&D collaborations (representing technological proximity).

The Key-Player Problem The problem of identifying key players in a network has a long history, at least in the sociology literature. Indeed, one of the focuses of this literature is to propose different measures of network centrality and to assert the descriptive and/or prescriptive suitability of each of these measures to different situations [see, in particular Wasserman and Faust, 1994]. Borgatti [2003, 2006] was among the first to investigate the issue of identifying key players, which is based on explicitly measuring the contribution of a set of agents to the cohesion of a network. The basic strategy is to take any network property, such as density or maximum flow, and derive a centrality measure by deleting nodes and measuring the change in the network property. Borgatti measures the amount of *reduction* in cohesiveness of the network that would occur if some nodes were not present.

Ballester et al. [2006, 2010] were the first to define the key-player problem in terms of the *behavior* of agents and the total activity is measured as the sum of efforts of all agents at the Nash equilibrium. As stated above, from a theoretical viewpoint, we extend their intercentrality measure of the key player by looking at welfare loss instead of total activity (output) loss and by including both the network spillover and the competition effect. In our context, a key firm can help measure the fragility of the system, since, if it disappears from the economy, welfare reduction will be the highest among all other possible firms. It has to be clear, however, that it is a short-run policy since the network is taken as given.

To the best of our knowledge, there are only two other papers that have empirically tested the key player policy but for crime. Liu et al. [2012] test the key player policy for juvenile crime in the United States, while Lindquist and Zenou [2013] identify key players for co-offending networks in Sweden.[6] We are the first to test the key player policy for R&D networks and propose a ranking of firms according to their intercentrality measures. We also consider another policy which consists of subsidizing the R&D expenditures of firms so that total welfare is maximized. In the empirical section, we also compare the ranking of firms in terms of the key player and the subsidy policies.

3. The Model

We consider a general Cournot oligopoly game where a set $\mathcal{N} = \{1, \ldots, n\}$ of firms is partitioned in M heterogeneous product markets.[7] We also allow for consumption goods to be imperfect substitutes (and thus differentiated products) by adopting the consumer utility maximization approach of Singh and Vives [1984]. We first consider the demand q_i for the good produced by firm i in market $\mathcal{M}_m$, $m = 1, \ldots, M$. A representative consumer in market

[6]For an overview of the key-player literature, see Zenou [2015b].

[7]In the empirical analysis in Section 7, we measure the market where each firm operates by the Standard Industrial Classification (SIC), which classifies industries by a four-digit code. As a result, a market corresponds to a particular industry or sector.

$\mathcal{M}_m$ obtains the following gross utility from consumption of the goods $(q_i)_{i\in\mathcal{M}_m}$

$$\bar{U}_m((q_i)_{i\in\mathcal{M}_m}) = \alpha_m \sum_{i\in\mathcal{M}_m} q_i - \frac{1}{2}\sum_{i\in\mathcal{M}_m} q_i^2 - \frac{\rho}{2}\sum_{i\in\mathcal{M}_m}\sum_{j\in\mathcal{M}_m, j\neq i} q_i q_j.$$

In this formulation, the parameter α_m captures the market size or the heterogeneity in products, whereas $\rho \in (0,1]$ measures the degree of substitutability between products. In particular, $\rho \to 1$ depicts a market of perfectly substitutable goods, while $\rho \to 0$ represents the case of local monopolies.

The consumer maximizes net utility $U_m = \bar{U}_m - \sum_{i\in\mathcal{M}_m} p_i q_i$, where p_i is the price of good i. This gives the inverse demand function for firm i

$$p_i = \bar{\alpha}_i - q_i - \rho \sum_{\substack{j\in\mathcal{M}_m,\\ j\neq i}} q_j, \tag{1}$$

where $\bar{\alpha}_i = \sum_{m=1}^{M} \alpha_m \mathbb{1}_{\{i\in\mathcal{M}_m\}}$. In the model, we will study both the general case where $\rho > 0$ but also the special case where $\rho = 0$. The latter case is when firms are local monopolies so that the price of the good produced by each firm i is only determined by its quantity q_i (and the size of the market) and not by the quantities of other firms, i.e. $p_i = \bar{\alpha}_i - q_i$.

Firms can reduce their production costs by investing in R&D as well as by establishing an R&D collaboration with another firm. The amount of this cost reduction depends on the R&D effort e_i of firm i and the R&D efforts of the firms that are collaborating with i, i.e., R&D collaboration partners.[8] Given the effort level $e_i \in \mathbb{R}_+$, the marginal cost c_i of firm i is given by[9,10]

$$c_i = \bar{c}_i - e_i - \varphi \sum_{j=1}^{n} a_{ij} e_j, \tag{2}$$

The network G is captured by $\mathbf{A}$, which is a symmetric $n \times n$ *adjacency matrix*. Its element $a_{ij} \in \{0,1\}$ indicates if there exists a link between nodes i and j and zero otherwise.[11] In the context of our model, $a_{ij} = 1$ if firms i and j set up an R&D collaboration (0 otherwise) and $a_{ii} = 0$. In Equation (2), the total cost reduction for firm i stems from its own research effort e_i and the research knowledge of other firms, i.e., *knowledge spillovers*, which is captured by the term $\sum_{j=1}^{n} a_{ij} e_j$, where $\varphi \geq 0$ is the marginal cost reduction due to neighbor's effort.[12]

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

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

[10] Throughout the paper we assume that the cost $\bar{c}_i$ is large enough such that marginal costs, c_i, are always positive for all firms $i \in \mathcal{N}$.

[11] See Appendix A.1 for more definitions and characterization of networks.

[12] In Equation (42) in Appendix E we present an extension of the model where firms benefit from both, direct technology spillovers between collaborating firms and indirect technology spillovers between non-collaborating firms. It is therefore important to note that we can define with $\mathbf{A}$ a more general matrix that captures

We assume that R&D effort is costly. In particular, the cost of R&D effort is an increasing function, exhibits decreasing returns, and is given by $\frac{1}{2}e_i^2$. Firm i's profit is then given by

$$\pi_i = (p_i - c_i)q_i - \frac{1}{2}e_i^2. \tag{3}$$

Inserting marginal cost from Equation (2) and inverse demand from Equation (1) into Equation (3) gives

$$\begin{aligned}\pi_i &= (\bar{\alpha}_i - q_i - \rho \sum_{j\in\mathcal{M}_m, j\neq i} q_j - \bar{c}_i + e_i + \varphi \sum_{j=1}^n a_{ij}e_j)q_i - \frac{1}{2}e_i^2 \\ &= (\bar{\alpha}_i - \bar{c}_i)q_i - q_i^2 - \rho\sum_{j=1}^n b_{ij}q_iq_j + q_ie_i + \varphi q_i \sum_{j=1}^n a_{ij}e_j - \frac{1}{2}e_i^2,\end{aligned} \tag{4}$$

where $b_{ij} \in \{0,1\}$ indicates whether firms i and j operate in the same market or not, and let $\mathbf{B}$ be the $n\times n$ matrix whose ij-th element is b_{ij}. In Equation (4), we have that $\sum_{j\in\mathcal{M}_m, j\neq i} q_j = \sum_{j=1}^n b_{ij}q_j$ since $b_{ij} = 1$ if $i,j\in\mathcal{M}_m$ and $i\neq j$, and $b_{ij} = 0$ otherwise, i.e. if i and j do not belong to the same market. In other words, the matrix $\mathbf{B}$ captures which firms operate in the same market and which firms do not. Take row i in matrix $\mathbf{B}$, for example. If there are only zeros, this means that firm i is alone in its market. If there is a 1 corresponding to column j, this means that firms i and j operate in the same market (or sector).

In the following, we consider quantity competition among firms à la Cournot.[13] The next proposition establishes the Nash equilibrium where each firm i simultaneously chooses *both* its quantities q_i and its R&D effort e_i in a given network of R&D collaborations.[14]

Proposition 1. *Consider the n–player simultaneous move game with payoffs given by Equation (4) and strategy space in $\mathbb{R}^n_+ \times \mathbb{R}^n_+$. Denote by $\mu_i \equiv \bar{\alpha}_i - \bar{c}_i$ for all $i\in\mathcal{N}$, $\boldsymbol{\mu}$ the corresponding $n\times 1$ vector, $\phi \equiv \varphi/(1-\rho)$, $|\mathcal{M}_m|$ the size of the largest market, $\mathbf{I}_n$ the $n\times n$ identity matrix, $\mathbf{u}$ the $(n\times 1)$ vector of ones and $\lambda_{PF}(\mathbf{A})$ the largest eigenvalue of $\mathbf{A}$. Denote also by $\underline{\mu} = \max_i\{\mu_i \mid i\in\mathcal{N}\}$ and $\overline{\mu} = \max_i\{\mu_i \mid i\in\mathcal{N}\}$, with $0<\underline{\mu}<\overline{\mu}$.*

(i) If

$$\rho + \varphi < \left(\max\left\{\lambda_{PF}(\mathbf{A}), \max_{m=1,\dots,M}\{(|\mathcal{M}_m|-1)\}\right\}\right)^{-1} \tag{5}$$

and

$$\rho \max_{m=1,\dots,M}\{(|\mathcal{M}_m|-1)\} < 1 - \varphi\lambda_{PF}(\mathbf{A}), \tag{6}$$

potential technology spillovers between firms.

[13] In Appendix B we show that the same functional forms for best response quantities and efforts can be obtained for price setting firms under Bertrand competition as we find them in the case of Cournot competition.

[14] See Appendix A.4 for a precise definition of the Bonacich centrality used in the proposition.

hold, there exists a unique interior Nash equilibrium with output levels given by

$$\mathbf{q} = (\mathbf{I}_n + \rho\mathbf{B} - \varphi\mathbf{A})^{-1}\boldsymbol{\mu}. \tag{7}$$

(ii) Assume that there exists only a single market so that $M = 1$. Let the $\boldsymbol{\mu}$-weighted Bonacich centrality be given by $\mathbf{b}_{\boldsymbol{\mu}}(G, \phi) \equiv (\mathbf{I}_n - \phi\mathbf{A})^{-1}\boldsymbol{\mu}$. If

$$\phi\lambda_{PF}(\mathbf{A}) + \frac{n\rho}{1-\rho}\left(\frac{\overline{\mu}}{\underline{\mu}} - 1\right) < 1, \tag{8}$$

holds, then there exists a unique interior Nash equilibrium with output levels given by

$$\mathbf{q} = \left(\frac{1}{1-\rho}\right)\left[\mathbf{b}_{\boldsymbol{\mu}}(G,\phi) - \frac{\rho\left\|\mathbf{b}_{\boldsymbol{\mu}}(G,\phi)\right\|_1}{1-\rho+\rho\left\|\mathbf{b}_{\mathbf{u}}(G,\phi)\right\|_1}\mathbf{b}_{\mathbf{u}}(G,\phi)\right]. \tag{9}$$

(iii) Assume a single market (i.e., $M = 1$) and that $\mu_i = \mu$ for all $i \in \mathcal{N}$. If $\phi\lambda_{PF}(\mathbf{A}) < 1$, then there exists a unique interior Nash equilibrium with output levels given by

$$\mathbf{q} = \frac{\mu}{1-\rho+\rho\|\mathbf{b}_{\mathbf{u}}(G,\phi)\|_1}\mathbf{b}_{\mathbf{u}}(G,\phi). \tag{10}$$

(iv) Assume a single market (i.e., $M = 1$), $\mu_i = \mu$ for all $i \in \mathcal{N}$ and that goods are non-substitutable (i.e., $\rho = 0$). If $\varphi < \lambda_{PF}(\mathbf{A})^{-1}$, then the unique equilibrium quantities are given by $\boldsymbol{q} = \mu\boldsymbol{b}_{\boldsymbol{u}}(G,\varphi)$.

(v) Let $\mathbf{q}$ be the unique Nash equilibrium quantities in any of the above cases (i) to (iv), then for all $i \in \mathcal{N} = \{1, \ldots, n\}$ the equilibrium profits are given by

$$\pi_i = \frac{1}{2}q_i^2, \tag{11}$$

and the equilibrium efforts are given by

$$e_i = q_i. \tag{12}$$

This proposition gives the results of the Nash equilibrium starting from the most general case where firms can operate and have links in any market (case (i)) to the case where all firms operate in the same market (case (ii)) and where they have the same fixed cost of production and no product heterogeneity (case (iii)) and, finally, when, on top of that, goods are not substitutable (case (iv)). Indeed, it is easily verified (see Appendix F; proof of Proposition 1) that the first-order condition with respect to R&D effort e_i is given by Equation (12),[15]

[15] The proportional relationship between R&D effort levels and output in Equation (12) has been confirmed in a number of empirical studies [see e.g. Cohen and Klepper, 1996a,b; Klette and Kortum, 2004].

while the first-order condition with respect to quantity q_i leads to

$$q_i = \mu_i - \rho \sum_{j=1}^{n} b_{ij} q_j + \varphi \sum_{j=1}^{n} a_{ij} q_j, \tag{13}$$

or, in matrix form, $\mathbf{q} = \boldsymbol{\mu} - \rho \mathbf{B}\mathbf{q} + \varphi \mathbf{A}\mathbf{q}$. In terms of the literature on games on networks [Jackson and Zenou, 2015], this proposition generalizes the results of Ballester et al. [2006] and Calvó-Armengol et al. [2009] for the case of local competition in different markets and choices of both effort and quantity. This proposition provides a total characterization of an interior Nash equilibrium as well as its existence and uniqueness in a very general framework when different markets and different products are considered. If we consider case (i), the new conditions are Equations (5) and (6), which guarantee the existence, uniqueness and interiority of the Nash equilibrium solutions in the most general case. In case (ii) where all firms operate in the same market, in order to obtain a unique interior solution, only the condition in Equation (8) is required, which generalizes the usual condition $\phi\lambda_{\mathrm{PF}}(\mathbf{A}) < 1$ given, for example, in Ballester et al. [2006]. In fact, the condition in Equation (8) imposes a more stringent requirement on $\rho, \varphi, \mathbf{A}$ as the left-hand side of the inequality is now augmented by $\frac{n\rho}{1-\rho}\left(\frac{\overline{\mu}}{\underline{\mu}} - 1\right) \geq 0$. That is, everything else equal, the higher the discrepancy $\overline{\mu}/\underline{\mu}$ of marginal payoffs at the origin, the lower is the level of network complementarities $\phi\lambda_{\mathrm{PF}}(\mathbf{A})$ that are compatible with a unique and interior Nash equilibrium.

More generally, the key insight of Proposition 1 is the interaction between the *network effect*, through the adjacency matrix $\mathbf{A}$, and the *market effect*, through the competition matrix $\mathbf{B}$ and that is why the first-order condition with respect to q_i given by Equation (13) takes both of them into account. To better understand this result, consider the following simple example of an industry composed of three firms and two sectors, $\mathcal{M}_1$ and $\mathcal{M}_2$, where firms 1 and 2, as well as firms 1 and 3 have an R&D collaboration, while firms 1 and 2 operate in the same market $\mathcal{M}_1$ (see Figure 1).

Then, the adjacency matrix $\mathbf{A}$ and the competition matrix $\mathbf{B}$ are given by

$$\mathbf{A} = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \qquad \mathbf{B} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Assume that firms are homogeneous such that $\mu_i = \mu$ for $i = 1, 2, 3$. Using Proposition 1, the equilibrium output is given by

$$\mathbf{q} = \mu(\mathbf{I} - \varphi\mathbf{A} + \rho\mathbf{B})^{-1}\mathbf{u} = \frac{\mu}{1 - 2\varphi^2 + 2\varphi\rho - \rho^2} \begin{pmatrix} 1 + 2\varphi - \rho \\ (\varphi + 1)(1 - \rho) \\ (1 + \rho)(1 + \varphi - \rho) \end{pmatrix}. \tag{14}$$

Profits are equal to $\pi_i = q_i^2/2$ for $i = 1, 2, 3$. The condition for an interior equilibrium is $\rho + \varphi < 1/\sqrt{2}$. Figure 1 shows an illustration of equilibrium outputs and profits for the

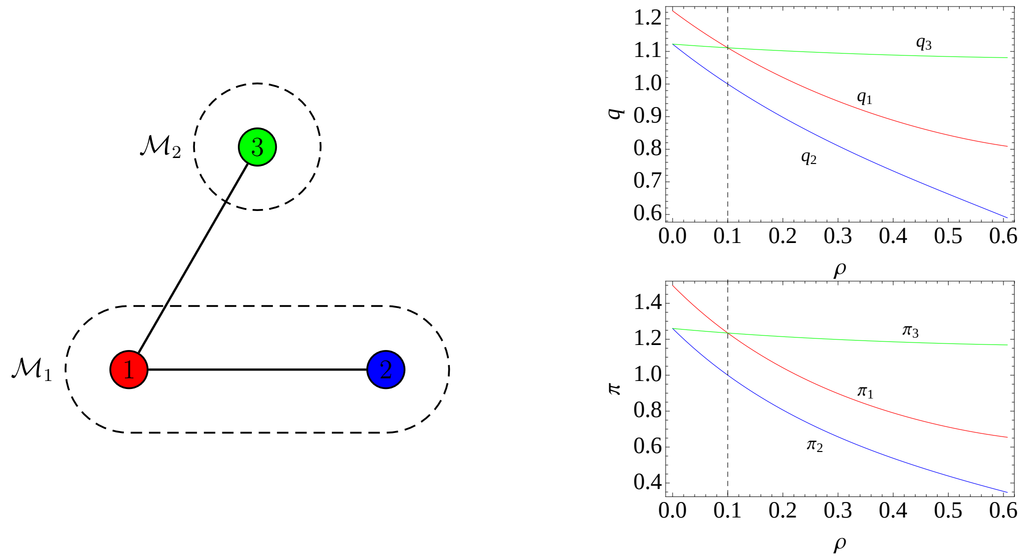


Figure 1: Equilibrium output from Equation (15) and profits for the three firms with varying values of the competition parameter $0 \leq \rho \leq \frac{1}{2}\left(\sqrt{2}-2\varphi\right)$, $\mu = 1$ and $\varphi = 0.1$. Profits of firms 1 and 3 intersect at $\rho = \varphi$ (indicated with a dashed line).

three firms with varying values of the competition parameter $0 \leq \rho \leq \frac{1}{2}\left(\sqrt{2}-2\varphi\right)$, $\mu = 1$ and $\varphi = 0.1$. We see that firm 1 has higher profits due to having the largest number of R&D collaborations when competition is weak (ρ is low compared to φ). However, when ρ increases, its profits decrease and become smaller than the profit of firm 3 when $\rho > \varphi$. This result highlights the key trade off faced by firms between the *technology (or knowledge) spillover effect* and the *product rivalry effect* of R&D [cf. Bloom et al., 2013] since the former increases with φ, which captures the intensity of the spillover effect while the latter increases with ρ, which indicates the degree of competition in the product market.

To better understand these two effects, consider the case of a single market, that is $M = 1$. It is easily verified that, in that case, $\mathbf{B} = (\mathbf{u}\mathbf{u}^\top - \mathbf{I}_n)$ where $\mathbf{u} = (1, \ldots, 1)^\top$ is an n-dimensional vector of ones. In our example, if there is only one market, all three firms will compete with each other in the same market so that:

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

If $\varphi/(1-\rho) < 1/\sqrt{2}$, then the unique equilibrium output will given by:

$$\mathbf{q} = \frac{\mu}{1 - 2\varphi^2 + 4\varphi\rho + \rho - 2\rho^2} \begin{pmatrix} 1 + 2\varphi - \rho \\ 1 + \varphi - \rho \\ 1 + \varphi - \rho \end{pmatrix}. \tag{15}$$

Since there is only one market, the position in the network will determine which firm will produce the most and have the highest profit. As firm 1 is the most central firm in the network and has the highest Bonacich centrality, it has the highest profit. In other words, when $M = 1$, only the *technology (or knowledge) spillover effect* is of importance and the position in the network is the only determi–t of output and profit. However, we saw that this was not the case in the previous example with two markets because, as compared to firm 3, even if firm 1 had the highest Bonacich centrality, it was competing with firm 2 on the product market while firm 3 had no competitor on its market. In other words, there is now a trade off between the position in the network (*technology (or knowledge) spillover effect*) and the position in the product market (*product rivalry effect*). We have seen that, depending on the values of ρ and φ, firm 1 can have a higher or lower output and profit than firm 3.

4. Welfare

Let us now determine the welfare of this economy. We will consider different cases from general to more specific ones. Inserting the inverse demand from Equation (1) into net utility U_m of the consumer in market $\mathcal{M}_m$ shows that

$$U_m = \frac{1}{2} \sum_{i \in \mathcal{M}_m} q_i^2 + \frac{\rho}{2} \sum_{i \in \mathcal{M}_m} \sum_{\substack{j \in \mathcal{M}_m, \\ j \neq i}} q_i q_j$$

For given quantities, the consumer surplus is strictly increasing in the degree ρ of substitutability between products. In the special case of non-substitutable goods, when $\rho \to 0$, we obtain

$$U_m = \frac{1}{2} \sum_{i \in \mathcal{M}_m} q_i^2,$$

while in the case of perfectly substitutable goods, when $\rho \to 1$, we get

$$U_m = \frac{1}{2} \left(\sum_{i \in \mathcal{M}_m} q_i \right)^2.$$

The total consumer surplus is then given by $U = \sum_{m=1}^{M} U_m$. The producer surplus is given by aggregate profits $\Pi = \sum_{i=1}^{n} \pi_i$. As a result, total welfare is equal to $W = U + \Pi$.

4.1. Non-Substitutable Goods

When products are not substitutable ($\rho \to 0$), total welfare is given by the producer and consumer surplus, which can then be written as

$$W(G) = \sum_{i=1}^{n} \left(\frac{q_i^2}{2} + \pi_i \right) = \sum_{i=1}^{n} q_i^2.$$

The following proposition provides upper and lower bounds on welfare for any given graph G, and determines the welfare maximizing graph.

Proposition 2. *Consider independent markets with $\rho \to 0$ and let μ_i and φ satisfy the* *restrictions of Proposition 1. Denote by $\mathcal{G}(n)$ the class of graphs with n nodes and the class of graphs with n nodes and m links by $\mathcal{H}(n,m) \subset \mathcal{G}(n)$.*

(i) Let the largest eigenvalue of the adjacency matrix $\mathbf{A}$ be given by λ_{PF} and let $\mathbf{v}_{PF}$ be the associated eigenvector. Then, social welfare can be written as

$$W(G) = \frac{(\boldsymbol{\mu}^\top \mathbf{v}_{PF})^2}{(1-\varphi\lambda_{PF})^2} + o\left(\frac{1}{1-\varphi\lambda_{PF}}\right),$$

and in the limit of large φ the efficient graph $G^* = \mathrm{argmax}_{G\in\mathcal{H}(n,m)} W(G)$ is a nested split graph where the ordering of degrees $\{d_i\}_{i=1}^n$ follows the ordering of $\{\mu_i\}_{i=1}^n$.

(ii) Assume that $\mu_i = \mu$ for all $i \in \mathcal{N}$. Then, welfare in the efficient graph $G^* = \mathrm{argmax}_{G\in\mathcal{H}(n,m)} W(G)$ can be bounded from above and from below as

$$\frac{\mu^2 n}{(1-\varphi\bar{d})^2} \leq W(G^*) \leq \frac{\mu^2 n}{\left(1-\varphi\sqrt{(n-1)\bar{d}}\right)^2},$$

where $\bar{d} = \frac{2m}{n}$ is the average degree in G.

(iii) The efficient graph $G^* = \mathrm{argmax}_{G\in\mathcal{G}(n)} W(G)$ is the one that maximizes the largest eigenvalue λ_{PF}, that is, the complete graph K_n.

This proposition provides several interesting results. First, when products are not substitutable ($\rho = 0$), we are able to write an explicit expression of total welfare as a function of the fundamental parameters of the model and provide a lower and an upper bound on this welfare function where the bounds depend on the parameters as well as the topology of the network (captured by the average degree in the network). Second, in terms of network design, when $\rho = 0$, there is no competition effect and thus, only spillover effects through the network matter. As a result, it should not be surprising that the complete network is the efficient network because of positive complementarities between firms. We also show that when φ is large (close to its maximum value in the limit), the nested split graph is the efficient network.[16,17] Basically, in a nested-split graph, the neighborhood of a node is contained in the neighborhoods of the nodes with higher degrees (see König et al. [2014] for a discussion of these graphs). If one looks at the leading term in the welfare function, then one can see that it depends on the product of $\boldsymbol{\mu}$ and the Perron eigenvector $\mathbf{v}_{\mathrm{PF}}$. In any nested split graph, the node with the highest degree also has the highest eigenvector. As a result, in order to maximize total welfare, one wants to have the node i with the highest μ_i to have the highest eigenvector component, which means that it should also have the highest degree. Note that similar results relating the largest eigenvalue to efficiency have been obtained in Corbo et al.

[16]The complete graph K_n is a particular (degenerate) case of a nested split graph.
[17]In Appendix A.3, we formally define nested-split graphs.

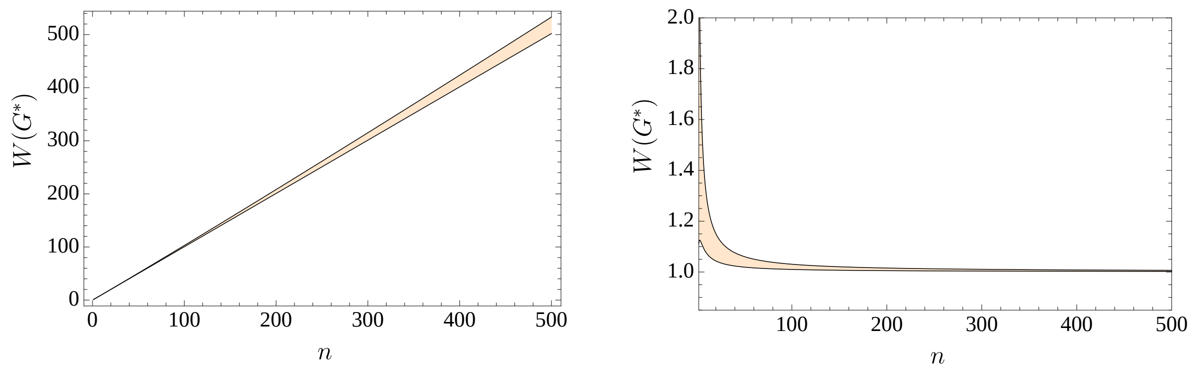


Figure 2: (Left panel) The two bounds from Proposition 2 for $\varphi = 0.001$, $\mu = 1$, $m = n-1$ for varying values of n. (Right panel) The two bounds from Proposition 4 for $\rho = 0.5$, $\varphi = 0.0001$ and $\mu = 1$. Note that the comparison of welfare in the case of $\rho \to 0$ and $\rho > 0$ in the above figures for increasing n is not meaningful as the first considers a growing number of products, while the latter a single product with an increasing number of firms producing it.

[2006], König et al. [2011] and Belhaj et al. [2013]. The two bounds from Proposition 2 part (ii) are shown in Figure 2 (left panel).

4.2. Imperfectly Substitutable Goods

In this section, we allow for products to be substitutable, i.e. $\rho > 0$. Then, social welfare is given by

$$W(G) = \frac{1}{2}\left(\sum_{i=1}^{n} q_i^2 + \rho \sum_{i=1}^{n}\sum_{j\neq i}^{n} b_{ij} q_i q_j\right) + \sum_{i=1}^{n} \pi_i,$$

where equilibrium output and profits are given by Equations (9) and (11). Inserting profits as a function of output leads to:

$$W(G) = \sum_{i=1}^{n} q_i^2 + \frac{\rho}{2}\sum_{i=1}^{n}\sum_{j\neq i}^{n} b_{ij} q_i q_j = \mathbf{q}^\top \mathbf{q} + \frac{\rho}{2}\mathbf{q}^\top \mathbf{B}\mathbf{q},$$

We are now able to state a similar result as in part (i) of Proposition 2 for the case of (imperfectly) substitutable goods.

Proposition 3. *Denote by $\mathbf{C} = \mathbf{A} - \frac{\rho}{\varphi}\mathbf{B}$, let $\{\nu_i\}_{i=1}^n$ be the eigenvalues of $\mathbf{C}$ and $\{\mathbf{v}_i\}_{i=1}^n$ the associated eigenvectors. Then, welfare can be written as*

$$W(G) = \frac{2-\rho}{2}\frac{(\boldsymbol{\mu}^\top \mathbf{v}_1)^2}{(1-\varphi\nu_1)^2}\left(1 + \frac{\rho}{2-\rho}\mathbf{v}_1^\top \mathbf{B}\mathbf{v}_1\right) + o\left(\frac{1}{1-\varphi\nu_1}\right)^2.$$

Proposition 3 shows that when spillover effects are strong such that the leading terms in $1/(1-\varphi\nu_1)$ dominate, then welfare is determined by the weighted sum of the eigenvector components $\boldsymbol{\mu}^\top \mathbf{v}_1 = \sum_{i=1}^n \mu_i v_{1,i}$ and the pairwise eigenvector complementarity effects in different markets $\mathbf{v}_1^\top \mathbf{B}\mathbf{v}_1 = \sum_{i=1}^n \sum_{j=1}^n v_{1,i} b_{ij} v_{1,j}$.[18]

To gain further insights, we will assume in the following that there is only a single market (with $M = 1$, $b_{ij} = 1$ for $i \neq j$ and $b_{ii} = 1$ for all $i, j \in \mathcal{N}$) and make the homogeneity assumption that $\mu_i = \mu$ for all $i \in \mathcal{N}$. Then, welfare can be written as follows

$$W(G) = \frac{2-\rho}{2}\|\mathbf{q}\|_2^2 + \frac{\rho}{2}\|\mathbf{q}\|_1^2,$$

where $\|\mathbf{q}\|_p \equiv (\sum_{i=1}^n q_i^p)^{\frac{1}{p}}$ is the ℓ^p-norm of $\mathbf{q}$ and $\mathbf{u} = (1, \ldots, 1)^\top$ is a vector of ones. Using the fact that $\|\mathbf{q}\|_2 \leq \|\mathbf{q}\|_1 \leq \sqrt{n}\|\mathbf{q}\|_2$, we obtain an upper bound on welfare given by

$$W(G) \leq \frac{2+(n-1)\rho}{2}\|\mathbf{q}\|_2^2 = (2+(n-1)\rho)\Pi,$$

where aggregate profits are given by $\Pi = \sum_{i=1}^n \pi_i$. Hence, welfare is upper bounded by a proportionality factor times the total profits generated in the economy.

We next consider the efficient network for small values of $\phi = \varphi/(1-\rho)$ (defined as in Proposition 1) and provide lower and upper bounds for welfare in the efficient graph.

Proposition 4. *Consider a large market with substitutable goods where $\rho > 0$. Further, assume that $\mu_i = \mu$ for all $i = 1, \ldots, n$, and let ρ, μ, φ and ϕ satisfy the restrictions of Proposition 1. Denote by $\mathcal{G}(n)$ the class of graphs with n nodes and the class of graphs with n nodes and m links by $\mathcal{H}(n,m) \subset \mathcal{G}(n)$.*

(i) For small values of ϕ, such that terms of the order $O(\phi^3)$ can be neglected, welfare $W(G)$ is maximized in the graph $G \in \mathcal{H}(n,m)$ with the smallest degree variance σ_d^2.

(ii) For small values of ϕ such that terms of the order $O(\phi^4)$ can be neglected, welfare $W(G)$ for two graphs $G, G' \in \mathcal{H}(n,m)$ with the same degree variance σ_d^2 is higher for the one which is less degree assortative.

(iii) Assume that $0 < \rho < 1$. Then, welfare in the efficient graph $G^* = \mathrm{argmax}_{G\in\mathcal{H}(n,m)} W(G)$ can be bounded from above and from below as follows

$$\frac{\mu^2 n(2+(n-1)\rho)}{2(1+(n-1)\rho-\bar{d}\varphi)^2} \leq W(G^*) \leq \frac{2-\rho}{2}\frac{\mu^2}{\rho^2}\left(\frac{\rho}{2-\rho} + \frac{1}{n\left(1-\frac{\varphi}{1-\rho}\sqrt{\bar{d}(n-1)}\right)}\right),$$

where $\bar{d} = \frac{2m}{n}$ is the average degree in G.

(iv) Assume that $0 < \rho < 1$. Then welfare in the efficient graph $G^* = \mathrm{argmax}_{G\in\mathcal{G}(n)} W(G)$ is bounded from above and from below by

$$\frac{\mu^2 n((n-1)\rho+2)}{2((n-1)\rho-n\varphi+\varphi+1)^2} \leq W(G^*) \leq \frac{\mu^2((n-1)n\rho\varphi+(\rho-1)((n-1)\rho+2))}{2n\rho^2((n-1)\varphi+\rho-1)}.$$

[18]Further results for the case of large spillover effects can be found in Appendix C.

In particular, in the limit of large industry size n, we have that $\lim_{n\to\infty} W(K_n)/W(G^*) = \frac{\rho^2}{(\rho-\varphi)^2}$*, and for weak spillovers, the complete graph* K_n *is efficient, that is* $\lim_{\varphi\to 0}\lim_{n\to\infty} W(K_n) = W(G^*)$.

Proposition 4 case (i) is in contrast to previous studies such as Westbrock [2010], where it is argued that welfare in R&D collaboration networks is increasing with the degree variance. Part (ii) of the proposition shows that once we allow for stronger spillover effects, the assortativity of the network is also of importance for welfare.[19] Part (iii) gives a general result for the competition effect when $0 < \rho < 1$. In that case, the welfare maximizing graph G^* can be bounded above and below. The two bounds from part (iv) of Proposition 4 are shown in Figure 2 (right panel). The last result in part (iv) once more shows that the complete network is the efficient one in large industries if the spillover effects are not too strong.[20]

5. The Key Player Policy

As stated in the introduction, the key-player problem was first been introduced in economics by Ballester et al. [2006, 2010]. In the context of crime, they have argued that concentrating efforts by targeting *key players*, i.e. criminals who once removed generate the highest possible reduction in the aggregate crime level in a network, can have large effects on crime because of feedback effects or *social multipliers*. Based on a peer-effect model, Ballester et al. have proposed a centrality measure (the *intercentrality measure*) that determines the key player in each network. Because we are not dealing with crime but with R&D networks, we will redefine the key-player policy in the following way. First, as shown in Proposition 1, where only cases (iii) and (iv) correspond to the model of Ballester et al., we will consider a much more general model where both network and competition effects are of importance in a context of different markets (or sectors) and different types of goods. Second, we define the key player in a different way: it will be the firm which once removed from the network reduces *total welfare* the most (and not *total activity* or *total output* as in Ballester et al.). As it will turn out, the centrality that we obtain to define the key player (or the key firm) will be quite different from the intercentrality measure proposed by Ballester et al..

Let G^{-i} be the network obtained from G by removing firm i. The key firm is the one whose removal from the network reduces welfare the most, i.e., the key firm $i^* \in \mathcal{N} = \{1, \ldots, n\}$ and is defined by $i^* = \arg\max_{i\in\mathcal{N}}\{W(G) - W(G^{-i})\}$. The following proposition characterizes the key firm i^* both when $\rho = 0$ and when $\rho > 0$.

[19] The assortativity coefficient $\rho_d(G) \in [-1, 1]$ is essentially the Pearson correlation coefficient of degree between nodes that are connected. Positive values of $\rho_d(G)$ indicate that nodes with similar degrees tend to be connected, while negative values indicate that nodes with different degrees tend to be connected. See Newman [2002] and Pastor-Satorras et al. [2001] for further details.

[20] In Appendix C, we provide additional results on welfare where we focus on a particular class of networks, namely those with a large spectral gap.

Proposition 5. *Let ρ, μ_i, $i \in \mathcal{N}$, φ and ϕ be defined as in Proposition 1.*

(i) Assume that goods are not substitutable, i.e. $\rho = 0$ and let $\varphi < 1/\lambda_{PF}$. Moreover, let $N_G(\varphi, i) = m_{ii}(G, \varphi)$ denote the generating function of the number of closed walks[21] *that start and terminate at node i and let $\mathbf{M}(G, \varphi) \equiv (\mathbf{I}_n - \varphi\mathbf{A})^{-1}$. Then, the key firm is given by $i^* = \arg\max_{i \in \mathcal{N}} c_i(G, \varphi)$, where the intercentrality of firm i is given by*

$$c_i(G, \varphi) = \frac{b_{\boldsymbol{\mu},i}(G, \varphi)}{N_G(\varphi, i)} \left[(\mathbf{M}(G, \varphi)\mathbf{b}_{\boldsymbol{\mu}}(G, \varphi))_i - \frac{1}{2} \frac{b_{\boldsymbol{\mu},i}(G, \varphi)}{N_G(\varphi, i)} (\mathbf{M}(G, \varphi)^2)_{ii} \right]. \qquad (16)$$

(ii) Assume that goods are substitutable, i.e. $\rho > 0$, that the matrix $\mathbf{M}(G, \rho, \varphi) = (\mathbf{I}_n + \rho\mathbf{B} - \varphi\mathbf{A})^{-1}$ exists[22] *and let $\mathbf{b}_{\boldsymbol{\mu}}(G, \rho, \varphi) = \mathbf{M}(G, \rho, \varphi)\boldsymbol{\mu}$. Then, the key firm is given by $i^* = \arg\max_{i \in \mathcal{N}} c_i(G, \rho, \varphi)$, where the intercentrality of firm i is given by*

$$\begin{aligned} c_i(G, \rho, \varphi) = \frac{b_{\boldsymbol{\mu},i}(G, \rho, \varphi)}{m_{ii}(G, \rho, \varphi)} \Big(&(\mathbf{M}(G, \rho, \varphi)(2\mathbf{I}_n + \rho\mathbf{B})\mathbf{b}_{\boldsymbol{\mu}}(G, \rho, \varphi))_i \\ &- \frac{1}{2} \frac{b_{\boldsymbol{\mu},i}(G, \rho, \varphi)}{m_{ii}(G, \rho, \varphi)} (\mathbf{M}(G, \rho, \varphi)(2\mathbf{I}_n + \rho\mathbf{B})\mathbf{M}(G, \rho, \varphi))_{ii} \Big). \end{aligned}$$

Let us start with case (i), which assumes that goods are not substitutable, i.e. $\rho = 0$. We propose a new intercentrality measure, which is an alternative to that of Ballester et al. [2006, 2010] defined as $\frac{b_{\mathbf{u},i}(G,\varphi)^2}{N_G(\varphi,i)}$. Our intercentrality measure is defined as $c_i = \frac{1}{2}\frac{d}{d\varphi}\left(\frac{\varphi b_{\boldsymbol{\mu},i}(G,\varphi)^2}{N_G(\varphi,i)}\right)$, which, after some calculations, can be written as in Equation (16) (see the proof of Proposition 5). As for the case of crime, the key firm need not necessarily be the one producing the highest output level or, equivalently, the one with the highest Bonacich centrality measure. This is because the removal of the key firm has both a *direct* and an *indirect* effect on total welfare and thus, the choice of key firm results from a compromise between these two effects. Indeed, if the choice of key firm was solely governed by the direct effect of firm removal on aggregate welfare, the most productive firms would be the natural candidates. But the choice of key firm must also take into account the indirect effect on aggregate welfare reduction induced by the network restructuring that follows from the removal of one firm from the original network. Our intercentrality measure defined in Equation (16) takes this trade off into account. Interestingly, the two effects appear in both intercentrality measures (ours and that of Ballester et al. [2006, 2010]), the two effecs appear since they include both the Bonacich centrality of the key firm (direct effect) and the generating function of the number of closed walks that start and terminate at the key firm (indirect effect through self-loops).

If we now consider the more general case ($\rho > 0$) where both the network effect and the competition effect are taken into account, it can be seen that there is a difference in

[21] See Appendix A.2 for a formal definition of walk generating functions of a graph and some results associated with them.

[22] See Proposition 1, item (i), for a sufficient condition that guarantees that this matrix is invertible.

the weighted Bonacich centralities $\mathbf{b}_{\boldsymbol{\mu}}(G, \cdot)$ between part (i) and part (ii) of Proposition 5. While the first is the standard weighted Bonacich centrality of the network G with firm-specific weights μ_i (see Appendix A.4), in part (ii), the Bonacich centrality depends on both, the adjacency matrix $\mathbf{A}$ of network G and the block diagonal matrix $\mathbf{B}$, which indicates which firm is competing with which other firms. This is an important generalization of the intercentrality measure of the key player, which we believe to be crucial when one deals with R&D networks (but also any network with both spillover and competition effects) since, as stated above, there is a key trade off faced by firms between the *technology (or knowledge) spillover effect* and the *product rivalry effect* of R&D that needs to be accounted for.

6. The R&D Subsidy Policy

In this section, we would like to consider an alternative policy to the key player one, that is R&D subsidies. Indeed, in order to foster innovative activities and economic growth, governments in numerous countries have introduced R&D support programs aimed at increasing the R&D effort in the private sector.[23] Moreover, national governments in a number of countries subsidize the R&D activities of domestic firms, particularly in industries where foreign and domestically owned firms are in competition for international markets. Such programs are, for example, the EUREKA program in the European Union or the SPIR program in the United States.

To better understand this issue, we would now like to extend our framework by considering an optimal R&D subsidy program in the short run, i.e. taking the network G as given. For our analysis, we first assume that all firms obtain a homogeneous subsidy per unit of R&D effort spent. Then, we proceed by allowing the social planner to differentiate between firms and implement firm-specific R&D subsidies.[24]

6.1. Homogeneous R&D Subsidies

Let us first consider the case of a single market, $M = 1$. An active government is introduced that can provide a subsidy, $s \geq 0$, per unit of R&D. It is assumed that each firm receives the same per unit R&D subsidy. The profit of firm i with an R&D subsidy can then be written as:

$$\pi_i = (\bar{\alpha} - \bar{c}_i)q_i - q_i^2 - \rho q_i \sum_{j \neq i} b_{ij} q_j + q_i e_i + \varphi q_i \sum_{j=1}^{n} a_{ij} e_j - \frac{1}{2} e_i^2 + s e_i. \tag{17}$$

[23] Public R&D grants covered about 7.5 % of private R&D in the OECD countries in 2004 [OECD, 2012]. For an overview of R&D tax credits which are another commonly used fiscal incentive for R&D investment, see Bloom et al. [2002]. Takalo et al. [2013] analyze the welfare effects of targeted R&D subsidies using project-level data from Finland.

[24] We would like to emphasize that, as we have normalized the cost of R&D to one in the profit function of Equation (3), the absolute values of R&D subsidies are not meaningful in the subsequent analysis, but rather relative comparisons across firms are.

This formulation is similar to that of Spencer and Brander [1983] where each firm i receives a fixed amount of money se_i proportional to firm i's effort e_i. The government (or the planner) is here introduced as an agent that can set subsidy rates on R&D effort in a period before the firms spend on R&D. The assumption that the government can pre-commit itself to such subsidies and thus can act in this leadership role is fairly natural. As a result, this subsidy will affect the levels of R&D conducted by firms, but not the resolution of the output game. In this context, the optimal R&D subsidy s^* determined by the planner is found by maximizing total welfare $W(G, s)$ less the cost of the subsidy $s\sum_{i=1}^{n} e_i$, taking into account the fact that firms choose output and effort for a given subsidy level by maximizing the profits in Equation (17). If we define net welfare as $\overline{W}(G, s) \equiv W(G, s) - s\sum_{i=1}^{n} e_i$, the social planner's problem is given by

$$s^* = \arg\max_{s\in\mathbb{R}_+} \overline{W}(G, s).$$

The following proposition derives the Nash equilibrium quantities and efforts and the optimal subsidy level that solves the planner's problem.

Proposition 6. *Consider the n–player simultaneous move game with profits given by Equation (17) where firms choose quantities and efforts in the strategy space in $\mathbb{R}_+^n \times \mathbb{R}_+^n$. Further, let μ_i, $i \in \mathcal{N}$ be defined as in Proposition 1.*

(i) If Equation (5) holds, then the matrix $\mathbf{M} = (\mathbf{I}_n + \rho\mathbf{B} - \varphi\mathbf{A})^{-1}$ exists, and the unique interior Nash equilibrium in quantities with subsidies (in the second stage) is given by

$$\mathbf{q} = \bar{\mathbf{q}} + s\mathbf{r}, \tag{18}$$

where $\bar{\mathbf{q}} = \mathbf{M}\boldsymbol{\mu}$ and $\mathbf{r} = \varphi\mathbf{M}\left(\frac{1}{\varphi}\mathbf{u} + \mathbf{A}\mathbf{u}\right)$. The equilibrium profits are given by

$$\pi_i = \frac{q_i^2 + s^2}{2}. \tag{19}$$

(ii) Assume that goods are not substitutable, i.e. $\rho = 0$. Then if $\sum_{i=1}^{n}\left(r_i^2(1-3) + 2r_i + 1\right) \geq 0$, the optimal subsidy level (in the first stage) is given by

$$s^* = \frac{\sum_{i=1}^{n} \bar{q}_i\,(1 - 2r_i)}{\sum_{i=1}^{n}\left(r_i\,(2r_i - 2) - 1\right)},$$

(iii) Assume that goods are substitutable, i.e. $\rho > 0$. Then if

$$\sum_{i=1}^{n}\left(r_i^2(1-3) + 2r_i + 1 - \rho\sum_{j=1}^{n} b_{ij} r_i r_j\right) \geq 0,$$

the optimal subsidy level (in the first stage) is given by

$$s^* = \frac{\sum_{i=1}^n \left(\bar{q}_i(2r_i - 1) + \frac{\rho}{2}\sum_{j=1}^n b_{ij}(\bar{q}_i r_j + \bar{q}_j r_i)\right)}{\sum_{i=1}^n \left(1 + r_i\left(2 - 2r_i - \rho\sum_{j=1}^n b_{ij} r_j\right)\right)},$$

In part (i) of Proposition 6, we solve the second stage of the game where firms decide their output given the homogenous subsidy s. In parts (ii) and (iii) of the proposition, we solve the first stage when the planner optimally determines the subsidy per R&D effort when goods are not substitutable, i.e. $\rho = 0$, and when they are ($\rho > 0$). We are able to determine the exact value of the optimal subsidy to be given to each firm embedded in a network of R&D collaborations in both cases. Interestingly, the optimal subsidy depends on the vector $\mathbf{r} = \mathbf{Mu} + \varphi\mathbf{MAu}$ where the vector $\mathbf{Au}$ determines the *degree* (i.e. number of links) of each firm.

6.2. Targeted R&D Subsidies

We now consider the case where the planner can discriminate between firms by offering different subsidies. In other words, we assume that each firm i, for all $i = 1, \ldots, n$, obtains a subsidy $s_i \geq 0$ per unit of R&D effort. The profit of firm i can then be written as:

$$\pi_i = (\bar{\alpha} - \bar{c}_i)q_i - q_i^2 - \rho q_i \sum_{j \neq i} b_{ij} q_j + q_i e_i + \varphi q_i \sum_{j=1}^n a_{ij} e_j - \frac{1}{2}e_i^2 + s_i e_i. \tag{20}$$

As above, the optimal R&D subsidies $\mathbf{s}^*$ are then found by maximizing welfare $W(G, \mathbf{s})$ less the cost of the subsidy $\sum_{i=1}^n s_i e_i$, when firms are choosing output and effort for a given subsidy level by maximizing the profits in Equation (20). If we define net welfare as $\overline{W}(G, \mathbf{s}) \equiv W(G, \mathbf{s}) - \sum_{i=1}^n e_i s_i$, then the solution to the social planner's problem is given by

$$\mathbf{s}^* = \arg\max_{\mathbf{s} \in \mathbb{R}_+^n} \overline{W}(G, \mathbf{s}).$$

The following proposition derives the Nash equilibrium quantities and efforts (second stage) and the optimal subsidy levels that solve the planner's problem (first stage).

Proposition 7. *Consider the n–player simultaneous move game with profits given by Equation (17) where firms choose quantities and efforts in the strategy space in $\mathbb{R}_+^n \times \mathbb{R}_+^n$. Further, let μ_i, $i \in \mathcal{N}$ be defined as in Proposition 1.*

(i) If Equation (5) holds, then the matrix $\mathbf{M} = (\mathbf{I}_n + \rho\mathbf{B} - \varphi\mathbf{A})^{-1}$ exists, and the unique interior Nash equilibrium in quantities with subsidies (in the second stage) is given by

$$\mathbf{q} = \bar{\mathbf{q}} + \mathbf{Rs}, \tag{21}$$

where $\mathbf{R} = \mathbf{M}(\mathbf{I}_n + \varphi\mathbf{A})$, $\bar{\mathbf{q}} = \mathbf{M}\boldsymbol{\mu}$, *and equilibrium profits are given by*

$$\pi_i = \frac{q_i^2 + s_i^2}{2}. \tag{22}$$

(ii) Assume that goods are not substitutable, i.e. $\rho = 0$. *Then if the matrix* $\mathbf{I}_n + 2\mathbf{R} - 2\mathbf{R}^2$ *is positive definite, the optimal subsidy levels (in the first stage) are given by*

$$\mathbf{s}^* = (\mathbf{I}_n + 2\mathbf{R} - 2\mathbf{R}^2)^{-1}(2\mathbf{R} - \mathbf{I}_n)\bar{\mathbf{q}}.$$

(iii) Assume that goods are substitutable, i.e. $\rho > 0$. *Then, if the matrix* $\mathbf{I}_n - 2\mathbf{R}^\top\left(\frac{1}{2}(2\mathbf{I}_n + \rho\mathbf{B}\right)\mathbf{R} - \mathbf{I}_n)$ *is positive definite, the optimal subsidy levels (in the first stage) are given by*

$$\mathbf{s}^* = \left(\mathbf{I}_n - 2\mathbf{R}^\top\left(\frac{1}{2}(2\mathbf{I}_n + \rho\mathbf{B}\right)\mathbf{R} - \mathbf{I}_n)\right)^{-1}\left(\mathbf{R}^\top(2\mathbf{I}_n + \rho\mathbf{B}) - \mathbf{I}_n\right)\bar{\mathbf{q}}.$$

As in the previous proposition, in part (i) of Proposition 7, we solve for the second stage of the game where firms decide their output given the *targeted* subsidy s_i. In parts (ii) and (iii), we solve the first stage of the model when the planner optimally decides the targeted subsidy per R&D effort when goods are substitutable (i.e. $\rho > 0$), and when they are not (i.e. $\rho = 0$). We are able to determine the exact value of the optimal subsidy to be given to each firm embedded in a network of R&D collaborations in both cases.[25] We will use the results of these two propositions below to empirically study R&D collaborations between firms in our dataset.

We would now like to test the different parts of the theoretical results. First, we will test Proposition 1 and will try to disentangle between the *technology (or knowledge) spillover effect* and the *product rivalry effect* of R&D. Second, once the parameters of the model have been estimated, we will determine which the key firms are in our dataset and make a comparison with those that should be subsidized.

7. Data

We use data on interfirm R&D collaborations stemming from the MERIT-CATI database.[26] Given its history and coverage, the MERIT-CATI database is one of the few databases that allows us to study patterns in R&D partnerships in several industries, both domestically and internationally, in different regions of the world over an extended period of several decades. This database contains information about strategic technology agreements, including any

[25] Note that when the condition for positive definiteness is not satisfied then we can sill use parts (ii) or (iii) of Proposition 7, respectively, as a candidate for a welfare improving subsidy program. However, there might exist other subsidy programs which yield even higher welfare gains.

[26] We would like to thank Christian Helmers for providing access to the dataset.

alliance that involves some arrangements for mutual transfer of technology or joint research, such as joint research pacts, joint development agreements, cross licensing, R&D contracts, joint ventures and research corporations [cf. Hagedoorn, 2002].[27] The database only records agreements for which a combined innovative activity or an exchange of technology is at least part of the agreement. Moreover, only agreements that have at least two industrial partners are included in the database, thus agreements involving only universities or government labs, or one company with a university or lab, are disregarded. From the MERIT-CATI database, we then obtain a total of 13,040 companies. The systematic collection of inter-firm alliances started in 1987 and ended in 2006. However, information about alliances prior to 1987 is available in the database, and we use all information available starting from the year 1950.[28] We construct the R&D alliance network by assuming that an alliance lasts 5 years [similar to e.g. Rosenkopf and Padula, 2008].[29] In the robustness section below (Section 9.2.2), we will test our model for different durations of alliances.

Figure 3 shows the number of firms n participating in an alliance in the R&D network constructed in this way, the average degree $\bar{d}$, the degree variance σ_d^2 and the degree coefficient of variation, i.e. $c_v = \sigma_d/\bar{d}$, over the years 1990 to 2005. It can be seen that there are very large variations over the years in the number of firms having an R&D alliance with other firms. Starting from 1990, we observe a strong increase followed by a sudden drop to a low level. Since 1998, it is once more increasing. Interestingly, the average number of alliances per firm (captured by the average degree $\bar{d}$), the degree variance σ_d^2 as well as the degree coefficient of variation c_v have decreased over the years, indicating lower inter-firm collaboration activity levels.

In Figure 4,[30,31] exemplary plots of the largest connected component in the R&D network for the years 1990, 1995, 2000 and 2005 are shown. In 1990, the giant component had a core-periphery structure with many R&D interactions between firms from different sectors. If we look at the same picture in 2005, the core-periphery structure seems less obvious and two cores and a periphery seem to emerge, where there are only few interactions between firms of different sectors in one of the cores. This may indicate more specialization in R&D

[27] Schilling [2009] compares different alliance databases, including MERIT-CATI, and finds that the different databases show similar patterns.

[28] As explained below, we do not have any information available on firms' financial reports from Standard & Poor's Compustat database prior to 1950. Hence, we choose this year as the first year of observation in our sample.

[29] Rosenkopf and Padula [2008] use a five-year moving window assuming that alliances have a five-year life span, and state that the choice of a five-year window is consistent with extant alliance studies [e.g. Gulati and Gargiulo, 1999; Stuart, 2000] and conforms to Kogut [1988] finding that the normal life span of most alliances is no more than five years. Moreover, Harrigan [1988] studies 895 alliances from 1924 to 1985 and concludes that the average life-span of the alliance is relatively short, 3.5 years, with a standard deviation of 5.8 years and 85 % of these alliances last less than 10 years. Park and Russo [1996] focus on 204 joint ventures among firms in the electronic industry for the period 1979–1988. They show that less than half of these firms remain active beyond a period of five years and for those that last less than 10 years (2/3 of the total), the average lifetime turns out to be 3.9 years.

[30] See Appendix A.1 for the definition of a connected component.

[31] Only firms for which we could obtain their industry classification are shown.

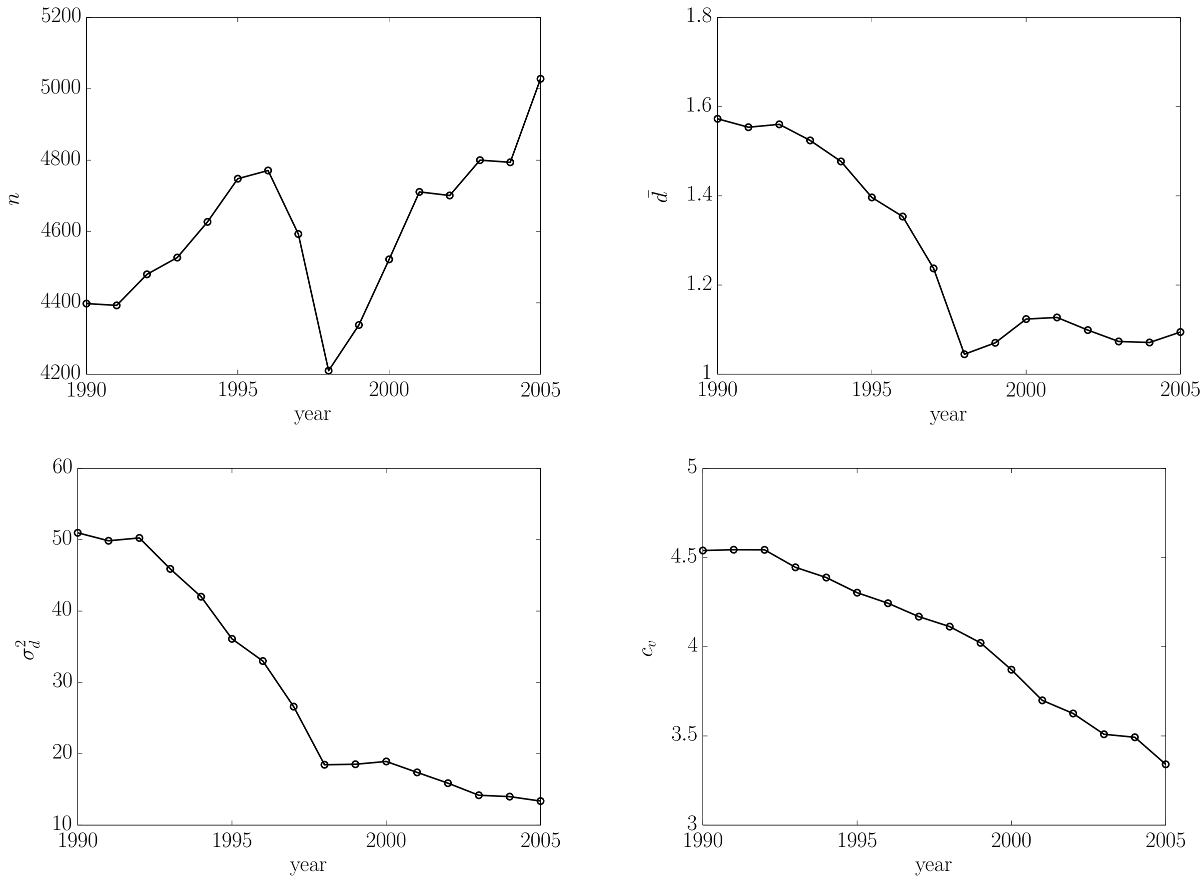


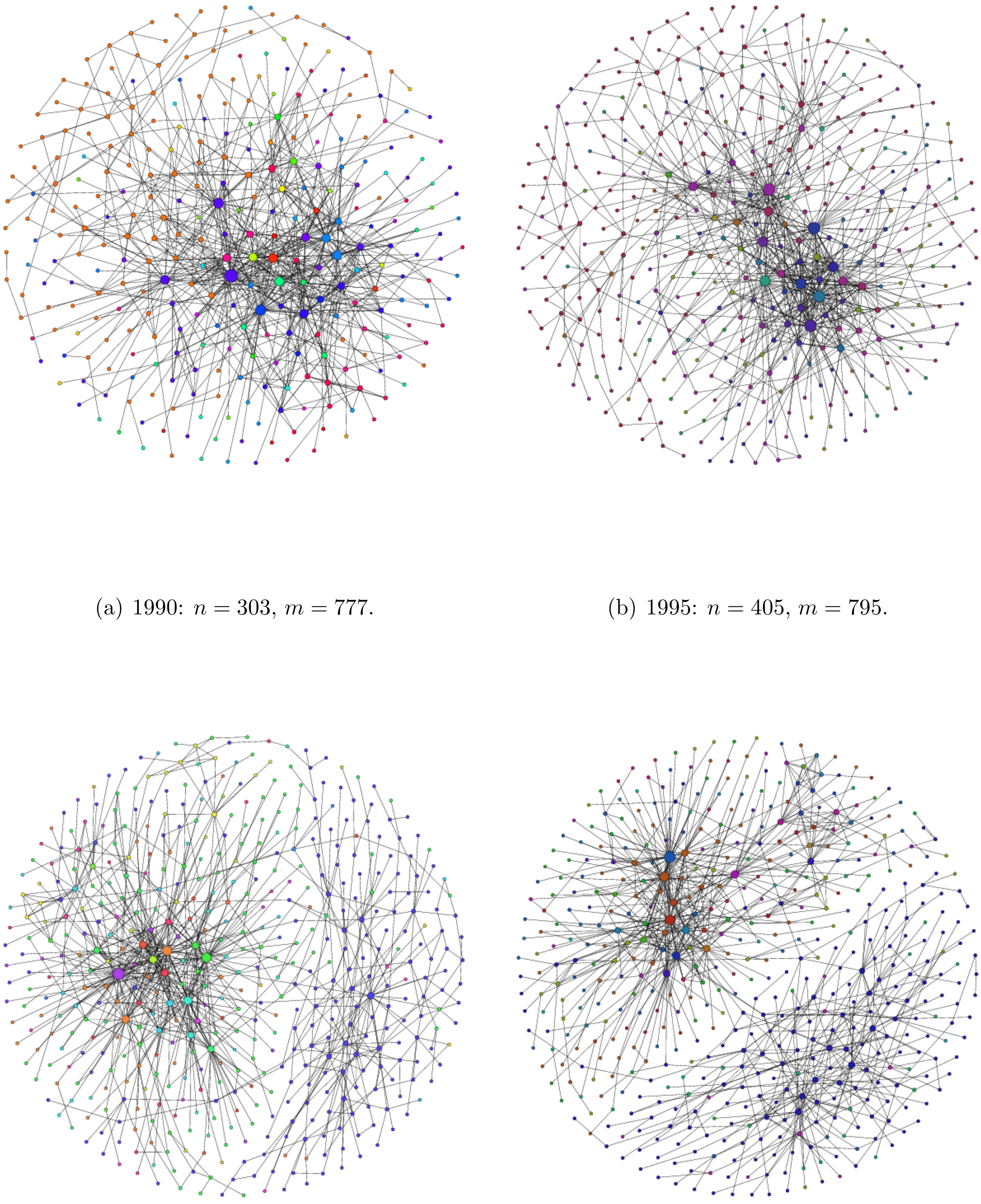
Figure 3: The number of firms n participating in an alliance, the average degree $\bar{d}$, the degree variance σ_d^2 and the degree coefficient of variation $c_v = \sigma_d/\bar{d}$.

alliance partnerships.

The CATI database provides the names for each firm in an alliance. We matched the firms' names in the CATI database with the firms' names in Standard & Poor's Compustat US and Global fundamentals databases, to obtain information about their balance sheets and income statements. For this purpose, we adopted and extended the name matching algorithm developed as part of the NBER patent data project.[32] We could match roughly 20% of the firms in the CATI data. From our match between the firms' names in the CATI database and the firms' names in the Compustat database, we obtained a firm's sales, cost of goods sold, number of employees and capital. As we do not observe physical output of the firms directly, we use the profit function from Equation (3), $\pi_i = (p_i - c_i)q_i - \frac{1}{2}e_i^2$, together with the Nash equilibrium expressions $e_i = q_i$ and $\pi_i = \frac{1}{2}q_i^2$ (cf. Proposition 1), to obtain a relationship between output q_i, sales, $p_i q_i$, and the cost of goods sold, $c_i q_i$, given by $q_i = \sqrt{p_i q_i - c_i q_i}$.[33] The empirical distributions for output $P(q)$ (using a logarithmic binning of the data with 100 bins) and the degree distribution $P(d)$ are shown in Figure 5. Both are highly skewed, indicating a large degree of inequality in the number of goods produced as well as the number

[32]See https://sites.google.com/site/patentdataproject.

[33]Our approach circumvents the standard solution in the literature to deflate firm-level sales by an industrywide producer price index in order to eliminate price effects to compute output. This procedure yields unbiased results only when every firm's price relative to the industry producer price index does not change over time [cf. Bloom et al., 2013; DeLoecker, 2011].



(a) 1990: $n = 303$, $m = 777$.

(b) 1995: $n = 405$, $m = 795$.

(c) 2000: $n = 497$, $m = 845$.

(d) 2005: $n = 513$, $m = 861$.

Figure 4: Network snapshots of the largest connected component for the years (a) 1990, (b) 1995, (c) 2000 and (d) 2005 with the number of firms n and the number of links m. Node colors represent different industry SIC codes at the 4-digit level. The nodes' sizes indicate their degree.

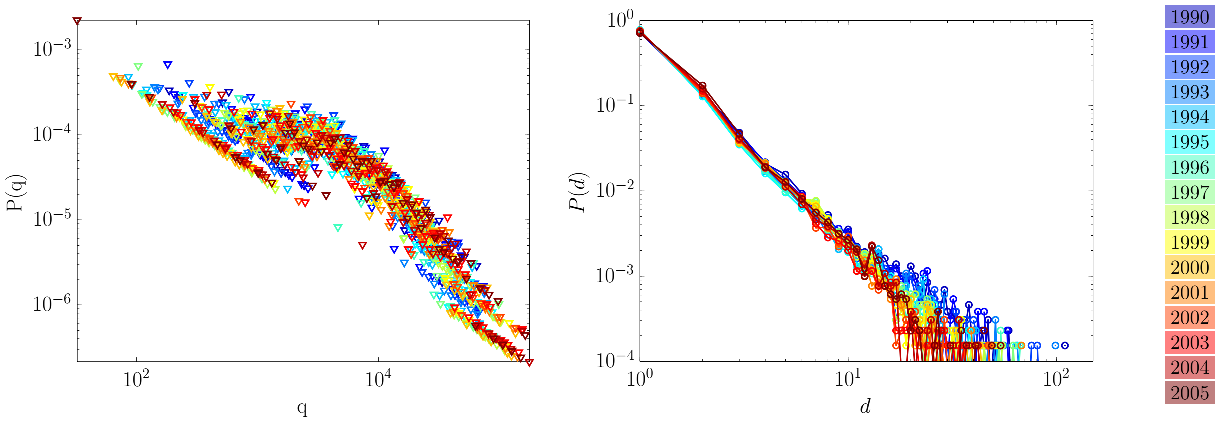


Figure 5: Empirical output distribution $P(q)$ and the distribution of degree $P(d)$ for the years $1975, \ldots, 2006$. The data for output has been logarithmically binned and non-positive data entries have been discarded.

of R&D collaborations.

8. Econometric Analysis

8.1. Econometric Specification

In this section, we introduce the econometric equivalent to the equilibrium quantity produced by each firm given in Equation (13). Our empirical counterpart of the marginal cost c_{it} of firm i from Equation (2) at period t has a fixed cost equal to $\bar{c}_{it} = \eta_i^* - \epsilon_{it} - \mathbf{x}_{it}^\top \boldsymbol{\beta}$, and thus, we get

$$c_{it} = \eta_i^* - \varepsilon_{it} - \mathbf{x}_{it}^\top \boldsymbol{\beta} - e_{it} - \varphi \sum_{j=1}^{n} a_{ij,t} e_{jt}, \tag{23}$$

where $\mathbf{x}_{it}$ is a k-dimensional vector of observed exogenous characteristics of firm i, η_i^* captures the unobserved (to the econometrician) firm-specific fixed effect, and ε_{it} captures the remaining unobserved (to the econometrician) characteristics of the firms. We use capital and labor to capture $\mathbf{x}_{it}$. Moreover, we assume that η_i^* and ε_{it} can be observed by other firms.

Similarly to Equation (1), the inverse demand function for firm i is given by:

$$p_{it} = \bar{\alpha}_m + \bar{\alpha}_t - q_{it} - \rho \sum_{j=1}^{n} b_{ij} q_{jt}, \tag{24}$$

where $b_{ij} = 1$ if i and j are in the same market and zero otherwise. In this equation, $\bar{\alpha}_m$ indicates the market-specific fixed effect and $\bar{\alpha}_t$ captures the time fixed effect due to exogenous demand shifters that affect consumer income, number of consumers (population), consumer taste and preferences and expectations over future prices of complements and substitutes or future income.

Denote by $\kappa_t \equiv \bar{\alpha}_t$ and $\eta_i \equiv \bar{\alpha}_m - \eta_i^*$. Observe that κ_t captures the time fixed effect while

η_i, which includes both $\bar{\alpha}_m$ and η_i^*, captures the firm fixed effect. Then, proceeding as in Section 3 (see, in particular the proof of Proposition 1), adding subscript t for time and using Equations (23) and (24), the econometric model equivalent to the best-response quantity in Equation (13) is given by:

$$q_{it} = \varphi \sum_{j=1}^{n} a_{ij,t} q_{jt} - \rho \sum_{j=1}^{n} b_{ij} q_{jt} + \mathbf{x}_{it}^{\top} \boldsymbol{\beta} + \eta_i + \kappa_t + \epsilon_{it}. \tag{25}$$

Observe that the econometric specification in Equation (25) has a similar specification as the product competition and technology spillover production function estimation in Bloom et al. [2013] where the estimation of φ will give the intensity of the *technology (or knowledge) spillover effect* of R&D, while the estimation of ρ will give the intensity of the *product rivalry effect*. However, as opposed to these authors, we explicitly take into account the technology spillovers stemming from R&D collaborations by using a network approach.

In vector-matrix form, we can write Equation (25) as

$$\mathbf{q}_t = \varphi \mathbf{A}_t \mathbf{q}_t - \rho \mathbf{B} \mathbf{q}_t + \mathbf{X}_t \boldsymbol{\beta} + \boldsymbol{\eta} + \kappa_t \mathbf{u}_n + \boldsymbol{\epsilon}_t, \tag{26}$$

where $\mathbf{q}_t = (q_{1t}, \cdots, q_{nt})^{\top}$, $\mathbf{A}_t = [a_{ij,t}]$, $\mathbf{B} = [b_{ij}]$, $\mathbf{X}_t = (\mathbf{x}_{1t}, \cdots, \mathbf{x}_{nt})^{\top}$, $\boldsymbol{\eta} = (\eta_1, \cdots, \eta_n)^{\top}$, $\boldsymbol{\epsilon}_t = (\epsilon_{1t}, \cdots, \epsilon_{nt})^{\top}$, and $\mathbf{u}_n$ is an n-dimensional vector of ones.

For the T periods, Equation (26) can be written as

$$\mathbf{q} = \varphi \mathrm{diag}\{\mathbf{A}_t\} \mathbf{q} - \rho (\mathbf{I}_T \otimes \mathbf{B}) \mathbf{q} + \mathbf{X} \boldsymbol{\beta} + \mathbf{u}_T \otimes \boldsymbol{\eta} + \boldsymbol{\kappa} \otimes \mathbf{u}_n + \boldsymbol{\epsilon}, \tag{27}$$

where $\mathbf{q} = (\mathbf{q}_1^{\top}, \cdots, \mathbf{q}_T^{\top})^{\top}$, $\mathbf{X} = (\mathbf{X}_1^{\top}, \cdots, \mathbf{X}_T^{\top})^{\top}$, $\boldsymbol{\kappa} = (\kappa_1, \cdots, \kappa_T)^{\top}$, and $\boldsymbol{\epsilon} = (\boldsymbol{\epsilon}_1^{\top}, \cdots, \boldsymbol{\epsilon}_T^{\top})^{\top}$, All vectors are of dimension $(nT \times 1)$, where T is the number of years available in the data.

In terms of data, our main variables will be measured as follows. Output q_{it} is the square root of net profits of firm i at time t (as explained in the previous section) and comes from the Compustat database. The network data comes from the CATI database and we set $a_{ij,t} = 1$ if there exists an R&D collaboration between firms i and j in the last s years before time t, where s is the duration of an alliance.[34] The exogenous variables captured by $\mathbf{x}_{it}$ are the firm's number of employees and the square root of capital. Finally, we measure b_{ij} as in the theoretical model so that $b_{ij} = 1$ if firms i and j are the same industry (measured by the industry SIC codes at the four-digit level) and zero otherwise.

8.2. Identification and IV Estimation

We here adopt a structural approach in the sense that we estimate exactly the first-order condition of the firm's maximization in terms of output and R&D effort, which lead to

[34]For the benchmark estimation results reported in Table 1, we set $s = 5$. We report estimation results with different lengths of alliance duration in Table 3 and the results are robust.

Equation (26). Interestingly, the best-response quantity in Equation (26) corresponds to a higher-order Spatial Auto-Regressive (SAR) model with two spatial lags $\mathbf{A}_t\mathbf{q}_t$ and $\mathbf{B}\mathbf{q}_t$ [Lee and Liu, 2010]. As in the SAR model, the spatial lags $\mathbf{A}_t\mathbf{q}_t$ and $\mathbf{B}\mathbf{q}_t$ are endogenous variables and need to be instrumented by $\mathbf{A}_t\mathbf{X}_t$ and $\mathbf{B}\mathbf{X}_t$.

To be more specific, let us consider Equation (25). The output of firm i at time t, q_{it}, is a function of the total output of all firms with an R&D collaboration with firm i at time t, i.e. $\bar{q}_{a,it} = \sum_{j=1}^{n} a_{ij,t}q_{jt}$, and the total output of all firms that operate in the same market as firm i, i.e. $\bar{q}_{b,it} = \sum_{j=1}^{n} b_{ij}q_{jt}$. Due the feedback effect, q_{jt} also depends on q_{it} and, thus, $\bar{q}_{a,it}$ and $\bar{q}_{b,it}$ are endogenous. We instrument $\bar{q}_{a,it}$ by the total number of employees and the total capital of all firms with an R&D collaboration with firm i, i.e. $\sum_{j=1}^{n} a_{ij,t}\mathbf{x}_{it}$, and instrument $\bar{q}_{b,it}$ by the total number of employees and the total capital of all firms that operate in the same industry as firm i, i.e. $\sum_{j=1}^{n} b_{ij}\mathbf{x}_{it}$. In other words, we estimate Equation (25) using a two-stage least squares (2SLS) approach where, in the first stage, we regress $\bar{q}_{a,it}$ and $\bar{q}_{b,it}$ on $\sum_{j=1}^{n} a_{ij,t}\mathbf{x}_{it}$ and $\sum_{j=1}^{n} b_{ij}\mathbf{x}_{it}$, respectively, to obtain $\hat{\bar{q}}_{a,it}$ and $\hat{\bar{q}}_{b,it}$. In the second stage of the estimation, we replace the spatial lags in Equation (25) by $\hat{\bar{q}}_{a,it}$ and $\hat{\bar{q}}_{b,it}$ and estimate

$$q_{it} = \varphi\hat{\bar{q}}_{a,it} - \rho\hat{\bar{q}}_{b,it} + \mathbf{x}_{it}^{\top}\boldsymbol{\beta} + \eta_i + \kappa_t + \epsilon_{it}. \tag{28}$$

Obviously, the above identification strategy based on IVs is valid only if $\mathbf{X}_t$ and $\mathbf{A}_t$ are exogenous. To control for the potential endogeneity of $\mathbf{X}_t$, we also experiment with IVs based on time-lagged employment and capital, e.g. $\mathbf{A}_t\mathbf{X}_{t-1}$ and $\mathbf{B}\mathbf{X}_{t-1}$, for $\mathbf{A}_t\mathbf{q}_t$ and $\mathbf{B}\mathbf{q}_t$ and the estimation results are robust. On the other hand, the potential endogeneity of $\mathbf{A}_t$ is somewhat more complicated to deal with. $\mathbf{A}_t$ is endogenous if there exists an *unobservable factor* that affects both q_{it} and $a_{ij,t}$. If the unobservable factor is firm-specific, then it is captured by the firm fixed-effect. If the unobservable factor is time-specific, then it is captured by the time fixed-effect. Therefore, the fixed effects in the panel data model are helpful for attenuating the potential endogeneity of $\mathbf{A}_t$.

Furthermore, to address the endogeneity of the adjacency matrix, we also model the network formation process of R&D collaborations between firms. That is, we consider an IV strategy based on the predicted adjacency matrix, i.e. $\hat{\mathbf{A}}_t\mathbf{X}_t$ following Kelejian and Piras [2012]. To be more specific, let us consider the estimation of Equation (25) using the predicted adjacency matrix by a three-stage least squares (3SLS) approach. In the first stage of the estimation, we obtain the predicted links $\hat{a}_{ij,t}$ from the regression of $a_{ij,t}$ on whether firms i and j collaborated before time $(t-s)$ where s is the duration of an alliance, whether i and j are in the same industry (measured by the first two digits of their SIC codes) and technological proximity of firms i and j represented by P_{ij} and P_{ij}^2 [cf. e.g. Nooteboom et al., 2006; Powell and Grodal, 2006, Sec. 3.5].[35,36] The proximity measure turns out to be a significant predictor for R&D collaborations.[37] In the second stage, we regress $\bar{q}_{a,it}$ on $\sum_{j=1}^{n} \hat{a}_{ij,t}\mathbf{x}_{jt}$ to obtain $\tilde{\bar{q}}_{a,it}$, and regress $\bar{q}_{b,it}$ on $\sum_{j=1}^{n} b_{ij}\mathbf{x}_{it}$ to obtain $\hat{\bar{q}}_{b,it}$. In the third stage, we replace the spatial lags in Equation (25) by $\tilde{\bar{q}}_{a,it}$ and $\hat{\bar{q}}_{b,it}$, respectively and estimate

$$q_{it} = \varphi\tilde{\bar{q}}_{a,it} - \rho\hat{\bar{q}}_{b,it} + \mathbf{x}_{it}^{\top}\boldsymbol{\beta} + \eta_i + \kappa_t + \epsilon_{it}. \tag{29}$$

Let us now give a formal definition for the estimator. In Equation (27), $\boldsymbol{\eta}$ and $\boldsymbol{\kappa}$ capture the firm and time fixed effects, respectively. We allow for $\boldsymbol{\eta}$ and $\boldsymbol{\kappa}$ to depend on diag$\{\mathbf{A}_t\}$, $\mathbf{B}$ and $\mathbf{X}$ by treating them as vectors of unknown parameters. To avoid the incidental problem, we transform Equation (27) using a within projector $\mathbf{J} = \mathbf{J}_T \otimes \mathbf{J}_n$ where $\mathbf{J}_T = \mathbf{I}_T - \frac{1}{T}\mathbf{u}_T\mathbf{u}_T^{\top}$ and $\mathbf{J}_n = \mathbf{I}_n - \frac{1}{n}\mathbf{u}_n\mathbf{u}_n^{\top}$. The transformed Equation (27) is

$$\mathbf{Jq} = \varphi\mathbf{J}\text{diag}\{\mathbf{A}_t\}\mathbf{q} - \rho\mathbf{J}(\mathbf{I}_T \otimes \mathbf{B})\mathbf{q} + \mathbf{JX}\boldsymbol{\beta} + \mathbf{J}\boldsymbol{\epsilon}. \tag{30}$$

where the firm and time fixed effects $\boldsymbol{\eta}$ and $\boldsymbol{\kappa}$ have been washed out.

As stated above, to estimate Equation (30), we consider the IV matrix with the actual adjacency matrix $\mathbf{A}_t$, i.e. $\mathbf{Q}_1 = \mathbf{J}[\text{diag}\{\mathbf{A}_t\}\mathbf{X}, (\mathbf{I}_T \otimes \mathbf{B})\mathbf{X}, \mathbf{X}]$, and the IV matrix with the predicted adjacency matrix $\hat{\mathbf{A}}_t = [\hat{a}_{ij,t}]$, i.e. $\mathbf{Q}_2 = \mathbf{J}[\text{diag}\{\hat{\mathbf{A}}_t\}\mathbf{X}, (\mathbf{I}_T \otimes \mathbf{B})\mathbf{X}, \mathbf{X}]$.

Let $\mathbf{P}_1 = \mathbf{Q}_1(\mathbf{Q}_1^{\top}\mathbf{Q})_1^{-1}\mathbf{Q}_1^{\top}$ $\mathbf{P}_2 = \mathbf{Q}_2(\mathbf{Q}_2^{\top}\mathbf{Q})_2^{-1}\mathbf{Q}_2^{\top}$ and $\mathbf{Z} = [\text{diag}\{\mathbf{A}_t\}\mathbf{q}, (\mathbf{I}_T \otimes \mathbf{B})\mathbf{q}, \mathbf{X}]$. The 2SLS estimator with IVs based on the actual adjacency matrix is given by $(\mathbf{Z}^{\top}\mathbf{P}_1\mathbf{Z})^{-1}\mathbf{Z}^{\top}\mathbf{P}_1\mathbf{q}$. The 3SLS estimator with IVs based on the predicted adjacency matrix is given by $(\mathbf{Z}^{\top}\mathbf{P}_2\mathbf{Z})^{-1}\mathbf{Z}^{\top}\mathbf{P}_2\mathbf{q}$. With the estimates of $\varphi, \rho, \boldsymbol{\beta}$, we can recover $\boldsymbol{\eta}$ and $\boldsymbol{\kappa}$ by the least squares dummy variable method.

---

[35] We matched the firms in our alliance data with the owners of patents from the US patent office (USPTO) and the European patent office (EPO). For the US patents, we matched the firm names with the firms in the NBER patent data project, while for the European patents, we matched them to the firms listed in the EPO Worldwide Patent Statistical Database (PATSTAT). This allowed us to obtain the number of patents and the patent portfolio held for about 30 % of the firms in the data. From the firms' patents, we then computed their technological proximity $P_{ij} = \frac{\mathbf{F}_i^{\top}\mathbf{F}_j}{\sqrt{\mathbf{F}_i^{\top}\mathbf{F}_i}\sqrt{\mathbf{F}_j^{\top}\mathbf{F}_j}}$, where $\mathbf{F}_i$ is a vector whose k-th component F_{ik} counts the number of patents firm i has in technology category k divided by the total number of technologies attributed to the firm [cf. Bloom et al., 2013; Jaffe, 1986]. We used the three-digit US patent classification system to identify technology categories [Hall et al., 2001].

[36] A year-by-year linear probability regression is carried out. The estimated coefficients of the regressors are all statistically significant with expected signs. In particular, the coefficient of P_{ij} is positive and that of P_{ij}^2 is negative. The estimation results of this first stage are omitted to save space and are available upon request.

[37] One would therefore expect to obtain similar results using patents as a measure for technology spillovers (instead of explicit data on R&D collaborations) as in Bloom et al. [2013]. Moreover, in Section 9.2.4 we allow for both, direct technology spillover from R&D collaborations, and indirect spillovers captured via the patent proximity matrix P_{ij}.

Table 1: Parameter estimates (with standard errors in parenthesis) from a panel regression with time dummies of Equation (26). Model A does not include firm fixed effects (f.e.), while Model B introduces also firm fixed effects. Model C uses the predicted instead of the actual adjacency matrix.

	Model A		Model B		Model C	
time eff.	yes		yes		yes	
firm f.e.	no		yes		yes	
φ	0.0214***	(0.0015)	0.0077***	(0.0013)	0.0077***	(0.0019)
ρ	0.0019***	(0.0001)	0.0008***	(0.0002)	0.0008***	(0.0002)
β_1	0.0856***	(0.0072)	0.1249***	(0.0119)	0.1248***	(0.0119)
β_2	0.4617***	(0.0064)	0.5280***	(0.0116)	0.5281***	(0.0118)

*** Statistically significant at 1% level.
** Statistically significant at 5% level.
* Statistically significant at 10% level.

9. Empirical Results

9.1. Results from Estimating our Model

The parameter estimates of Equation (26) are reported in Table 1, which reports three different models. Models A and B are the 2SLS estimation of Equation (26) with time fixed effects only and time and firm fixed effects, respectively. Model C is the 3SLS estimation of Equation (26) where we estimate the adjacency matrix (network formation) in the first stage. In all these models, we obtain the expected signs, that is the *technology (or knowledge) spillover effect* (estimate of φ) always has a *positive* impact on own output while the *product rivalry effect* (estimate of $-\rho$) always has *negative* impact on own output. Indeed, the more a given firm collaborates with other firms in R&D, the higher is its output production. This indicates that R&D by allied firms in the network is associated with higher product value and indicate that there are strategic complementarities between own and allied firms. However, conditional on technology spillovers, the more firms that compete in the same market, the lower is the production of the good by the given firm. As in Bloom et al. [2013], this table shows that the magnitude of the first effect (technology spillover) is much higher than that of the second effect (product rivalry). Keeping all other firms' output levels constant, suppose that firm j is both a collaboration partner of firm i and operates in the same market as firm i. Then, we find that the net effect of firm j increasing its output by one unit is captured by the difference of the two effects. As the technology spillover effect is much higher than the rivalry effect, we find that the latter dominates the former so that the net returns to R&D collaborations are strictly positive. Furthermore, this table also shows that capital and labor have a positive and significant impact on own output.

Table 2: Parameter estimates (with standard errors in parenthesis) from a panel regression with time dummies of Equation (26). Model A does not include firm fixed effects (f.e.), while Model B introduces also firm fixed effects. Model C uses the predicted instead of the actual adjacency matrix. Firms for which one of the neighbors has missing data are dropped from the sample.

	Model A		Model B		Model C	
time eff.	yes		yes		yes	
firm f.e.	no		yes		yes	
φ	0.0174***	(0.0022)	0.0048**	(0.0024)	0.0242***	(0.0051)
ρ	0.0021***	(0.0001)	0.0014***	(0.0002)	0.0014***	(0.0002)
β_1	0.1325***	(0.0113)	0.1758***	(0.0108)	0.1766***	(0.0109)
β_2	0.4300***	(0.0079)	0.5073***	(0.0123)	0.4966***	(0.0128)

*** Statistically significant at 1% level.
** Statistically significant at 5% level.
* Statistically significant at 10% level.

9.2. Robustness Checks

9.2.1. Missing Data

The actual number of observations used in the estimation is much lower than the 13,040 companies in the MERIT-CATI database due to missing data in the dependent variables q_{it} and controls $\mathbf{x}_{it}$. The presence of missing data does not only introduce some technical difficulty since the panel is unbalanced,[38] but may also lead to more severe consequences. Suppose that $a_{ij,t} = 1$ and the observation on q_{jt} is missing. Then, the missing observation introduces a measurement error to the spatial lag $\sum_{j=1}^{n} a_{ij,t}q_{jt}$ in Equation (25) and the above estimators may not be consistent.[39] Note that this is a different sampling issue from that studied by Chandrasekhar and Lewis [2011] and Liu [2013], where the dependent variable and controls can be observed and the observations on network links might be missing. This missing data issue is more in line with that in Wang and Lee [2013]. However, the method in Wang and Lee [2013] cannot be applied here as they consider a random-effect panel data model rather than a fixed-effect model and they assume that there is no missing data in the control variables.

As a robustness check, we estimate our model using the subsample of firms whose collaboration partners have no missing outputs. For those firms, the collaboration effect is correctly specified. The estimation results are reported in Table 2 and are similar with respect to those reported in Table 1, except for an increase in the spillover coefficient from R&D collaborations.

[38]For notational simplicity, we present the estimation procedure for a balanced panel. For an unbalanced panel due to missing data, the projector introduced by Wansbeek and Kapteyn [1989] can easily be extended to the current model to eliminate the individual and time fixed effects.

[39]The missing observation of an individual firm output is less a concern for the product rivalry effect because we directly use the industry-level total output from the data to get $\sum_{j=1}^{n} b_{ij}q_{jt}$.

9.2.2. Time Span of Alliances

We here analyze the impact of considering different time spans (other than 5 years as in the previous section) for the duration of an alliance. The estimation results from Table 1 in Section 8.2 for alliance durations ranging from 3 to 7 years are shown in Table 3. We find that the estimates are robust over the different durations considered.

9.2.3. Intra- versus Interindustry Collaborations

So far, we have assumed that network effects or knowledge spillovers were the same whether they were intra- or inter-industry collaborations. In the real-world, the knowledge spillovers between two firms in the same industry (say Volvo and Honda in the car manufacturing sector) may be different than between two firms from different industries (for example, between Volvo and Toshiba in the car manufacturing and ICT sectors, respectively). The rationale is that the involved firms might differ in the similarity of their areas of technological competences and knowledge domains depending on whether the collaborating firms operate in the same or in different industries [cf. Nooteboom et al., 2006; Powell and Grodal, 2006].[40]

In this section, we extend our empirical model of Equation (25) by allowing for intra-industry technology spillovers to differ from inter-industry spillovers. The generalized model is given by[41]

$$q_{it} = \varphi_1 \sum_{j=1}^{n} a_{ij,t}^{(1)} q_{jt} + \varphi_2 \sum_{j=1}^{n} a_{ij,t}^{(2)} q_{jt} - \rho \sum_{j=1}^{n} b_{ij} q_{jt} + \mathbf{x}_{it}^{\top} \boldsymbol{\beta} + \eta_i + \kappa_t + \epsilon_{it}, \tag{31}$$

where $a_{ij,t}^{(1)} = a_{ij,t} b_{ij}$, $a_{ij,t}^{(2)} = a_{ij,t}(1 - b_{ij})$, and the coefficients φ_1 and φ_2 capture the intra-industry and the inter-industry technology spillover effect, respectively. In vector-matrix form, we have:

$$\mathbf{q}_t = \varphi_1 \mathbf{A}_t^{(1)} \mathbf{q}_t + \varphi_2 \mathbf{A}_t^{(2)} \mathbf{q}_t - \rho \mathbf{B} \mathbf{q}_t + \mathbf{X}_t \boldsymbol{\beta} + \boldsymbol{\eta} + \kappa_t \mathbf{u}_n + \boldsymbol{\epsilon}_t. \tag{32}$$

The parameter estimates from a fixed-effect panel regression with time dummies of Equation (32) are given in Table 4. We observe that the signs and the significance (except for the within industry spillovers) of the coefficients remain the same as before. Interestingly, the inter-industry R&D spillover coefficient is significant, while the intra-industry R&D spillover coefficient becomes insignificant. This highlights the importance of technology spillovers from firms in different industries driven by dissimilarities in their technology portfolios and the recombination of heterogeneous technologies for innovation [cf. Weitzman, 1998].

[40] This specification also allows for testing the possibility that allied firms which operate in the same market might form a collusive agreement and thus affect each other's quantity levels differently than firms operating in different markets [cf. Duso et al., 2012; Goeree and Helland, 2012].

[41] The theoretical foundation of Equation (31) can be found in Appendix D.

Table 3: Parameter estimates (with standard errors in parenthesis) from a panel regression with time dummies of Equation (26) including firm fixed effects assuming different durations of an alliance ranging from 3 to 7 years.

	3 years		4 years		5 years		6 years		7 years	
φ	0.0089***	(0.0016)	0.0084***	(0.0014)	0.0077***	(0.0013)	0.0075***	(0.0012)	0.0073***	(0.0012)
ρ	0.0008***	(0.0002)	0.0008***	(0.0002)	0.0008***	(0.0002)	0.0008***	(0.0002)	0.0008***	(0.0002)
β_1	0.1271***	(0.0116)	0.1259***	(0.0117)	0.1249***	(0.0119)	0.1240***	(0.0119)	0.1233***	(0.0120)
β_2	0.5267***	(0.0113)	0.5271***	(0.0113)	0.5280***	(0.0116)	0.5275***	(0.0117)	0.5275***	(0.0119)
# obs.	25166		25098		25027		24946		24862	

*** Statistically significant at 1% level.

** Statistically significant at 5% level.

* Statistically significant at 10% level.

Table 4: Parameter estimates (with standard errors in parenthesis) from a fixed effects panel regression with time dummies of Equation (32). Model E does not include firm fixed effects (f.e.), while Model F introduces also firm fixed effects.

	Model E		Model F	
time eff.	yes		yes	
firm f.e.	no		yes	
φ_1	0.0052	(0.0048)	0.0081	(0.0060)
φ_2	0.0235***	(0.0015)	0.0078***	(0.0015)
ρ	0.0018***	(0.0001)	0.0008***	(0.0002)
β_1	0.0862***	(0.0072)	0.1248***	(0.0111)
β_2	0.4634***	(0.0065)	0.5278***	(0.0111)

*** Statistically significant at 1% level.
** Statistically significant at 5% level.
* Statistically significant at 10% level.

9.2.4. Direct and Indirect Technology Spillovers

So far, our adjacency matrix $\mathbf{A}_t$ only captured R&D collaborations between firms, which, in the data, where measured by our MERIT-CATI database. As stated in the Introduction, a lot of R&D spillovers do not take place through R&D collaborations and can come from many channels such as reading other firms' patents or imitating their products. In the present section, we extend our empirical model of Equation (25) by allowing for both direct (between collaborating firms) and indirect technology spillovers (between non-collaborating firms), which will be captured by the matrices $\mathbf{A}_t$ and $\mathbf{W}_t$, respectively.

The generalized empirical model is now given by[42]

$$q_{it} = \varphi \sum_{j=1}^{n} a_{ij,t} q_{jt} + \chi \sum_{j=1}^{n} w_{ij,t} q_{jt} - \rho \sum_{j=1}^{n} b_{ij} q_{jt} + \mathbf{x}_{it}^{\top} \boldsymbol{\beta} + \eta_i + \kappa_t + \epsilon_{it}, \tag{33}$$

where w_{ij} are weights characterizing alternative channels for technology spillovers than R&D collaborations and the coefficients φ and χ capture the direct and the indirect technology spillover effect, respectively. In vector-matrix form, we then have:

$$\mathbf{q}_t = \varphi \mathbf{A}_t \mathbf{q}_t + \chi \mathbf{W}_t \mathbf{q}_t - \rho \mathbf{B} \mathbf{q}_t + \mathbf{X}_t \boldsymbol{\beta} + \boldsymbol{\eta} + \kappa_t \mathbf{u}_n + \boldsymbol{\epsilon}_t. \tag{34}$$

In terms of data, we measure the weights w_{ij} by the cells of the technological proximity matrix, P_{ij}, introduced already in Footnote 35. Recall that we matched the firms in our alliance data with the patents from the US patent office (USPTO) and the European patent office (EPO). From the firms' patents, we then computed their technological proximity P_{ij},

[42]The theoretical foundation of Equation (33) can be found in Appendix E.

Table 5: Parameter estimates (with standard errors in parenthesis) from a panel regression with time dummies of Equation (34). Model G does not include firm fixed effects (f.e.), while Model H introduces also firm fixed effects. Model I uses the predicted instead of the actual adjacency matrix. Firms for which one of the neighbors has missing data are dropped from the sample.

	Model G		Model H		Model I	
time eff.	yes		yes		yes	
firm f.e.	no		yes		yes	
φ	0.0149***	(0.0022)	0.0039*	(0.0023)	0.0211***	(0.0051)
χ	0.0045***	(0.0003)	0.0029***	(0.0004)	0.0023***	(0.0005)
ρ	0.0021***	(0.0001)	0.0015***	(0.0002)	0.0014***	(0.0002)
β_1	0.1306***	(0.0111)	0.1764***	(0.0108)	0.1769***	(0.0109)
β_2	0.4276***	(0.0078)	0.5064***	(0.0124)	0.4972***	(0.0128)

*** Statistically significant at 1% level.
** Statistically significant at 5% level.
* Statistically significant at 10% level.

defined by

$$P_{ij} = \frac{\mathbf{F}_i^\top \mathbf{F}_j}{\sqrt{\mathbf{F}_i^\top \mathbf{F}_i}\sqrt{\mathbf{F}_j^\top \mathbf{F}_j}}, \tag{35}$$

where $\mathbf{F}_i$ is a vector whose k-th component F_{ik} counts the number of patents firm i has in technology category k divided by the total number of technologies attributed to the firm [cf. Bloom et al., 2013; Jaffe, 1986]. We used the three-digit US patent classification system to identify technology categories [cf. Hall et al., 2001].

The results of a fixed-effect panel regression with time dummies of Equation (34) are shown in Table 5. Both spillover coefficients, φ and χ, are positive and significant. We further observe that the spillover effect from direct R&D collaborations (φ) is also typically higher than the one from indirect technology spillovers (χ). We will use the most conservative estimation in the last column in Table 5, Model I, for our policy analysis in Section 10, where we use the predicted instead of the actual adjacency matrix, firms for which one of the neighbors has missing data are dropped from the sample, and we allow for both direct and indirect technology spillovers.

10. Policy Results

In the following sections, we use our estimation results from the most general model of Section 9.2.4 to analyze the different policies highlighted in Section 5 (key-player policy) and Section 6 (subsidy policy).

10.1. Determining the Key Firms

Now that we have estimated the parameters of the model, using the most conservative regression in Table 5, Model I, together with the results of Section 5, we can calculate the intercentrality of each firm in our dataset.[43] The corresponding formula is given in part (ii) of Proposition 5. This will determine the key players or the key firms in our dataset.[44] Therefore, we can rank the firms according to their intercentrality measures. This means that the firm that will be ranked first is such that, if it exits from the market, it will generate the highest loss in total welfare in the economy. The firm ranked second will be the one that, if removed, will generate the second highest loss in total welfare, etc.. Remember that the key-player policy is a short-term one as we do not take into account the possible reaction of firms in terms of network formation. This is why we calculate for each year a new list of key players.

A ranking of the first 25 firms with the highest impact on welfare upon exit in the year 1990 can be found in Table 6, while the corresponding ranking in the year 2005 is shown in Table 7. In these two tables, we also calculate the market share of each firm in the primary four-digit sector where it operates (across all firms in the Compustat database within that sector), the number of patents accumulated, its degree $\mathbf{d}$ (i.e. the number of R&D collaborations), its eigenvector centrality $\mathbf{v}_{\mathrm{PF}}$, its betweenness and closeness centralities (see Wasserman and Faust [1994] and Jackson [2008] for a list and definitions of these and other centrality measures), the relative output or Bonacich centrality of the firm, the key player according to Ballester et al. [2006, 2010] (i.e. the firm which once removed generates the highest decrease in total activity/output) and, finally, the key player defined in the present paper (i.e. the firm which once removed generates the highest decrease in total welfare).

It should be clear that key firms are not always those with the highest centralities. If we look, for example, at Table 6, then one can see that the key firm is *General Motors* but it is not the one with the largest number of R&D collaborations (degree), number of patents, nor the highest eigenvector, betweenness or closeness centrality. More importantly, *General Motors* is not the firm that has the highest market share in its sector since it has "only" 12.14 % of the market share while, for example, *Hitachi*, *Altria* or *Pepsico* have a much higher share (up to more than 50 %). This means that it is not straightforward to determine which firm should be "targeted" in the network by only observing its market share, size or even its position in the network. Interestingly, our intercentrality and that of Ballester et al. [2006, 2010] give roughly the same rankings of firms. If *General Motors* were to be removed from the market, then total welfare would be reduced by 8.14 %, while total output would decrease by 2.13 %. As stated in Section 1, the bailout of *General Motors* by president Obama was a success because of the indirect effects of a possible bankrupcy of this automoblie company on other companies in the industry. These spillover effects are the main reason for why *General Motors* is a key firm in our analysis. If *Sony* (which has 32 % of the market share of its market) or *Procter and Gamble* (which has nearly 59 % of the market share of its market)

[43]Note that we use both, the matrix $\mathbf{A} = (a_{ij})_{1\leq i,j\leq n}$ of R&D collaborations and the matrix $\mathbf{W} = (w_{ij})_{1\leq i,j\leq n}$ capturing indirect technology spillovers for our policy analysis, as discussed in Section 9.2.4.

[44]If some firms turn out to have non-positive output levels in a counterfactual equilibrium after the removal of a firm, then they are assumed to exit the market.

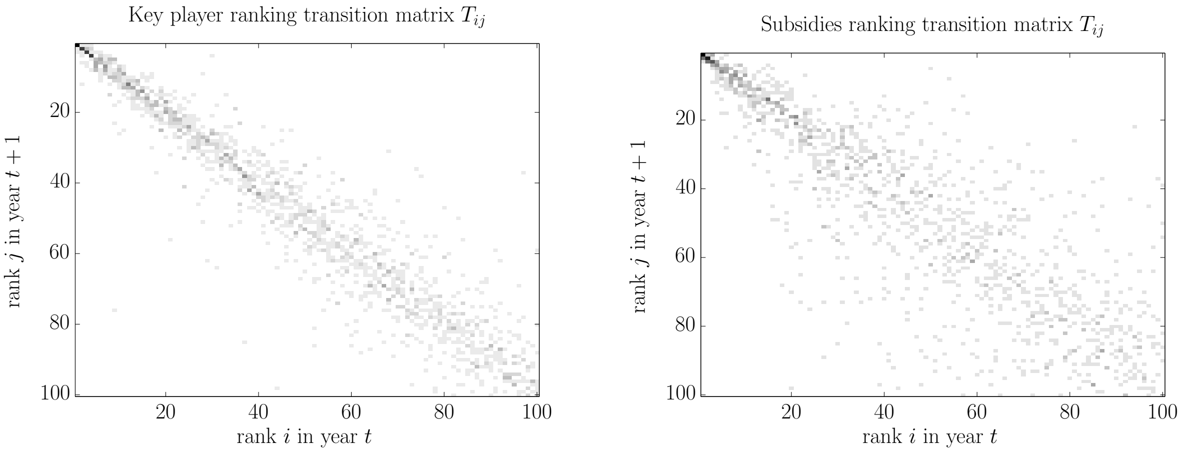


Figure 6: The transition matrix T_{ij} from the rank i in year t to the rank j in year $t+1$ for the key player ranking (left panel) and the subsidies ranking (right panel) for the first 100 ranks.

were removed from the economy, then less than a 1.7 % decrease in output or welfare would follow.

If we now compare the key player ranking between 1990 and 2005 (15 years after), then, from Tables 6 and 7, we find that key firms change over time. For example, *General Motors*, which was the key firm in 1990, is ranked seventh in 2005 and its removal will reduce welfare by 1.83 % and total output by 0.68 %, while these numbers were 8.14 % and 2.13 % in 1990. More generally, it can be seen that the decline in welfare and total output due to the removal of the highest ranked firms is generally much lower in 2005. Apart from the fact that some key firms in 1990 are no longer present in 2005 (for example, *Texaco*, *Hoechst A.G.*, *Elf Aquitaine*), most key firms are still "key" in 2005. Figure 7 captures this idea by showing the change in the ranking of the 25 highest ranked firms from 1990 to 2005. The ranking of firms can be quite stable for some, while it is rather volatile for others. For example, *Exxon Corp.* was the second highest ranked firm in 1990 and occupies the third place in 2005. In contrast, *Hoechst A.G.*, which was among the three highest ranked firms in 1990, slipped down to rank 82 in 2003. The left-hand panel in Figure 6 shows the transition probability T_{ij} from a rank i in year t to a rank j in year $t+1$ for the first 100 ranked firms. The figure illustrates that the rankings are quite stable over time, where most transitions occur along the diagonal of T_{ij}. There is a larger variation in the bottom right corner of T_{ij} and less variation in the top left corner. This shows that the upper ranks are more stable than the lower ranks.

The left-hand panel of Figure 8 shows the (ordered) percentage decrease in welfare due to the removal of a firm over the years 1990 to 2005. The exit of most firms only has a minor impact on welfare, while the highest ranked firms can have a considerable effect on total welfare.

10.2. R&D Subsidies

As an alternative policy to the key player analysis in the previous section, we now study empirically the optimal subsidy policy, both for the homogenous subsidy, s^*, (see Proposition

Table 6: Key player ranking for the year 1990 for the first 25 firms.

Firm	Share [%][a]	num. pat.	$\mathbf{d}$	$\mathbf{v}_{\mathrm{PF}}(G)$	Betweenness[b]	Closeness[c]	$q_i/\lVert\mathbf{q}\rVert_1$ [%][d]	$\frac{\lVert\mathbf{q}(G)\rVert_1-\lVert\mathbf{q}(G^{-i})\rVert_1}{\lVert\mathbf{q}(G)\rVert_1}$ [%][e]	$\frac{W(G)-W(G^{-i})}{W(G)}$ [%]	Rank
General Motors Corp.	12.1445	50185	12	0.1185	0.3381	0.0025	1.5952	2.1320	8.1366	1
Exxon Corp.	10.1151	6927	3	0.0187	0.0000	0.0015	1.3351	1.4587	5.2459	2
Hoechst A.G.	13.8715	9634	2	0.0000	0.0006	0.0002	1.1600	1.2474	3.6020	3
Altria Group	57.0787	0	0	0.0000	0.0000	0.0000	0.9675	0.9675	2.4648	4
Chevron	3.7009	4410	5	0.0191	0.0430	0.0015	0.8618	0.9773	2.3855	5
Pepsico Inc.	52.5069	798	0	0.0000	0.0000	0.0000	0.9035	0.9203	2.1723	6
Unilever N.V./Plc.	8.2910	982	0	0.0000	0.0000	0.0000	0.8933	0.9300	2.1323	7
Daimler Corp.	5.2310	2	0	0.0000	0.0000	0.0000	0.8931	0.8877	2.0879	8
Texaco Inc.	3.9206	8939	6	0.0427	0.1857	0.0019	0.7446	0.8666	1.8826	9
Toyota Motor Corp.	6.2806	58	9	0.1578	0.1420	0.0023	0.6746	0.8293	1.6732	10
Sony Corp.	32.0711	5840	8	0.1021	0.0382	0.0018	0.7243	0.8922	1.6603	11
Motorola Inc.	18.5193	7903	16	0.2996	0.1027	0.0025	0.6349	0.9300	1.6273	12
Hitachi Ltd.	37.6873	40838	7	0.1309	0.0141	0.0020	0.6840	0.8554	1.5681	13
Bellsouth Corp.	3.2244	42	2	0.0059	0.0000	0.0009	0.7353	0.7782	1.5395	14
McDonnell Douglas Corp.	21.8941	899	14	0.1851	0.1452	0.0024	0.6194	0.8771	1.5056	15
Alcatel-Lucent	31.0329	1238	0	0.0000	0.0000	0.0000	0.7462	0.7683	1.4979	16
Renault	2.9712	524	2	0.0034	0.0000	0.0009	0.7250	0.7481	1.4789	17
Merrill Lynch Inc.	13.1555	8	4	0.0206	0.0257	0.0014	0.7033	0.7689	1.4645	18
Volkswagen A.G.	4.1641	414	4	0.0184	0.0344	0.0014	0.6855	0.7554	1.4348	19
Xerox Corp.	84.2264	24341	8	0.1378	0.0246	0.0020	0.6350	0.7820	1.2918	20
Procter & Gamble	58.8860	14744	1	0.0000	0.0000	0.0004	0.6888	0.7208	1.2861	21
Texas Instruments Inc.	20.5932	14822	22	0.3450	0.2083	0.0028	0.5342	0.8365	1.2766	22
Volvo A.B.	1.3887	119	4	0.0147	0.0395	0.0014	0.6205	0.6808	1.2521	23
Elf Aquitaine	3.1007	2471	1	0.0000	0.0000	0.0002	0.6675	0.7056	1.2442	24
Novartis AG	4.6058	311	0	0.0000	0.0000	0.0000	0.6518	0.6472	1.1257	25

[a] Market share in the primary 4-digit sector in which the firm is operating.

[b] The normalized betweenness centrality is the fraction of all shortest paths in the network that contain a given node, divided by $(n-1)(n-2)$, the maximum number of such paths.

[c] The closeness centrality of node i is computed as $\sum_{j=1}^{n} 2^{-\ell_{ij}(G)}$, where $\ell_{ij}(G)$ is the length of the shortest path between i and j in the network G [Dangalchev, 2006].

[d] The relative output of a firm i is computed as $q_i/\lVert\mathbf{q}\rVert_1 = b_{\boldsymbol{\mu},i}/\lVert\mathbf{b}_{\boldsymbol{\mu}}\rVert_1$ (see Proposition 1).

[e] The decrease in output due to the removal of firm i is computed as $\frac{\lVert\mathbf{q}(G)\rVert_1-\lVert\mathbf{q}(G^{-i})\rVert_1}{\lVert\mathbf{q}(G)\rVert_1} = \frac{b_{\mathbf{u},i}(G)b_{\boldsymbol{\mu},i}(G)}{m_{ii}(G)}/\lVert\mathbf{b}_{\boldsymbol{\mu}}(G)\rVert_1$.

Table 7: Key player ranking for the year 2005 for the first 25 firms.

Firm	Share [%][a]	num. pat.	$\mathbf{d}$	$\mathbf{v}_{\mathrm{PF}}(G)$	Betweenness[b]	Closeness[c]	$q_i/\Vert\mathbf{q}\Vert_1$ [%][d]	$\frac{\Vert\mathbf{q}(G)\Vert_1-\Vert\mathbf{q}(G^{-i})\Vert_1}{\Vert\mathbf{q}(G)\Vert_1}$ [%][e]	$\frac{W(G)-W(G^{-i})}{W(G)}$ [%]	Rank
Daimler Corp.	7.5743	9952	9	0.0093	0.1239	0.0009	0.8474	1.0518	3.8156	1
Toyota Motor Corp.	7.7760	567	4	0.0036	0.0019	0.0006	0.8056	0.8829	3.3055	2
Exxon Corp.	7.8647	11457	0	0.0000	0.0000	0.0000	0.8478	0.8744	3.2712	3
NTT DoCoMo	4.3962	1505	7	0.2752	0.1323	0.0016	0.6197	0.7763	2.1623	4
Volkswagen A.G.	4.8178	3931	3	0.0023	0.0000	0.0006	0.6201	0.6750	1.9932	5
Intel Corp.	9.8341	31709	15	0.3040	0.1593	0.0016	0.5607	0.7691	1.9605	6
General Motors Corp.	7.7341	66784	6	0.0101	0.0480	0.0009	0.5757	0.6823	1.8318	7
Sony Corp.	32.1340	32362	12	0.3462	0.1095	0.0015	0.5436	0.7413	1.8186	8
Chevron	4.4312	4987	0	0.0000	0.0000	0.0000	0.6201	0.6345	1.7510	9
Hitachi Ltd.	27.8692	106477	12	0.3199	0.1439	0.0015	0.5309	0.7532	1.7308	10
Total SA	3.6544	1205	0	0.0000	0.0000	0.0000	0.6107	0.6415	1.7153	11
Altria Group	40.0416	0	0	0.0000	0.0000	0.0000	0.5546	0.5546	1.3849	12
Fujitsu Ltd.	17.3622	26830	14	0.3722	0.1840	0.0016	0.3941	0.5520	1.1174	13
Unilever N.V./Plc.	9.0941	6697	0	0.0000	0.0000	0.0000	0.4910	0.5071	1.1033	14
Metro AG	17.6754	1	2	0.0543	0.0000	0.0009	0.4574	0.4842	1.0323	15
Endesa	1.5322	5	0	0.0000	0.0000	0.0000	0.4501	0.4775	0.9436	16
Sanofi	4.7555	7655	0	0.0000	0.0000	0.0000	0.4460	0.4340	0.9004	17
Tyco	78.0977	7517	0	0.0000	0.0000	0.0000	0.4388	0.4625	0.8971	18
Novartis AG	4.5096	10352	4	0.0048	0.1026	0.0008	0.4362	0.4842	0.8940	19
Procter & Gamble	54.0431	51544	1	0.0042	0.0000	0.0006	0.4345	0.4600	0.8915	20
Canon Inc.	28.2269	105340	1	0.0303	0.0000	0.0009	0.4269	0.4662	0.8826	21
Time Warner Inc	31.5018	242	0	0.0000	0.0000	0.0000	0.4335	0.4481	0.8637	22
Bellsouth Corp.	1.8106	4060	0	0.0000	0.0000	0.0000	0.4203	0.4354	0.8162	23
Comcast Corp	16.9208	0	3	0.0534	0.0176	0.0012	0.4042	0.4352	0.8128	24
Telstra Corp	1.4888	40	0	0.0000	0.0000	0.0000	0.4084	0.4219	0.7693	25

[a] Market share in the primary 4-digit sector in which the firm is operating.

[b] The normalized betweenness centrality is the fraction of all shortest paths in the network that contain a given node, divided by $(n-1)(n-2)$, the maximum number of such paths.

[c] The closeness centrality of node i is computed as $\sum_{j=1}^{n} 2^{-\ell_{ij}(G)}$, where $\ell_{ij}(G)$ is the length of the shortest path between i and j in the network G [Dangalchev, 2006].

[d] The relative output of a firm i is computed as $q_i/\Vert\mathbf{q}\Vert_1 = b_{\boldsymbol{\mu},i}/\Vert\mathbf{b}_{\boldsymbol{\mu}}\Vert_1$ (see Proposition 1).

[e] The decrease in output due to the removal of firm i is computed as $\frac{\Vert\mathbf{q}(G)\Vert_1-\Vert\mathbf{q}(G^{-i})\Vert_1}{\Vert\mathbf{q}(G)\Vert_1} = \frac{b_{\mathbf{u},i}(G)b_{\boldsymbol{\mu},i}(G)}{m_{ii}(G)}/\Vert\mathbf{b}_{\boldsymbol{\mu}}(G)\Vert_1$.

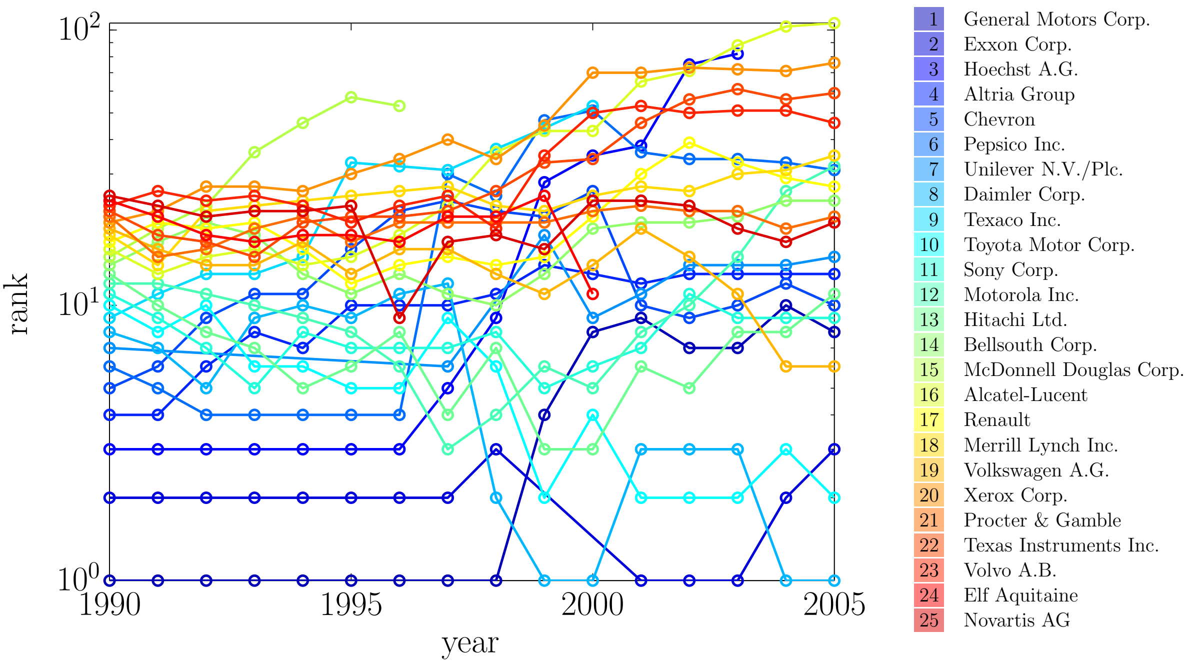


Figure 7: Change in ranking of the 25 key firms (Table 6) from the year 1990 to the year 2005.

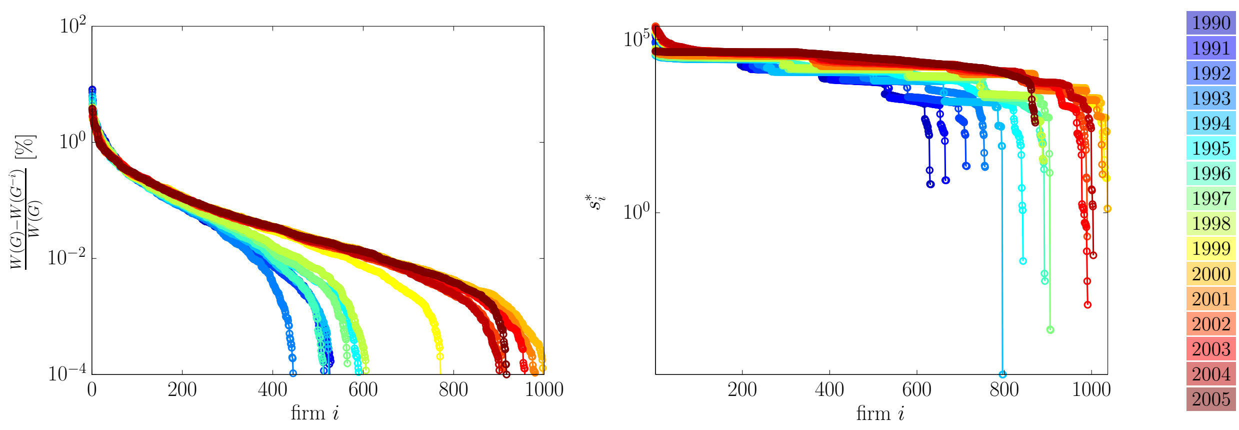


Figure 8: (Left panel) The ordered percentage decrease in welfare due to the removal of firm i over the years 1990 to 2005. (Right panel) The ordered targeted subsidy level of firm i over the years 1990 to 2005.

6) and for the targeted subsidy, $\mathbf{s}^*$ (see Proposition 7).[45] In Figure 9, in the top panel, we calculate the optimal homogenous subsidy times R&D effort over time (top left panel) and the percentage increase in welfare due to the homogenous subsidy over time (top right panel). Interestingly, the total subsidized R&D effort more than doubled over the time between 1990 and 2005. In terms of welfare, the highest increase (around 46 %) is in 2000 and 2005, while the increase in welfare in 1995 is smaller (around 37 %). The bottom panel of Figure 9 does the same exercise for the targeted subsidy policy.[46] We first note that in this case we find that while most of the firms receive a subsidy, a few are actually taxed (always below 0.5 % across all years). Moreover, the total expenditures on the targeted subsidies are typically of the same order of magnitude as the ones for the homogeneous subsidy. However, the targeted subsidy program turns out to have a much higher impact on total welfare, as it can improve welfare by up to 120 % while the homogeneous subsidies can improve total welfare only by up to 46 %. Moreover, the optimal subsidy levels show a strong variation over time. Both the homogeneous and the aggregate targeted subsidy seem to follow a cyclical trend, similar to the strong variation we have observed for the number of firms participating in R&D collaborations in a given year in Figure 3. This cyclical trend is also reminiscent of the R&D expenditures observed in the empirical literature on business cycles [cf. Barlevy, 2007; Galí, 1999].

We can compare the optimal subsidy level predicted from our model with the R&D tax subsidies actually implemented in the United States and selected other countries between 1979 to 1997 [see Bloom et al., 2002; Impullitti, 2010]. While these time series typically show a steady increase of R&D subsidies over time, they do not seem to incorporate the cyclicality that we obtain for the optimal subsidy levels. Our analysis thus suggests that policy makers should adjust R&D subsidies to these cycles.

At the firm level we can further compare our firm-specific optimal subsidies with those that are actually provided by government agencies. For this purpose we have matched the firms in our dataset with the firms that have obtained R&D subsidies from the European intergovernmental organization for market-driven industrial R&D, EUREKA.[47] A ranking of the first 25 firms according to our optimal subsidy policy considering only those that received funding from EUREKA is shown in Table 10. We observe that the ranking of our subsidy policy does not necessarily reflect the ranking of the actual subsidies implemented by EUREKA. For example, *Fujitsu Ltd.* received funding of 0.96 millions USD and is ranked second according to our optimal subsidy policy, while *Sony Corp.* received funding of only 488.1916 millions USD while being ranked third, behind *Fujitsu Ltd.*. However, this discrep-

[45]As in the case of the key player analysis, we adopt the convention that if some firms turn out to have non-positive output levels in a counterfactual equilibrium where the subsidy policy is implemented, then they are assumed to exit the market.

[46]We find that the condition for positive definiteness in the case of a targeted subsidy in part (iii) of Proposition 7 was violated in our data. Hence, our subsidy policy yields a lower bound on the potential welfare gains. See also Footnote 25.

[47]See `http://www.eurekanetwork.org/`.

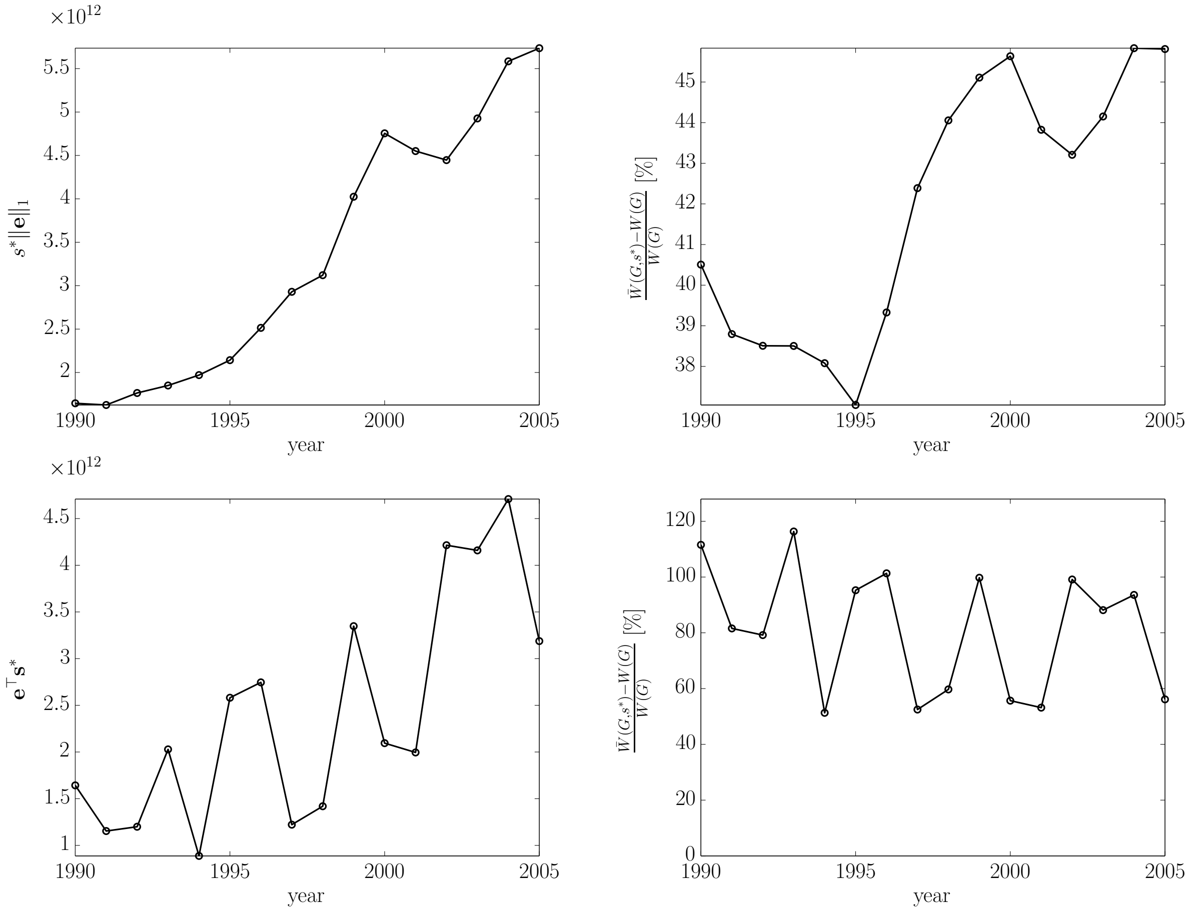


Figure 9: (Top left panel) The total optimal subsidy payments, $s^*\|\mathbf{e}\|_1$, in the homogeneous case over time. (Top right panel) The percentage increase in welfare due to the homogeneous subsidy, s^*, over time. (Bottom left panel) The total subsidy payments, $\mathbf{e}^\top \mathbf{s}^*$, when the subsidies are targeted towards specific firms. (Bottom right panel) The percentage increase in welfare due to the targeted subsidies, $\mathbf{s}^*$, over time.

ancy is not surprising, as current public funding instruments such as EUREKA do not take into account network effects stemming from R&D collaborations that determine our optimal subsidy policy.

We proceed by providing a similar ranking as for the key player policy by ranking firms in terms of targeted subsidies. In other words, if the planner wants to maximize total welfare, which firms should receive the highest subsidies and how much should it be. The ranking of the first 25 firms by their optimal subsidy levels in 1990 can be found in Table 8 while the one for 2005 is shown in Table 9. As for the key player policy, we see that the ranking of firms in terms of subsidies does not correspond to other rankings in terms of network centrality, patent stocks or market share. However, the ranking is similar to the one for the key firms. The correlation between the subsides and the key player ranking is 0.91 in the year 1990 and 0.88 in the year 2005. There is also volatility in the ranking since many firms that are ranked in the top 25 in 1990 are no longer there in 2005 (for example *McDonnell Douglas Corp.*, *Texaco Inc.*, *Honeywell Inc.*, etc.). Figure 11 shows the change in the ranking of the 25 highest subsidized firms (Table 8) from 1990 to 2005. The right-hand panel of Figure 6 shows the transition probability T_{ij} from a rank i in year t to a rank j in year $t+1$ for the first 100 ranks. As in the case of the key player rankings, the subsidy rankings are quite stable over time, where most transitions occur along the diagonal of T_{ij}. There is a larger variation at the bottom right corner of T_{ij} and less variation at the top left corner, showing

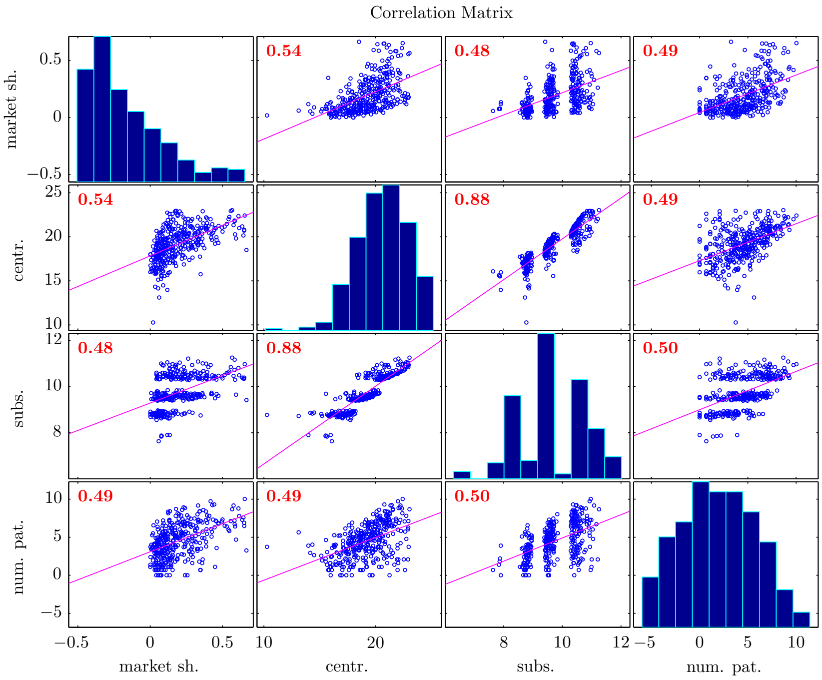


Figure 10: Pair correlation plot of market shares, the number of patents, the key player centrality, $(W(G) - W(G^{-i}))/W(G)$, used for the computation of the key player ranking (cf. Table 6), and the targeted subsidies, s_i^* (cf. Table 8), in the year 1990. The Spearman correlation coefficients are shown for each scatter plot. The data have been log and square root transformed to account for the heterogeneity in the data.

that the upper ranks are more stable than the lower ranks.

A comparison of market shares, the number of patents, the centrality used for the computation of the key player ranking and the targeted subsidies yields a high correlation between the key player centrality and the targeted subsidies. A slightly weaker correlation can also be found with the firms' market shares and patent stocks. The corresponding pair correlation plots for the year 1990 can be seen in Figure 10. Highly ranked firms tend to have a larger market share and also a larger patent stock. However, these measures can only partially explain the ranking of the firms, as the market share is more related to the product market rivalry effect, while the patent stock is more related to the technology spillover effect, and both enter into the computation of the key player ranking and the optimal subsidy program.

Observe that our subsidy rankings typically favor larger firms as they tend to be better connected in the R&D network than small firms.[48] This adds to the discussion of whether large or small firms are contributing more to the innovativeness of an economy [cf. Mandel, 2011],[49] by adding another dimension along which larger firms can have an advantage over small ones. Namely by creating R&D spillover effects that contribute to the overall productivity of the economy.[50] While studies such as Spencer and Brander [1983] and Acemoglu

[48] We further find a significant correlation between market share and the optimal subsidy levels of 0.48 in the year 1990 and 0.58 in the year 2005. See also Figure 10.

[49] See also "Big and clever. Why large firms are often more inventive than small ones." The Economist (2011, Dec. 17th). Retrieved from http://www.economist.com.

[50] Our findings regarding the pro-welfare effect of R&D conducted by large firms is in line with the results obtained by Bloom et al. [2013], where it is noted that "...smaller firms generate lower social returns to R&D because they operate more in technological niches."

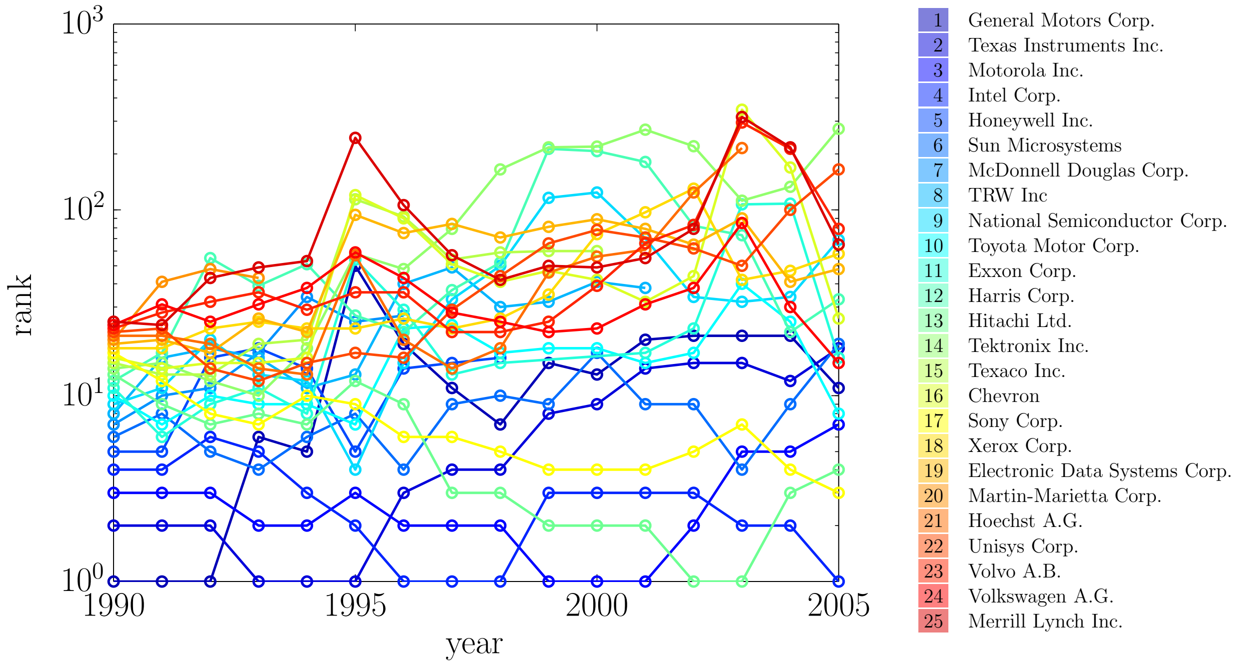


Figure 11: Change in the ranking of the 25 highest subsidized firms (Table 8) from 1990 to 2005.

et al. [2012] find that R&D should often be taxed rather than subsidized, we find in line with e.g. Hinloopen [2001] that R&D subsidies can have a significantly positive effect on welfare. As argued by Hinloopen [2001], the reason why our results differ from those of Spencer and Brander [1983] is that we take into account the consumer surplus when deriving the optimal R&D subsidy. Moreover, in contrast to Acemoglu et al. [2012], we do not focus on entry and exit but incorporate the network of R&D collaborating firms. This allows us to take into account the R&D spillover effects of incumbent firms, which are typically ignored in studies of the innovative activity of incumbent firms versus entrants. Therefore, we see our analysis as complementary to that of Acemoglu et al. [2012], and we show that R&D subsidies can trigger considerable welfare gains when technology spillovers through R&D alliances are incorporated.

Finally, if we compare the key player ranking and the subsidy ranking, we see that many firms appear in both rankings (such as *General Motors*, *Exxon Corp.*, *Toyota Motor Corp.*, etc.) but that there are also many firms that do not (such as *Altria Group*, *Honeywell Inc.*, etc.). In general, we believe that the key player policy is more relevant than the subsidy policy. First, it captures the fragility of the system. Second, it allows the planner to help or bail out the key firms whose removal or disappearance would be extremely costly in terms of total welfare and total activity for the economy. Third, by taking the network as given in this short run policy analysis, we do not incorporate the strategic formation of collaborations when firms expect that these collaboration might have an impact on the R&D subsidies they might receive. This is less of an issue for the key player analysis, where we study the response of the economy to exogenous, unanticipated shocks.

Table 8: Subsidies ranking for the year 1990 for the first 25 firms.

Firm	Share [%][a]	num pat.	**d**	$\mathbf{v}_{\mathrm{PF}}(G)$	Betweenness[b]	Closeness[c]	$q_i/\|\mathbf{q}\|_1$ [%][d]	$\frac{\|\mathbf{q}(G)\|_1-\|\mathbf{q}(G^{-i})\|_1}{\|\mathbf{q}(G)\|_1}$ [%][e]	$\mathbf{s}^*$ $[10^6]$	Rank
General Motors Corp.	12.1445	50185	12	0.1185	0.3381	0.0025	1.5986	2.1365	0.1368	1
Texas Instruments Inc.	20.5932	14822	22	0.3450	0.2083	0.0028	0.5353	0.8382	0.1201	2
Motorola Inc.	18.5193	7903	16	0.2996	0.1027	0.0025	0.6362	0.9320	0.1137	3
Intel Corp.	12.2966	1397	21	0.3029	0.2529	0.0028	0.4336	0.6611	0.1103	4
Honeywell Inc.	63.9769	17194	11	0.2464	0.1256	0.0024	0.4080	0.5552	0.1025	5
Sun Microsystems	11.0880	413	21	0.2791	0.1483	0.0026	0.2899	0.4483	0.1010	6
McDonnell Douglas Corp.	21.8941	899	14	0.1851	0.1452	0.0024	0.6207	0.8790	0.1005	7
TRW Inc	7.0559	1659	11	0.2016	0.0648	0.0023	0.5006	0.6758	0.0965	8
National Semiconductor Corp.	5.3366	1266	13	0.2527	0.0405	0.0023	0.3270	0.4391	0.0928	9
Toyota Motor Corp.	6.2806	58	9	0.1578	0.1420	0.0023	0.6760	0.8311	0.0920	10
Exxon Corp.	10.1151	6927	3	0.0187	0.0000	0.0015	1.3379	1.4618	0.0909	11
Harris Corp.	5.1937	2606	11	0.2257	0.0540	0.0023	0.3318	0.4460	0.0867	12
Hitachi Ltd.	37.6873	40838	7	0.1309	0.0141	0.0020	0.6854	0.8572	0.0825	13
Tektronix Inc.	17.5728	4486	14	0.2055	0.0787	0.0024	0.2902	0.4033	0.0823	14
Texaco Inc.	3.9206	8939	6	0.0427	0.1857	0.0019	0.7462	0.8684	0.0790	15
Chevron	3.7009	4410	5	0.0191	0.0430	0.0015	0.8636	0.9793	0.0779	16
Sony Corp.	32.0711	5840	8	0.1021	0.0382	0.0018	0.7258	0.8940	0.0770	17
Xerox Corp.	84.2264	24341	8	0.1378	0.0246	0.0020	0.6363	0.7836	0.0761	18
Electronic Data Systems Corp.	6.8935	8	7	0.1300	0.0616	0.0023	0.3882	0.4721	0.0738	19
Martin-Marietta Corp.	34.0667	818	8	0.1360	0.0423	0.0020	0.3406	0.4287	0.0704	20
Hoechst A.G.	13.8715	9634	2	0.0000	0.0006	0.0002	1.1624	1.2500	0.0675	21
Unisys Corp.	10.9318	9622	7	0.0836	0.0045	0.0017	0.4719	0.5777	0.0651	22
Volvo A.B.	1.3887	119	4	0.0147	0.0395	0.0014	0.6218	0.6822	0.0628	23
Volkswagen A.G.	4.1641	414	4	0.0184	0.0344	0.0014	0.6869	0.7570	0.0621	24
Merrill Lynch Inc.	13.1555	8	4	0.0206	0.0257	0.0014	0.7047	0.7705	0.0600	25

[a] Market share in the primary 4-digit sector in which the firm is operating.

[b] The normalized betweenness centrality is the fraction of all shortest paths in the network that contain a given node, divided by $(n-1)(n-2)$, the maximum number of such paths.

[c] The closeness centrality of node i is computed as $\sum_{j=1}^{n} 2^{-\ell_{ij}(G)}$, where $\ell_{ij}(G)$ is the length of the shortest path between i and j in the network G [Dangalchev, 2006].

[d] The relative output of a firm i is computed as $q_i/\|\mathbf{q}\|_1 = b_{\boldsymbol{\mu},i}/\|\mathbf{b}_{\boldsymbol{\mu}}\|_1$ (see Proposition 1).

[e] The decrease in output due to the removal of firm i is computed as $\frac{\|\mathbf{q}(G)\|_1-\|\mathbf{q}(G^{-i})\|_1}{\|\mathbf{q}(G)\|_1} = \frac{b_{\mathbf{u},i}(G)b_{\boldsymbol{\mu},i}(G)}{m_{ii}(G)}/\|\mathbf{b}_{\boldsymbol{\mu}}(G)\|_1$.

Table 9: Subsidies ranking for the year 2005 for the first 25 firms.

Firm	Share [%][a]	num pat.	$\mathbf{d}$	$\mathbf{v}_{\mathrm{PF}}(G)$	Betweenness[b]	Closeness[c]	$q_i/\lVert\mathbf{q}\rVert_1$ [%][d]	$\frac{\lVert\mathbf{q}(G)\rVert_1-\lVert\mathbf{q}(G^{-i})\rVert_1}{\lVert\mathbf{q}(G)\rVert_1}$ [%][e]	$\mathbf{s}^*$ [10^6]	Rank
Intel Corp.	9.8341	31709	15	0.3040	0.1593	0.0016	0.5607	0.7691	0.0470	1
Fujitsu Ltd.	17.3622	26830	14	0.3722	0.1840	0.0016	0.3941	0.5520	0.0468	2
Sony Corp.	32.1340	32362	12	0.3462	0.1095	0.0015	0.5436	0.7413	0.0468	3
Hitachi Ltd.	27.8692	106477	12	0.3199	0.1439	0.0015	0.5309	0.7532	0.0467	4
NTT DoCoMo	4.3962	1505	7	0.2752	0.1323	0.0016	0.6197	0.7763	0.0465	5
Daimler Corp.	7.5743	9952	9	0.0093	0.1239	0.0009	0.8474	1.0518	0.0464	6
Motorola Inc.	12.4529	26313	16	0.1841	0.5527	0.0019	0.3129	0.4458	0.0460	7
Toyota Motor Corp.	7.7760	567	4	0.0036	0.0019	0.0006	0.8056	0.8829	0.0458	8
Infineon Technologies AG	2.1293	13376	13	0.2028	0.1826	0.0016	0.2971	0.3855	0.0458	9
Cisco Systems Inc	63.1857	4432	10	0.1648	0.2269	0.0017	0.3313	0.4350	0.0457	10
General Motors Corp.	7.7341	66784	6	0.0101	0.0480	0.0009	0.5757	0.6823	0.0456	11
Mitsubishi Electric Corp	5.6782	30529	6	0.2016	0.0742	0.0014	0.3309	0.4075	0.0455	12
Sharp Corp.	8.5948	14454	6	0.2413	0.0066	0.0012	0.2951	0.3532	0.0455	13
Lockheed Martin Co	95.5769	5482	11	0.0856	0.0553	0.0011	0.2899	0.3975	0.0454	14
Volkswagen A.G.	4.8178	3931	3	0.0023	0.0000	0.0006	0.6201	0.6750	0.0453	15
Exxon Corp.	7.8647	11457	0	0.0000	0.0000	0.0000	0.8478	0.8744	0.0453	16
Boeing Company	31.5363	14222	9	0.0798	0.1254	0.0013	0.3572	0.4511	0.0452	17
Texas Instruments Inc.	3.3920	40456	6	0.1820	0.0278	0.0012	0.2632	0.2955	0.0452	18
Sun Microsystems	7.3032	8227	7	0.1645	0.0533	0.0011	0.2264	0.2835	0.0451	19
Mitsubishi Corp	87.2569	154	4	0.1914	0.0000	0.0011	0.3396	0.3777	0.0451	20
Oracle Corp.	7.8059	1586	5	0.1310	0.0321	0.0011	0.2784	0.3130	0.0450	21
Northrop Grumman	37.8576	11573	7	0.0705	0.0240	0.0011	0.2766	0.3426	0.0449	22
STMicroelectronics NV	2.2482	4942	4	0.1306	0.0206	0.0014	0.2984	0.3211	0.0449	23
Total SA	3.6544	1205	0	0.0000	0.0000	0.0000	0.6107	0.6415	0.0448	24
Chevron	4.4312	4987	0	0.0000	0.0000	0.0000	0.6201	0.6345	0.0447	25

[a] Market share in the primary 4-digit sector in which the firm is operating.

[b] The normalized betweenness centrality is the fraction of all shortest paths in the network that contain a given node, divided by $(n-1)(n-2)$, the maximum number of such paths.

[c] The closeness centrality of node i is computed as $\sum_{j=1}^{n} 2^{-\ell_{ij}(G)}$, where $\ell_{ij}(G)$ is the length of the shortest path between i and j in the network G [Dangalchev, 2006].

[d] The relative output of a firm i is computed as $q_i/\lVert\mathbf{q}\rVert_1 = b_{\boldsymbol{\mu},i}/\lVert\mathbf{b}_{\boldsymbol{\mu}}\rVert_1$ (see Proposition 1).

[e] The decrease in output due to the removal of firm i is computed as $\frac{\lVert\mathbf{q}(G)\rVert_1-\lVert\mathbf{q}(G^{-i})\rVert_1}{\lVert\mathbf{q}(G)\rVert_1} = \frac{b_{\mathbf{u},i}(G)b_{\boldsymbol{\mu},i}(G)}{m_{ii}(G)}/\lVert\mathbf{b}_{\boldsymbol{\mu}}(G)\rVert_1$.

Table 10: Optimal subsidies ranking for the year 2005 including the first 25 firms which also received funding trough EUREKA.

Firm	Share [%][a]	num pat.	$\mathbf{d}$	$\mathbf{v}_{\mathrm{PF}}(G)$	Betweenness[b]	Closeness[c]	$q_i/\lVert\mathbf{q}\rVert_1$ [%][d]	$\frac{\lVert\mathbf{q}(G)\rVert_1-\lVert\mathbf{q}(G^{-i})\rVert_1}{\lVert\mathbf{q}(G)\rVert_1}$ [%][e]	$\mathbf{s}^*$ $[10^6]$	EUREKA $[10^6]$[f]	Rank[g]
Fujitsu Ltd.	17.3622	26830	14	0.3722	0.1840	0.0016	0.3941	0.5520	0.0468	0.9600	2
Sony Corp.	32.1340	32362	12	0.3462	0.1095	0.0015	0.5436	0.7413	0.0468	488.1916	3
Hitachi Ltd.	27.8692	106477	12	0.3199	0.1439	0.0015	0.5309	0.7532	0.0467	2.0880	4
Daimler Corp.	7.5743	9952	9	0.0093	0.1239	0.0009	0.8474	1.0518	0.0464	10951.9789	6
Motorola Inc.	12.4529	26313	16	0.1841	0.5527	0.0019	0.3129	0.4458	0.0460	596.3673	7
Toyota Motor Corp.	7.7760	567	4	0.0036	0.0019	0.0006	0.8056	0.8829	0.0458	7.9403	8
Infineon Technologies AG	2.1293	13376	13	0.2028	0.1826	0.0016	0.2971	0.3855	0.0458	64.3849	9
Volkswagen A.G.	4.8178	3931	3	0.0023	0.0000	0.0006	0.6201	0.6750	0.0453	2003.9790	15
Texas Instruments Inc.	3.3920	40456	6	0.1820	0.0278	0.0012	0.2632	0.2955	0.0452	173.3877	18
Metro AG	17.6754	1	2	0.0543	0.0000	0.0009	0.4574	0.4842	0.0447	0.44737	27
Johnson Controls Inc.	43.0902	1165	5	0.0080	0.0351	0.0008	0.2962	0.3462	0.0447	2.7370	28
Continental A.G.	4.3929	4152	3	0.0023	0.0000	0.0006	0.2994	0.3273	0.0447	6.3594	29
Novartis AG	4.5096	10352	4	0.0048	0.1026	0.0008	0.4362	0.4842	0.0447	23.5423	32
Endesa	1.5322	5	0	0.0000	0.0000	0.0000	0.4501	0.4775	0.0445	12.6000	39
Thales	15.7304	16086	4	0.0282	0.0029	0.0008	0.2194	0.2566	0.0445	1083.8765	40
Unilever N.V./Plc.	9.0941	6697	0	0.0000	0.0000	0.0000	0.4910	0.5071	0.0445	58.8410	41
Roche Holding AG	4.1559	6119	4	0.0117	0.0422	0.0007	0.3859	0.4114	0.0445	3.1983	42
Nortel Networks Corp.	12.3793	7274	3	0.0372	0.0261	0.0011	0.2310	0.2598	0.0445	2901.7316	45
Electronic Data Systems Corp.	6.9497	284	2	0.0282	0.0176	0.0010	0.2598	0.2851	0.0444	0.30975	48
Sanofi	4.7555	7655	0	0.0000	0.0000	0.0000	0.4460	0.4340	0.0444	18.6991	52
Johnson & Johnson Inc.	7.0606	11940	0	0.0000	0.0000	0.0000	0.3706	0.3744	0.0443	0.6662	55
Xerox Corp.	23.0754	50254	1	0.0455	0.0000	0.0009	0.2540	0.2700	0.0443	0.0490	58
Agere Systems	0.4245	2582	3	0.0592	0.0349	0.0012	0.1193	0.1263	0.0442	3.8720	63
Renault	2.0905	2355	0	0.0000	0.0000	0.0000	0.4167	0.4075	0.0442	9668.1773	68
Omron Corp.	0.9875	1	2	0.0490	0.0176	0.0008	0.2004	0.2155	0.0442	0.2750	70

[a] Market share in the primary 4-digit sector in which the firm is operating.

[b] The normalized betweenness centrality is the fraction of all shortest paths in the network that contain a given node, divided by $(n-1)(n-2)$, the maximum number of such paths.

[c] The closeness centrality of node i is computed as $\sum_{j=1}^{n} 2^{-\ell_{ij}(G)}$, where $\ell_{ij}(G)$ is the length of the shortest path between i and j in the network G [Dangalchev, 2006].

[d] The relative output of a firm i is computed as $q_i/\lVert\mathbf{q}\rVert_1 = b_{\boldsymbol{\mu},i}/\lVert\mathbf{b}_{\boldsymbol{\mu}}\rVert_1$ (see Proposition 1).

[e] The decrease in output due to the removal of firm i is computed as $\frac{\lVert\mathbf{q}(G)\rVert_1-\lVert\mathbf{q}(G^{-i})\rVert_1}{\lVert\mathbf{q}(G)\rVert_1} = \frac{b_{\mathbf{u},i}(G)b_{\boldsymbol{\mu},i}(G)}{m_{ii}(G)}/\lVert\mathbf{b}_{\boldsymbol{\mu}}(G)\rVert_1$.

[f] The EUREKA subsidies comprise the total contribution to project costs (in Mio. USD), where all project costs involving a particular firm are accumulated. For more detailed information see `http://www.eurekanetwork.org/`.

[g] The rank corresponds to the ranking of Table 9.

11. Conclusion

In this paper, we have developed a model where firms jointly form R&D collaborations (networks) to lower their production costs while at the same time competing on the product market. We have highlighted the positive role of the network in terms of technology spillovers and the negative role of product rivalry in terms of market competition. We have also determined the importance of the key firms and targeted subsidies on the total welfare of the economy.

Using a panel of R&D alliance networks and annual reports, we have then tested our theoretical results and first showed that the magnitude of the technology spillover effect is much higher than that of the product rivalry effect, indicating that the latter dominates the former so that the net returns to R&D collaborations are strictly positive. We have also identified the key firms whose default would reduce social welfare and aggregate industry output the most. Finally, we have drawn some policy conclusions about optimal R&D subsidies from the results obtained over different sectors, as well as their temporal variation.

We believe that the methodology developed in this paper offers a fruitful way of analyzing the existence of R&D spillovers and their policy implications in terms of firms' subsidies. We also believe that putting forward the role of networks in terms of R&D collaborations is key to understanding the different aspects of these markets.

References

Acemoglu, D., Akcigit, U., Bloom, N., Kerr, W., 2012. Innovation, reallocation and growth. Stanford University Working Paper.

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

Anderson, G., Guionnet, A., Zeitouni, O., 2010. An Introduction to Random Matrices. Vol. 200. Cambridge University Press.

Anselin, L. 1988. Spatial Econometric: Methods and Model. Kluwer Academic Publishers, Dordrecht.

Arrow, K.J. 1962. Economic welfare and the allocation of resources for invention. In: R. Nelson (Ed.), The Rate and Direction of Inventive Activity: Economic and Social Factors, Princeton: Princeton University Press.

Bagwell, K., Staiger, R., 1994. The sensitivity of strategic and corrective R&D policy in oligopolistic industries. Journal of International Economics 36 (1-2), 133–150.

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

Ballester, C., Calvó-Armengol, A., Zenou, Y., 2010. Delinquent Networks. Journal of the European Economic Association, 8, 34–61.

Ballester, C., Vorsatz, M., 2014. Random walk based segregation measures. Review of Economics and Statistics, forthcoming.

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

Belhaj, M., Bervoets, S. and Deroïan, F., 2013. Using network design as a public policy tool. Unpublished manuscript, Aix-Marseille School of Economics.

Blondel, V., Gajardo, A., Heymans, M., Senellart, P., Van Dooren, P., 2004. A measure of similarity between graph vertices: Applications to synonym extraction and web searching. Siam Review 46 (4), 647–666.

Bloom, N., Griffith, R., Van Reenen, J., 2002. Do R&D tax credits work? evidence from a panel of countries 1979–1997. Journal of Public Economics 85 (1), 1–31.

Bloom, N., Schankerman, M., Van Reenen, J., 2013. Identifying technology spillovers and product market rivalry. Econometrica 81 (4), 1347–1393.

Blume, L.E., Brock, W.A., Durlauf, S.N., Ioannides, Y.M., 2011. Identification of social interactions. In: J. Benhabib, A. Bisin, and M.O. Jackson (Eds.), Handbook of Social Economics, Amsterdam: Elsevier Science.

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

Borgatti, S.P., 2003. The key player problem. In: R. Breiger, K. Carley and P. Pattison (Eds.), Dynamic Social Network Modeling and Analysis: Workshop Summary and Papers, New York: National Academy of Sciences Press, pp. 241-252.

Borgatti, S.P., 2006. Identifying sets of key players in a network. Computational, Mathematical and Organizational Theory 12, 21–34.
Bramoullé, Y., Djebbari, H., Fortin, B., 2009. Identification of peer effects through social networks. Journal of Econometrics 150, 41–55.
Bramoullé, Y., Kranton, R., D'Amours, M., 2014. Strategic interaction and networks. American Economic Review 104, 898–930.
Brualdi, Richard A. and Hoffman, A.J., 1985. On the spectral radius of (0,1)-matrices. Linear Algebra and its Applications, 65, 133–146.
Boyd, S., Vandenberghe, L., 2004. Convex Optimization. Cambridge University Press.
Calvó-Armengol, A., Patacchini, E., Zenou, Y., 2009. Peer effects and social networks in education. Review of Economic Studies 76, 1239–1267.
Cvetković, Dragoš and Rowlinson, P., 1990. The largest eigenvalue of a graph: A survey. Linear and Multilinear Algebra, 28, 3–33.
Chandrasekhar, A., Lewis, R., 2011. Econometrics of sampled networks. MIT Working Paper.
Cohen, W. M. & Levinthal, D. A., 1990. Absorptive Capacity: A New Perspective on Learning and Innovation. Administrative Science Quarterly 35, 128–152.
Cohen, W., Klepper, S., 1996a. A reprise of size and R & D. The Economic Journal 106 (437), 925–951.
Cohen, W., 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-Cole, E., 2006. Multiple groups identification in the linear-in-means model. Economics Letters. 92, 157-162.
Corbo, J., Calvó-Armengol, A., Parkes, D., 2006. A study of nash equilibrium in contribution games for peer-to-peer networks. SIGOPS Operation Systems Review 40 (3), 61–66.
Cvetkovic, D., Doob, M., Sachs, H., 1995. Spectra of graphs: theory and applications. Johann Ambrosius Barth.
Cvetkovic, D., Rowlinson, P., Simic, S., 1997. Eigenspaces of graphs. Vol. 66. Cambridge University Press, Encyclopedia of Mathematics.
Dangalchev, C., 2006. Residual closeness in networks. Physica A 365 (2), 556–564.
Dawid, H. and Hellmann, T., 2012. The Evolution of R&D Networks, Mimeo, University of Bielefeld.
D'Aspremont, C., Jacquemin, A., 1988. Cooperative and noncooperative R&D in duopoly with spillovers. The American Economic Review 78 (5), 1133–1137.
Debreu, G., Herstein, I. N., 1953. Nonnegative square matrices. Econometrica 21 (4), 597–607.
De Martí, J., Zenou, Y. 2011, Social networks. In: I. Jarvie et J. Zamora-Bonilla (Eds.), Handbook of Philosophy of Social Science. Londres: SAGE Publications, Chap. 16, pp. 339-361.
Dorogovtsev, S., Goltsev, A., Mendes, J., Samukhin, A., 2003. Spectra of complex networks. Physical Review E 68 (4), 46109.
Durlauf, S.N., Ioannides, Y.M., 2010. Social interactions. Annual Review of Economics, 2, 451-478.
Duso, T., Röller, L.-H., Seldeslachts, J., 2012. Collusion through joint R&D: An empirical assessment. KU Leuven-Faculty of Business and Economics Research Paper No. OR1229.
Echenique, F., Fryer Jr., R., 2007. A Measure of Segregation Based on Social Interactions. The Quarterly Journal of Economics 122 (2), 441–485.
Echenique, F., Fryer Jr., R., Kaufman, A., 2006. Is school segregation good or bad? The American Economic Review 96 (2), 265–269.
Einiö, E., 2013. R&D subsidies and company performance: Evidence from geographic variation in government funding based on the ERDF population-density rule. Review of Economics and Statistics, forthcoming.
Fleming, L., King, C., and A.I. Juda, 2007. Small worlds and regional innovation. Organization Science 18(6), 938-?954.
Galí, J., 1999. Technology, employment, and the business cycle: Do technology shocks explain aggregate fluctuations? American Economic Review 89 (1), 249–271.
Goeree, M. S., Helland, E., 2012. Do research joint ventures serve a collusive function? Mimeo, Institute for Empirical Research in Economics-University of Zurich.
Goyal, S., 2007. Connections: An Introduction to the Economics of Networks. Princeton: Princeton University Press.
Goyal, S., Joshi, S., 2003. Networks of collaboration in oligopoly. Games and Economic Behavior 43 (1), 57–85.
Goyal, S., Moraga-Gonzalez, J. L., 2001. R&D networks. RAND Journal of Economics 32 (4), 686–707.
Griffith, R., Harrison, R. and J. Van Reenen, 2006. How special is the special relationship? Using the impact of US R&D spillovers on UK firms as a test of technology sourcing. American Economic Review 96(5), 1859–1875.
Gulati, R., Gargiulo, M., 1999. Where do interorganizational networks come from? The American Journal of Sociology 104 (5), 1439–1493.
Hagberg, A., Swart, P. and Schult, D., 2006. Designing threshold networks with given structural and dynamical properties. Physical Review E, 74, 56116.
Hagedoorn, J., May 2002. Inter-firm R&D partnerships: an overview of major trends and patterns since 1960. Research Policy 31 (4), 477–492.
Hall, B. H., Mairesse, J., Pierre, M., 2010. Measuring the returns to R&D. Handbook in Economics 2.
Hall, B.H. and Jaffe, A.B. and Trajtenberg, M., 2001. The NBER Patent Citation Data File: Lessons, Insights and Methodological Tools. NBER Working Paper No. 8498.
Hanaki, N., Nakajima, R. and Y. Ogura, 2010. The dynamics of R&D network in the IT industry. Research policy 39(3), 386-?399.
Harrigan, K., 1988. Strategic alliances and partner asymmetries. Management international review 28 (4), 53–72.
Hinloopen, J., 2001. Subsidizing R&D cooperatives. De Economist 149 (3), 313–345.
Horn, R. A., Johnson, C. R., 1990. Matrix Analysis. Cambridge: Cambridge University Press.

Impullitti, G., 2010. International competition and U.S. R&D subsidies: A quantitative welfare analysis. International Economic Review 51 (4), 1127–1158.
Jackson, M. O., 2008. Social and Economic Networks. Princeton: Princeton University Press.
Jackson, M. O., Wolinsky, A., 1996. A strategic model of social and economic networks. Journal of Economic Theory 71 (1), 44–74.
Jackson, M. O., Zenou, Y. 2013. Economic Analyses of Social Networks. The International Library of Critical Writings in Economics. London: Edward Elgar Publishing.
Jackson, M. O., Zenou, Y., 2015. Games on networks. In: P. Young and S. Zamir (Eds.), Handbook of Game Theory Vol. 4, Amsterdam: Elsevier Publisher, pp. 91-157.
Jaffe, Adam B. 1986. Technological Opportunity and Spillovers of R & D: Evidence from Firms' Patents, Profits, and Market Value. The American Economic Review 76(5), 984–1001.
Jones, C. and J. Williams, 1998. Measuring the social rate of return to R&D. Quarterly Journal of Economics 113(4), 119–135.
Kamien, M. I., Muller, E., Zang, I., 1992. Research joint ventures and R&D cartels. The American Economic Review 82 (5), 1293–1306.
Kelejian, H. H., Piras, G., 2012. Estimation of spatial models with endogenous weighting matrices, and an application to a demand model for cigarettes. Working paper, University of Maryland.
Kitsak, M., Riccaboni, M., Havlin, S., Pammolli, F., Stanley, H., 2010. Scale-free models for the structure of business firm networks. Physical Review E 81, 036117.
Klette, T.J. and Kortum, S., 2004. Innovating firms and aggregate innovation. Journal of Political Economy 112 (5), 986–1018.
Kogut, B., 1988. Joint ventures: Theoretical and empirical perspectives. Strategic Management Journal 9 (4), 319–332.
König, M. D., 2013. Dynamic R&D networks, University of Zurich, Department of Economics Working Paper No. 109.
König, M. D., Battiston, S., Napoletano, M., Schweitzer, F., 2011. The efficiency and stability of R&D networks. Games and Economic Behavior 75, 694–713.
König, M. D., Tessone, C., Zenou, Y., 2014. Nestedness in networks: A theoretical model and some applications, Theoretical Economics, forthcoming.
König, M. D., 2013. Technology cycles in dynamic R&D networks. *Mimeo, University of Zurich.*
Leahy, D. and P. Neary, 1997. Public policy towards R&D in oligopolistic industries. American Economic Review 87 (4), 642–662.
Lee, L., Liu, X., 2010. Efficient GMM estimation of high order spatial autoregressive models with autoregressive disturbances. Econometric Theory 26 (1), 187–230.
Leicht, E. A., Holme, P., Newman, M. E. J., Feb 2006. Vertex similarity in networks. Phys. Rev. E 73 (2), 026120.
Leonard, W. N., 1971. Research and development in industrial growth. The Journal of Political Economy, 232–256.
Lindquist, M.J. and Y. Zenou, 2013. Key players in co-offending networks. Unpublished manuscript, Stockholm University.
Lin, X., 2010. Identifying peer effects in student academic achievement by a spatial autoregressive model with group unobservables. Journal of Labor Economics, 28, 825-860.
Liu, X., 2013. Estimation of a local-aggregate network model with sampled networks. Economics Letters 118 (1), 243–246.
Liu, X., Patacchini, E., Zenou, Y., Lee, L., 2011. Criminal networks: Who is the key player? Unpublish manuscript, Stockholm University.
De Loecker, J., 2011. Product differentiation, multiproduct firms, and estimating the impact of trade liberalization on productivity. Econometrica 79, 1407–1451.
Mahadev, N.V.R. and Peled, Uri N., 1995. Threshold Graphs and Related Topics. North Holland, Amsterdam.
Mandel, M., 2011. Scale and Innovation in Today's Economy, Progressive Policy Institute Policy Memo.
Mansfield, E., Rapoport, J., Schnee, J., Wagner, S., Hamburger, M., 1971. Research and development in the modern corporation. New York: WW Norton.
Manshadi, V., Johari, R., 2010. Supermodular network games. In: 47th Annual Allerton Conference on Communication, Control, and Computing, 2009. IEEE, pp. 1369–1376.
Manski, C.F., 1993. Identification of endogenous effects: The reflection problem. Review of Economic Studies, 60, 531-542
Maroulas, J., Psarrakos, P. J., Tsatsomeros, M. J., 2000. Separable characteristic polynomials of pencils and property. Electronic Journal of Linear Algebra 7, 182–190.
Mihail, M., Papadimitriou, C., 2002. On the Eigenvalue Power Law. Randomization and Approximation Techniques in Computer Science: 6th International Workshop, RANDOM 2002, Cambridge, MA, USA, September 13-15, 2002: Proceedings.
Mises, R., Pollaczek-Geiringer, H., 1929. Praktische verfahren der gleichungsauflösung. ZAMM-Journal of Applied Mathematics and Mechanics/Zeitschrift für Angewandte Mathematik und Mechanik 9 (1), 58–77.
Newman, M., 2010. Networks: An Introduction. Oxford University Press.
Newman, M. E. J., Oct 2002. Assortative mixing in networks. Physical Review Letters 89 (20), 208701.
Newman, M. E. J., Feb 2003. Mixing patterns in networks. Physical Review E 67 (2), 026126.
Nooteboom, B., Vanhaverbeke, W., Duysters, G., Gilsing, V., Oord, A. v. d., 2006. Optimal cognitive distance and absorptive capacity, Discussion Paper.
OECD, 2012. Gross domestic expenditure on R&D. Science and Technology: Key Tables from OECD No. 1.
Park, S., Russo, M., 1996. When competition eclipses cooperation: An event history analysis of joint venture failure. Management science, 875–890.
Pastor-Satorras, R., Vazquez, A., Vespig–i, A., 2001. Dynamical and correlation properties of the internet. Physical Review Letters 87.
Peled, U. N., Petreschi, R., Sterbini, A., 1999. (n, e)-graphs with maximum sum of squares of degrees. Journal of Graph Theory 31 (4), 283–295.

Podolny, J. M., 1993. A status-based model of market competition. American Journal of Sociology 98 (4), pp. 829–872.
Powell, W., Grodal, S., April 2006. Oxford Handbook of Innovation. Oxford University Press, USA, Ch. Networks of Innovators.
Powell, W., Koput, K. W., 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., 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.
Qiu, L. D., Tao, Z., 1998. Policy on international R&D cooperation: Subsidy or tax? European Economic Review 42 (9), 1727 – 1750.
Ravenscraft, D., Scherer, F. M., 1982. The lag structure of returns to research and development. Applied Economics 14 (6), 603–620.
Rencher, A. C., Christensen, W. F., 2012. Methods of multivariate analysis. Vol. 709. John Wiley & Sons.
Riccaboni, M., Pammolli, F., 2002. On firm growth in networks. Research Policy 31 (8-9), 1405–1416.
Roijakkers, N., Hagedoorn, J., April 2006. Inter-firm R&D partnering in pharmaceutical biotechnology since 1975: Trends, patterns, and networks. Research Policy 35 (3), 431–446.
Rosenkopf, L., Padula, G., 2008. Investigating the Microstructure of Network Evolution: Alliance Formation in the Mobile Communications Industry. Organization Science 19 (5), 669.
Rosenkopf, L., Schilling, M., 2007. Comparing alliance network structure across industries: Observations and explanations. Strategic Entrepreneurship Journal 1, 191–209.
Rucker, G., Rucker, C., Gutman, I., 2002. On kites, comets, and stars. Sums of eigenvector coefficients in (molecular) graphs. Zeitschrift für Naturforschung 57(3/4), 143–153.
Samuelson, P., 1942. A method of determining explicitly the coefficients of the characteristic equation. The Annals of Mathematical Statistics, 424–429.
Schilling, M., 2009. Understanding the alliance data. Strategic Management Journal 30 (3), 233–260.
Singh, N., Vives, X., 1984. Price and quantity competition in a differentiated duopoly. RAND Journal of Economics 15 (4), 546–554.
Song, H., Vannetelbosch, V., 2007. International R&D collaboration networks. The Manchester School 75 (6), 742–766.
Spence, M. 1984. Cost reduction, competition, and industry performance. Econometrica 52, 101-121.
Spencer, B. J., Brander, J. A., 1983. International R & D rivalry and industrial strategy. The Review of Economic Studies 50 (4), 707–722.
Steinmueller, W. E., 2010. Economics of technology policy. Handbook of the Economics of Innovation 2, 1181–1218.
Suzumura, K. 1992. Cooperative and noncooperative R&D in an oligopoly with spillovers. American Economic Review 82, 1307-1320.
Stuart, T. E., 2000. Interorganizational alliances and the performance of firms: A study of growth and innovation rates in a high-technology industry. Strategic management journal 21 (8), 791–811.
Takalo, T., Tanayama, T. and Toivanen, O., 2013. Estimating the benefits of targeted R&D subsidies. Review of Economics and Statistics, 95, 255–272.
Vega-Redondo, F., 2007. Complex Social Networks. Cambridge University Press.
Van Mieghem, P., 2011. Graph spectra for complex networks. Cambridge University Press.
Wang, W., Lee, L.-f., 2013. Estimation of spatial panel data models with randomly missing data in the dependent variable. Regional Science and Urban Economics.
Wansbeek, T., Kapteyn, A., 1989. Estimation of the error-components model with incomplete panels. Journal of Econometrics 41 (3), 341–361.
Wasserman, S. and K. Faust, 1994. Social Network Analysis: Methods and Applications. Cambridge University Press.
Weitzman, M. L., 1998. Recombinant growth. The Quarterly Journal of Economics 113 (2), 331–360.
Westbrock, B., 2010. Natural concentration in industrial research collaboration. The RAND Journal of Economics 41 (2), 351–371.
Zenou, Y. 2015a. Networks in Economics. In: J.D. Wright (Ed.), International Encyclopedia of Social and Behavioral Sciences, 2nd Edition. Amsterdam: Elsevier Publisher, forthcoming.
Zenou, Y. 2015b. Key players. In: Y. Bramoullé, Y., Galeotti, A. and B. Rogers (Eds.), International Encyclopedia of Social and Behavioral Sciences, Oxford: Oxford University Press, forthcoming.

Appendix

A. Definitions and Characterizations

A.1. Network Definitions

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

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

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

Let $G=(\mathcal{N},\mathcal{E})$ be a graph whose distinct positive degrees are $d_{(1)}<d_{(2)}<\ldots<d_{(k)}$, and let $d_0=0$ (even if no agent with degree 0 exists in G). Further, define $\mathcal{D}_i=\{v\in\mathcal{N}:d_v=d_{(i)}\}$ for $i=0,\ldots,k$. Then the set-valued vector $\mathcal{D}=(\mathcal{D}_0,\mathcal{D}_1,\ldots,\mathcal{D}_k)$ is called the *degree partition* of G. Consider a *nested split graph* $G=(\mathcal{N},\mathcal{E})$ and let $\mathcal{D}=(\mathcal{D}_0,\mathcal{D}_1,\ldots,\mathcal{D}_k)$ be its degree partition. Then the nodes $\mathcal{N}$ can be partitioned in independent sets $\mathcal{D}_i$, $i=1,\ldots,\lfloor\frac{k}{2}\rfloor$ and a dominating set $\bigcup_{i=\lfloor\frac{k}{2}\rfloor+1}^{k}\mathcal{D}_i$ in the graph $G'=(\mathcal{N}\backslash\mathcal{D}_0,\mathcal{E})$. Moreover, the neighborhoods of the nodes are nested. In particular, for each node $v\in\mathcal{D}_i$, $\mathcal{N}_v=\bigcup_{j=1}^{i}\mathcal{D}_{k+1-j}$ if $i=1,\ldots,\lfloor\frac{k}{2}\rfloor$ if $i=1,\ldots,k$, while $\mathcal{N}_v=\bigcup_{j=1}^{i}\mathcal{D}_{k+1-j}\setminus\{v\}$ if $i=\lfloor\frac{k}{2}\rfloor+1,\ldots,k$.

In a *complete* graph K_n, every node is adjacent to every other node. The graph in which no pair of nodes is adjacent is the empty graph $\bar{K}_n$. A *clique* $K_{n'}$, $n'\leq n$, is a complete subgraph of the network G. A graph is *k-regular* if every node i has the same number of links $d_i=k$ for all $i\in\mathcal{N}$. The complete graph K_n is $(n-1)$-regular. The cycle C_n is 2-regular. In a *bipartite graph* there exists a partition of the nodes in two disjoint sets V_1 and V_2 such that each link connects a node in V_1 to a node in V_2. V_1 and V_2 are independent sets with cardinalities n_1 and n_2, respectively. In a complete bipartite graph K_{n_1,n_2} each node in V_1 is connected to each other node in V_2. The *star* $K_{1,n-1}$ is a complete bipartite graph in which $n_1=1$ and $n_2=n-1$.

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

Let $\mathbf{A}$ be the symmetric $n\times n$ *adjacency matrix* of the network G. The element $a_{ij}\in\{0,1\}$ indicates if there exists a link between nodes i and j such that $a_{ij}=1$ if $(i,j)\in\mathcal{E}$ and $a_{ij}=0$

if $(i,j) \notin \mathcal{E}$. The k-th power of the adjacency matrix is related to walks of length k in the graph. In particular, $\left(\mathbf{A}^k\right)_{ij}$ gives the number of walks of length k from node i to node j. The *eigenvalues* of the adjacency matrix $\mathbf{A}$ are the numbers $\lambda_1, \lambda_2, \ldots, \lambda_n$ such that $\mathbf{A}\mathbf{v}_i = \lambda_i \mathbf{v}_i$ has a nonzero solution vector $\mathbf{v}_i$, which is an *eigenvector* associated with λ_i for $i = 1, \ldots, n$. Since the adjacency matrix $\mathbf{A}$ of an undirected graph G is real and symmetric, the eigenvalues of $\mathbf{A}$ are real, $\lambda_i \in \mathbb{R}$ for all $i = 1, \ldots, n$. Moreover, if $\mathbf{v}_i$ and $\mathbf{v}_j$ are eigenvectors for different eigenvalues, $\lambda_i \neq \lambda_j$, then $\mathbf{v_i}$ and $\mathbf{v}_j$ are orthogonal, i.e. $\mathbf{v}_i^\top \mathbf{v}_j = 0$ if $i \neq j$. In particular, $\mathbb{R}^n$ has an orthonormal basis consisting of eigenvectors of $\mathbf{A}$. Since $\mathbf{A}$ is a real symmetric matrix, there exists an orthogonal matrix $\mathbf{S}$ such that $\mathbf{S}^\top\mathbf{S} = \mathbf{S}\mathbf{S}^\top = \mathbf{I}$ (that is $\mathbf{S}^\top = \mathbf{S}^{-1}$) and $\mathbf{S}^\top\mathbf{A}\mathbf{S} = \mathbf{D}$, where $\mathbf{D}$ is the diagonal matrix of eigenvalues of $\mathbf{A}$ and the columns of $\mathbf{S}$ are the corresponding eigenvectors. The *Perron-Frobenius eigenvalue* $\lambda_{\mathrm{PF}}(G)$ is the *largest real eigenvalue* of $\mathbf{A}$ associated with G, i.e. all eigenvalues λ_i of $\mathbf{A}$ satisfy $|\lambda_i| \leq \lambda_{\mathrm{PF}}(G)$ for $i = 1, \ldots, n$ and there exists an associated nonnegative eigenvector $\mathbf{v}_{\mathrm{PF}} \geq 0$ such that $\mathbf{A}\mathbf{v}_{\mathrm{PF}} = \lambda_{\mathrm{PF}}(G)\mathbf{v}_{\mathrm{PF}}$. For a connected graph G the adjacency matrix $\mathbf{A}$ has a unique largest real eigenvalue $\lambda_{\mathrm{PF}}(G)$ and a positive associated eigenvector $\mathbf{v}_{\mathrm{PF}} > 0$. There exists a relation between the number of walks in a graph and its eigenvalues. The number of closed walks of length k from a node i in G to herself is given by $\left(\mathbf{A}^k\right)_{ii}$ and the total number of closed walks of length k in G is $\mathrm{tr}\left(\mathbf{A}^k\right) = \sum_{i=1}^n \left(\mathbf{A}^k\right)_{ii} = \sum_{i=1}^n \lambda_i^k$. We further have that $\mathrm{tr}\left(\mathbf{A}\right) = 0$, $\mathrm{tr}\left(\mathbf{A}^2\right)$ gives twice the number of links in G and $\mathrm{tr}\left(\mathbf{A}^3\right)$ gives six times the number of triangles in G.

A.2. Walk Generating Functions

Denote by $\mathbf{u} = (1, \ldots, 1)^\top$ the n-dimensional vector of ones and define $\mathbf{M}(G, \phi) = (\mathbf{I}_n - \phi\mathbf{A})^{-1}$. Then, the quantity $N_G(\phi) = \mathbf{u}^\top \mathbf{M}(G, \phi)\mathbf{u}$ is the *walk generating function* of the graph G [cf. Cvetkovic et al., 1995]. Let us show this result. Let N_k denote the number of walks of length k in G. Then we can write N_k as follows

$$N_k = \sum_{i=1}^n \sum_{j=1}^n a_{ij}^{[k]} = \mathbf{u}^\top \mathbf{A}^k \mathbf{u},$$

where $a_{ij}^{[k]}$ is the ij-th element of $\mathbf{A}^k$. The walk generating function is then defined as

$$N_G(\phi) \equiv \sum_{k=0}^\infty N_k \phi^k = \mathbf{u}^\top \left(\sum_{k=0}^\infty \phi^k \mathbf{A}^k\right) \mathbf{u} = \mathbf{u}^\top \left(\mathbf{I}_n - \phi\mathbf{A}\right)^{-1} \mathbf{u} = \mathbf{u}^\top \mathbf{M}(G, \phi)\mathbf{u}.$$

For a k-regular graph G_k, the walk generating function is equal to

$$N_{G_k}(\phi) = \frac{n}{1 - k\phi}.$$

It holds that $N_G(0) = n$, and one can show that $N_G(\phi) \geq 0$. We further have that

$$\mathbf{M}(G, \phi) = (\mathbf{I}_n - \phi\mathbf{A})^{-1} = \sum_{k=0}^\infty \phi^k \mathbf{A}^k = \sum_{k=0}^\infty \phi^k \mathbf{S}\mathbf{\Lambda}^k \mathbf{S}^\top,$$

where $\mathbf{\Lambda} \equiv \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ is the diagonal matrix containing the eigenvalues of the real, symmetric matrix $\mathbf{A}$, and $\mathbf{S}$ is an orthogonal matrix with columns given by the orthogonal eigenvectors of $\mathbf{A}$ (with $\mathbf{S}^\top = \mathbf{S}^{-1}$), and we have used the fact that $\mathbf{A} = \mathbf{S}\mathbf{\Lambda}\mathbf{S}^\top$ [Horn and Johnson, 1990]. The eigenvectors $\mathbf{v}_i$ have the property that $\mathbf{A}\mathbf{v}_i = \lambda_i \mathbf{v}_i$ and are normalized such that

$\mathbf{v}_i^\top \mathbf{v}_i = 1$. Note that $\mathbf{A} = \mathbf{S}\boldsymbol{\Lambda}\mathbf{S}^\top$ is equivalent to $\mathbf{A} = \sum_{i=1}^n \lambda_i \mathbf{v}_i \mathbf{v}_i^\top$. It then follows that

$$\mathbf{u}^\top \mathbf{M}(G,\phi)\mathbf{u} = \mathbf{u}^\top \mathbf{S} \sum_{k=0}^{\infty} \phi^k \boldsymbol{\Lambda}^k \mathbf{S}^\top \mathbf{u},$$

where

$$\mathbf{S}^\top \mathbf{u} = \left(\mathbf{u}^\top \mathbf{v}_1, \ldots, \mathbf{u}^\top \mathbf{v}_n\right)^\top,$$

and

$$\boldsymbol{\Lambda}^k = \begin{pmatrix} \lambda_1^k & 0 & \ldots & 0 \\ 0 & \lambda_2^k & \ldots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \ldots & & \lambda_n^k \end{pmatrix} = \lambda_1^k \begin{pmatrix} 1 & 0 & \ldots & 0 \\ 0 & \left(\frac{\lambda_2}{\lambda_1}\right)^k & \ldots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \ldots & & \left(\frac{\lambda_n}{\lambda_1}\right)^k \end{pmatrix}.$$

We then can write

$$\mathbf{u}^\top \mathbf{M}(G,\phi)\mathbf{u} = \sum_{k=0}^{\infty} \phi^k \lambda_1^k \left(\mathbf{u}^\top \mathbf{v}_1, \ldots, \mathbf{u}^\top \mathbf{v}_n\right) \begin{pmatrix} 1 & 0 & \ldots & 0 \\ 0 & \left(\frac{\lambda_2}{\lambda_1}\right)^k & \ldots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \ldots & & \left(\frac{\lambda_n}{\lambda_1}\right)^k \end{pmatrix} \left(\mathbf{u}^\top \mathbf{v}_1, \ldots, \mathbf{u}^\top \mathbf{v}_n\right)^\top,$$

which gives

$$\begin{aligned} \mathbf{u}^\top \mathbf{M}(G,\phi)\mathbf{u} &= \sum_{k=0}^{\infty} \phi^k \lambda_1^k \left((\mathbf{u}^\top \mathbf{v}_1)^2 + \left(\frac{\lambda_2}{\lambda_1}\right)^k (\mathbf{u}^\top \mathbf{v}_2)^2 + \ldots + \left(\frac{\lambda_n}{\lambda_1}\right)^k (\mathbf{u}^\top \mathbf{v}_n)^2 \right) \\ &= \sum_{i=1}^{n} (\mathbf{u}^\top \mathbf{v}_i)^2 \sum_{k=0}^{\infty} \phi^k \lambda_i^k \\ &= \sum_{i=1}^{n} \frac{(\mathbf{u}^\top \mathbf{v}_i)^2}{1 - \phi\lambda_i}. \end{aligned}$$

The above computation also shows that

$$N_k = \mathbf{u}^\top \mathbf{A}^k \mathbf{u} = \sum_{i=1}^{n} (\mathbf{u}^\top \mathbf{v}_i)^2 \lambda_i^k.$$

Hence, we can write the walk generating function as follows

$$N_G(\phi) = \mathbf{u}^\top \mathbf{M}(G,\phi)\mathbf{u} = \sum_{k=0}^{\infty} N_k \phi^k = \sum_{i=1}^{n} \frac{(\mathbf{v}_i^\top \mathbf{u})^2}{1 - \lambda_i \phi}.$$

If λ_1 is much larger than λ_j for all $j \geq 2$, then we can approximate

$$N_G(\phi) \approx (\mathbf{u}^\top \mathbf{v}_1)^2 \sum_{k=0}^{\infty} \phi^k \lambda_1^k = \frac{(\mathbf{u}^\top \mathbf{v}_1)^2}{1 - \phi\lambda_1}.$$

Cvetkovic et al. [1995, p. 45] has found an alternative expression for the walk generating function given by

$$N_G(\phi) = \frac{1}{\phi}\left((-1)^n \frac{c_{\mathbf{A}^c}\left(-\frac{1}{\phi} - 1\right)}{c_{\mathbf{A}}\left(\frac{1}{\phi}\right)} - 1 \right),$$

where $c_{\mathbf{A}}(\phi) \equiv \det(\mathbf{A} - \phi \mathbf{I}_n)$ is the characteristic polynomial of the matrix $\mathbf{A}$, whose roots are the eigenvalues of $\mathbf{A}$. It can be written as $c_{\mathbf{A}}(\phi) = \phi^n - a_1\phi^{n-1} + \ldots + (-1)^n a_n$, where $a_1 = \mathrm{tr}(\mathbf{A})$ and $a_n = \det(\mathbf{A})$. Further, $\mathbf{A}^c = \mathbf{u}\mathbf{u}^\top - \mathbf{I}_n - \mathbf{A}$ is the complement of $\mathbf{A}$, and $\mathbf{u}\mathbf{u}^\top$ is an $n \times n$ matrix of ones. This is a convenient expression for the walk generating function, as there exist fast algorithms to compute the characteristic polynomial [Samuelson, 1942].

A.3. Nested Split Graphs

Let us define *nested split graphs* [Cvetkovic and Rowlinson, 1990; Mahadev and Peled, 1995], which include many common networks such as the star network. Moreover, as their name already indicates, they have a *nested neighborhood structure*. This means that the set of neighbors of each agent is contained in the set of neighbors of each higher degree agent. Nested split graphs have particular topological properties and an associated adjacency matrix with a well defined structure.

In order to characterize nested split graphs, it will be necessary to consider the degree partition of a graph, which is defined as follows:

Definition 1 (Mahadev and Peled [1995]). *Let $G = (\mathcal{N}, \mathcal{E})$ be a graph whose distinct positive degrees are $d_{(1)} < d_{(2)} < \ldots < d_{(k)}$, and let $d_0 = 0$ (even if no agent with degree 0 exists in G). Further, define $\mathcal{D}_i = \{v \in \mathcal{N} : d_v = d_{(i)}\}$ for $i = 0, \ldots, k$. Then the set-valued vector $\mathcal{D} = (\mathcal{D}_0, \mathcal{D}_1, \ldots, \mathcal{D}_k)$ is called the degree partition of G.*

With the definition of a degree partition, we can now give a more formal definition of a nested split graph.[51]

Definition 2 (Mahadev and Peled [1995]). *Consider a nested split graph $G = (\mathcal{N}, \mathcal{E})$ and let $\mathcal{D} = (\mathcal{D}_0, \mathcal{D}_1, \ldots, \mathcal{D}_k)$ be its degree partition. Then the nodes $\mathcal{N}$ can be partitioned in independent sets $\mathcal{D}_i$, $i = 1, \ldots, \left\lfloor \frac{k}{2} \right\rfloor$ and a dominating set $\bigcup_{i=\lfloor \frac{k}{2} \rfloor + 1}^{k} \mathcal{D}_i$ in the graph $G' = (\mathcal{N} \backslash \mathcal{D}_0, \mathcal{E})$. Moreover, the neighborhoods of the nodes are nested. In particular, for each node $v \in \mathcal{D}_i$, $i = 1, \ldots, k$,*

$$\mathcal{N}_v = \begin{cases} \bigcup_{j=1}^{i} \mathcal{D}_{k+1-j} & \text{if } i = 1, \ldots, \left\lfloor \frac{k}{2} \right\rfloor, \\ \bigcup_{j=1}^{i} \mathcal{D}_{k+1-j} \setminus \{v\} & \text{if } i = \left\lfloor \frac{k}{2} \right\rfloor + 1, \ldots, k. \end{cases} \tag{36}$$

In the following, we will call the sets $\mathcal{D}_i$, $i = \left\lfloor \frac{k}{2} \right\rfloor + 1, \ldots, k$, *dominating subsets*, since the set $\mathcal{D}_i$ induces a dominating set in the graph obtained by removing the nodes in the set $\bigcup_{j=0}^{k-i} \mathcal{D}_j$ from G.

A nested split graph has an associated adjacency matrix which is called *stepwise matrix* and it is defined as follows:

Definition 3 (Brualdi and Hoffman [1985]). *A stepwise matrix $\mathbf{A}$ is a symmetric, binary $(n \times n)$-matrix with elements a_{ij} satisfying the condition: if $i < j$ and $a_{ij} = 1$ then $a_{hk} = 1$ whenever $h < k \leq j$ and $h \leq i$.*

If a nested split graph is connected we call it a connected nested split graph. From the stepwise property of the adjacency matrix, it follows that a connected nested split graph contains at least one spanning star, that is, there is at least one agent that is connected to all other agents (see e.g. König et al. [2014] for further properties).

A.4. Bonacich Centrality

We introduce a network measure capturing the centrality of a firm in the network due to Bonacich [1987]. Let $\mathbf{A}$ be the symmetric $n \times n$ *adjacency matrix* of the network G and λ_{PF} its

[51] Let x be a real valued number $x \in \mathbb{R}$. Then, $\lceil x \rceil$ denotes the smallest integer larger or equal than x (the ceiling of x). Similarly, $\lfloor x \rfloor$ denotes the largest integer smaller or equal than x (the floor of x).

largest real eigenvalue. The matrix $\mathbf{M}(G,\phi) = (\mathbf{I}-\phi\mathbf{A})^{-1}$ exists and is non-negative if and only if $\phi < 1/\lambda_{\mathrm{PF}}$.[52] Then

$$\mathbf{M}(G,\phi) = \sum_{k=0}^{\infty} \phi^k \mathbf{A}^k. \tag{37}$$

The Bonacich centrality vector is given by

$$\mathbf{b_u}(G,\phi) = \mathbf{M}(G,\phi) \cdot \mathbf{u}, \tag{38}$$

where $\mathbf{u} = (1,\ldots,1)^\top$. We can write the Bonacich centrality vector as

$$\mathbf{b_u}(G,\phi) = \sum_{k=0}^{\infty} \phi^k \mathbf{A}^k \cdot \mathbf{u} = (\mathbf{I} - \phi\mathbf{A})^{-1} \cdot \mathbf{u}.$$

For the components $b_{\mathbf{u},i}(G,\phi)$, $i = 1,\ldots,n$, we get

$$b_{\mathbf{u},i}(G,\phi) = \sum_{k=0}^{\infty} \phi^k (\mathbf{A}^k \cdot \mathbf{u})_i = \sum_{k=0}^{\infty} \phi^k \sum_{j=1}^{n} \left(\mathbf{A}^k\right)_{ij}. \tag{39}$$

Because $\sum_{j=1}^{n} \left(\mathbf{A}^k\right)_{ij}$ counts the number of all walks of length k in G starting from i, $b_{\mathbf{u},i}(G,\phi)$ is the number of all walks in G starting from i, where the walks of length k are weighted by their geometrically decaying factor ϕ^k.

Observe that we can also define the *weighted* Bonacich centrality exactly as above but when $\mathbf{u}$ is not anymore the $(n \times 1)$ vector of 1 but any $(n \times 1)$ vector.

The Bonacich matrix of Equation (37) is also a measure of structural similarity of the firms in the network, called *regular equivalence*. Blondel et al. [2004]; Leicht et al. [2006] define a similarity score b_{ij}, which is high if nodes i and j have neighbors that themselves have high similarity, given by $b_{ij} = \phi \sum_{k=1}^{n} a_{ik} b_{kj} + \delta_{ij}$. In matrix-vector notation this reads $\mathbf{M} = \phi\mathbf{A}\mathbf{M} + \mathbf{I}$. Rearranging yields $\mathbf{M} = (\mathbf{I} - \phi\mathbf{A})^{-1} = \sum_{k=0}^{\infty} \phi^k \mathbf{A}^k$, assuming that $\phi < 1/\lambda_{\mathrm{PF}}$. We hence obtain that the similarity matrix $\mathbf{M}$ is equivalent to the Bonacich matrix from Equation (37). The average similarity of firm i is $\frac{1}{n}\sum_{j=1}^{n} b_{ij} = \frac{1}{n} b_{\mathbf{u},i}(G,\phi)$, where $b_{\mathbf{u},i}(G,\phi)$ is the Bonacich centrality of i. It follows that the Bonacich centrality of i is proportional to the average regular equivalence of i. Firms with a high Bonacich centrality are then the ones which also have a high average structural similarity with the other firms in the R&D network.

The interpretation of eingenvector-like centrality measures as a similarity index is also important in the study of correlations between observations in principal component analysis and factor analysis [cf. Rencher and Christensen, 2012]. Variables with similar factor loadings can be grouped together. This basic idea has also been used in the economics literature on segregation [e.g. Ballester and Vorsatz, 2014; Echenique and Fryer Jr., 2007; Echenique et al., 2006].

Since equilibrium profits are closely related the the Bonacich centralities of the firms in the network, it is worth introducing a connection between the Bonacich centrality of a node and its coreness in the network. Coreness is defined as follows: Given a network G, the induced subgraph $G_k \subseteq G$ is the k-core of G if it is the largest subgraph such that the degree of all nodes in G_k is at least k. Note that the cores of a graph are nested such that $G_{k+1} \subseteq G_k$. Cores can be used as a measure of centrality in the network G. Note that k-cores can be obtained by a simple pruning algorithm: at each step, we remove all nodes with degree less than k. We repeat this procedure until there exist no such nodes or all nodes are removed. We define the coreness of each node as follows: The coreness of node i, cor_i, is k if and only

[52] The proof can be found e.g. in Debreu and Herstein [1953].

if $i \in G_k$ and $i \notin G_{k+1}$. We have that $\text{cor}_i \leq d_i$. However, there is no other relation between the degree and coreness of nodes in a graph. We then have the following result due to Manshadi and Johari [2010], which relates the Nash equilibrium to the k-cores of the graph: If $\text{cor}_i = k$ then $b_i(G,\phi) \geq \frac{1}{1-\phi k}$, where the inequality is tight when i belongs to a disconnected clique of size $k+1$. The coreness of networks of firms has also been studied empirically in Kitsak et al. [2010] and Rosenkopf and Schilling [2007]. In particular, Kitsak et al. [2010] find that the coreness of a firm correlates with its market value. We can easily explain this from our model because we know that firms in higher cores tend to have higher Bonacich centrality, and therefore higher sales and profits (cf. Proposition 1).

B. Bertrand Competition

In the case of price setting firms we obtain from the profit function in Equation (3) the FOC with respect to price p_i for firm i

$$\frac{\partial \pi_i}{\partial p_i} = (p_i - c_i)\frac{\partial q_i}{\partial p_i} - q_i = 0.$$

When $i \in \mathcal{M}_m$, then observe that from the inverse demand in Equation (1) we find that

$$q_i = \frac{\alpha_m(1-\rho_m) - (1-(n_m-2)\rho_m)p_i + \rho_m \sum_{j\in\mathcal{M}_m,\, j\neq i} p_j}{(1-\rho)(1+(n_m-1)\rho_m)},$$

where $n_m \equiv |\mathcal{M}_m|$. It then follows that

$$\frac{\partial q_i}{\partial p_i} = -\frac{1-(n_m-2)\rho_m}{(1-\rho_m)(1+(n_m-1)\rho_m)}.$$

Inserting into the FOC with respect to p_i gives

$$q_i = -\frac{1-(n_m-2)\rho_m}{(1-\rho_m)(1+(n_m-1)\rho_m)}(p_i - c_i).$$

Inserting Equations (1) and (2) yields

$$\begin{aligned} q_i = & \frac{(1-(n_m-2)\rho_m)(\alpha_m - \bar{c}_i)}{\rho_m(4-(2-\rho_m)n_m - \rho_m)} - \frac{1-(n_m-2)\rho_m}{4-(2-\rho_m)n_m-\rho_m} \sum_{\substack{j\in\mathcal{M}_m,\\ j\neq i}} q_j \\ & + \frac{(1-(n_m-2)\rho_m)}{\rho_m(4-(2-\rho_m)n_m - \rho_m} e_i + \frac{(1-(n_m-2)\rho_m)\varphi}{\rho_m(4-(2-\rho_m)n_m-\rho_m} \sum_{j=1}^n a_{ij}e_j. \end{aligned}$$

The FOC with respect to R&D effort is the same as in the case of perfect competition, so that we get $e_i = q_i$. Inserting equilibrium effort and rearranging terms gives

$$\begin{aligned} q_i = & \frac{(1-(n_m-2)\rho_m)(\alpha_m - \bar{c}_i)}{\rho_m(4-(2-\rho_m)n_m-\rho_m) - 1(1-(n_m-2)\rho_m)} \\ & - \frac{\rho_m(1-(n_m-2)\rho_m)}{\rho_m(4-(2-\rho_m)n_m-\rho_m) - 1(1-(n_m-2)\rho_m)} \sum_{\substack{j\in\mathcal{M}_m,\\ j\neq i}} q_j \\ & + \frac{\varphi(1-(n_m-2)\rho_m)}{\rho_m(4-(2-\rho_m)n_m-\rho_m) - 1(1-(n_m-2)\rho_m)} \sum_{j=1}^n a_{ij} q_j. \end{aligned}$$

If we denote by

$$\mu_i \equiv \frac{(1-(n_m-2)\rho_m)(\alpha_m-\bar{c}_i)}{\rho_m(4-(2-\rho_m)n_m-\rho_m)-1(1-(n_m-2)\rho_m)},$$
$$\rho \equiv \frac{\rho_m(1-(n_m-2)\rho_m)}{\rho_m(4-(2-\rho_m)n_m-\rho_m)-1(1-(n_m-2)\rho_m)},$$
$$\lambda \equiv \frac{\varphi(1-(n_m-2)\rho_m)}{\rho_m(4-(2-\rho_m)n_m-\rho_m)-1(1-(n_m-2)\rho_m)}.$$

Then we can write equilibrium quantities as follows

$$q_i = \mu_i - \rho \sum_{j=1}^{n} b_{ij} q_j + \lambda \sum_{j=1}^{n} a_{ij} q_j. \tag{40}$$

Observe that the reduced form Equation (40) is identical to the Cournot case in Equation (44).

C. Additional Results on Welfare

Here, we analyze welfare issues for a particular class of networks, namely the ones with a large spectral gap, such that λ_1 is much larger than λ_j for all $j \geq 2$.[53] These networks not only allow for a more explicit computation of welfare, but they are also representative for many real-world networks with a fat-tailed degree distribution,[54] as we observe it in our data (see Figure 5).

Proposition 8. *Consider substitutable goods and assume that $\mu_i = \mu$ for all $i = 1, \ldots, n$, and let ϕ and μ be defined as in Proposition 1. Then in the limit of ϕ approaching the inverse of the largest eigenvalue λ_{PF} from below welfare can be written as*

$$\lim_{\phi \uparrow 1/\lambda_{PF}} W(G) = \frac{2-\rho}{2} \frac{\mu^2}{\rho^2} \left(\frac{\rho}{2-\rho} + \frac{1}{\|\mathbf{v}_1\|_1^2} \right).$$

Further, denote by $\mathcal{G}(n)$ the class of graphs with n nodes and the class of graphs with n nodes and m links by $\mathcal{H}(n,m) \subset \mathcal{G}(n)$. Consider the class $\mathcal{S}(n,m) \subset \mathcal{H}(n,m)$ of graphs with a large spectral gap, such that $\lambda_1 = \lambda_{PF}$ is much larger than λ_j for all $j \geq 2$. Then the welfare maximizing graph $G^* = \text{argmax}_{G \in \mathcal{S}(n,m)} W(G)$ in this class is the one that minimizes the ℓ^1-norm $\|\mathbf{v}_1\|_1$ of the principal eigenvector $\mathbf{v}_1$ associated with the largest eigenvalue λ_1.

Proposition 8 implies that the social planner's problem reduces to finding the principal eigenvector of $\mathbf{A}$. For this problem there exist efficient algorithms, e.g. by using the power iteration method [Mises and Pollaczek-Geiringer, 1929].

Note that the norm $\|\mathbf{v}_1\|_1$ is the projection of the principal eigenvector $\mathbf{v}_1$ onto the all ones vector $\mathbf{u}$,

$$\|\mathbf{v}_1\|_1 = \|\mathbf{v}_1\|_2 \|\mathbf{u}\|_2 \cos(\alpha_1) = \sqrt{n} \cos(\alpha_1),$$

where α_1 is the angle between the vector $\mathbf{v}_1$ and $\mathbf{u}$. α_1 is called the principal *graph angle* [Cvetkovic et al., 1997, Chap. 4.5]. Welfare can then be written in terms of the graph angle α_1 as follows

$$\lim_{\phi \uparrow 1/\lambda_{\mathrm{PF}}} W(G) = \frac{2-\rho}{2} \frac{\mu^2}{\rho^2} \left(\frac{\rho}{2-\rho} + \frac{1}{n \cos(\alpha_1)^2} \right).$$

[53] The spectral gap is defined as $\lambda_1 - \lambda_2$. It is maximal in the complete graph K_n where it is equal to n. In the star $K_{1,n-1}$ we get $\lambda_1 - \lambda_2 = \sqrt{n-1}$. In a k-regular graph we obtain $\lambda_1 - \lambda_2 = \mu_{n-1}$, where μ_{n-1} is the second smallest eigenvalue of the Laplacian $\mathbf{Q} = \text{diag}(\mathbf{d}) - \mathbf{A}$ and $\mathbf{d}$ is the vector of degrees in G.

[54] Mihail and Papadimitriou [2002] have shown that networks with a power-law degree distribution also have a power-law eigenvalue distribution. See also Anderson et al. [2010]; Dorogovtsev et al. [2003].

Consider the spectral decomposition of the matrix $\mathbf{A}$ given by $\mathbf{A} = \sum_{i=1}^{n} \lambda_i \mathbf{v}_i \mathbf{v}_i^\top$, then the principal graph angle satisfies $\cos(\alpha_1)^2 = \frac{1}{n}\|\mathbf{v}_1\mathbf{v}_1^\top\|_1$. Moreover, its value is maximal for the regular graph G_k, where it is one. We thus have that $\|\mathbf{v}_1\|_1^2 \leq n$, and we obtain a lower bound for welfare given by

$$\lim_{\phi \uparrow 1/\lambda_{\mathrm{PF}}} W(G) \geq W(G_k) = \frac{2-\rho}{2}\frac{\mu^2}{\rho^2}\left(\frac{\rho}{2-\rho} + \frac{1}{n}\right),$$

which is the welfare function in the regular graph. For the star $G = K_{1,n-1}$ the principal eigenvector is given by $\mathbf{v}_1 = \frac{1}{\sqrt{2(n-1)}}(\sqrt{n-1}, 1, \ldots, 1)^\top$ where the corresponding largest eigenvalue is $\lambda_1 = \sqrt{n-1}$. In this case $(\mathbf{v}_1^\top \mathbf{u})^2 = \frac{1}{2}(2\sqrt{n-1} + n)$, and we obtain a lower bound on welfare in the efficient graph given by

$$\lim_{\phi \uparrow 1/\lambda_{\mathrm{PF}}} W(K_{1,n-1}) = \frac{2-\rho}{2}\frac{\mu^2}{\rho^2}\left(\frac{\rho}{2-\rho} + \frac{2}{2\sqrt{n-1}+n}\right).$$

This is larger than the value we have obtained for the regular graph.[55] Note that the star has a higher degree variance than the regular graph. This indicates that the result of Proposition 4 does not hold for large values of the spillover parameter ϕ. Moreover, the star is dissortative while the regular or complete graphs are not.

The quantity $\|\mathbf{v}_1\|_1^2 = \left(\sum_{i=1}^{n} v_{1i}\right)^2$ has been called *mixedness* of G by Rucker et al. [2002], since it relates to the variance of the principal eigenvector components as follows[56]

$$\sigma_{\mathbf{v}_1}^2 = \frac{1}{n-1}\left(\sum_{i=1}^{n} v_{1i}^2 - \frac{1}{n}\left(\sum_{i=1}^{n} v_{1i}\right)^2\right) = \frac{n - \|\mathbf{v}_1\|_1^2}{n(n-1)}.$$

The variance $\sigma_{\mathbf{v}_1}^2$ is decreasing in $\|\mathbf{v}_1\|_1$, and it is minimal for the regular graph where $v_{1i} = 1/\sqrt{n}$ for all $i = 1, \ldots, n$, that is to say they are maximally mixed. Welfare can then be written as

$$\lim_{\phi \uparrow 1/\lambda_{\mathrm{PF}}} W(G) = \frac{2-\rho}{2}\frac{\mu^2}{\rho^2}\left(\frac{\rho}{2-\rho} + \frac{1}{n(1-(n-1)\sigma_{\mathbf{v}_1}^2)}\right).$$

This suggests that the welfare maximizing graph (among the graphs with a large spectral gap) is eigenvector heterogeneous, or minimally mixed. Rucker et al. [2002] have shown by means of numerical computations for all networks of size $n \leq 10$ that graphs called *k-kites* minimize the mixedness.

A graph with a principal eigenvalue λ_1 contains the more walks, the larger is $\|\mathbf{v}_1\|_1^2$. Moreover, the reciprocal $1/\|\mathbf{v}_1\|_1^2$ measures the share of self returning walks among all walks. It follows that, a small value of $\|\mathbf{v}_1\|_1^2$ implies a large share of self returning walks, or a small probability that a randomly chosen walk ends at a vertex other than its origin. In terms of our model, where the network governs the way knowledge spillovers and diffusion are directed between firms, we thus find that the welfare maximizing graph has a large share of self returning walks, that is, knowledge originating in a firm passes through others before returning to its originator. This indicates that maximizing the cross-fertilization of knowledge and *knowledge recombination* between firms is welfare enhancing [cf. Weitzman, 1998]. It must be noted, however, that we do not model explicitly the heterogeneous technology portfolios that firms possess, but a reduced form in which firms benefit from R&D of their collaboration

[55] Observe that $W(K_{1,n-1}) = \frac{2}{n+2\sqrt{n-1}} > W(K_n) = \frac{1}{n}$ and $\lim_{n\to\infty} W(K_{1,n-1})/W(K_n) = \frac{2n}{n+2\sqrt{n-1}} = 2$.

[56] An alternative way to write the norm is $\|\mathbf{v}_1\|_1^2 = n - \sum_{j=1}^{n}\sum_{l=1}^{j-1}(v_{kj} - v_{kl})^2$ [Van Mieghem, 2011, p.40], which shows that $\|\mathbf{v}_1\|_1^2$ is maximal for an eigenvector $\mathbf{v}_1$ with minimal difference between its components.

partners.[57]

D. Intra- versus Interindustry Collaborations: Theory

We extend our model by allowing for intra-industry technology spillovers to differ from inter-industry spillovers. The profit of firm $i \in \mathcal{N}$ is still given by $\pi_i = (p_i - c_i)q_i - \frac{1}{2}e_i^2$, where the inverse demand is $p_i = \bar{\alpha}_i - q_i - \rho \sum_{j=1}^n b_{ij} q_j$. The main change is in the marginal cost of production, which is now equal to

$$c_i = \bar{c}_i - e_i - \varphi_1 \sum_{j=1}^n a_{ij}^{(1)} e_j - \varphi_2 \sum_{j=1}^n a_{ij}^{(2)} e_j,$$

where we have introduced two matrices $\mathbf{A}^{(1)}$ and $\mathbf{A}^{(2)}$ with elements $a_{ij}^{(1)}$ and $a_{ij}^{(2)}$, respectively, indicating a collaboration within the same industry (with the superscript (1)) or across different industries (with the superscript (2)). Note that we can write $\mathbf{A}^{(1)} = \mathbf{A} \circ \mathbf{B}$ and $\mathbf{A}^{(2)} = \mathbf{A} \circ (\mathbf{U} - \mathbf{B})$, with the matrix $\mathbf{B}$ having elements $b_{ij} \in \{0,1\}$ indicating whether firms i and j operate in the same market or not, $\mathbf{U}$ being a matrix of all ones, and $\circ$ denotes the Hadamard elementwise matrix product.[58] Inserting this marginal cost of production into the profit function gives

$$\pi_i = (\bar{\alpha}_i - \bar{c}_i)q_i - q_i^2 - \rho q_i \sum_{j=1}^n b_{ij} q_j + q_i e_i + \varphi_1 q_i \sum_{j=1}^n a_{ij}^{(1)} e_j + \varphi_2 q_i \sum_{j=1}^n a_{ij}^{(2)} e_j - \frac{1}{2} e_i^2.$$

As above, from the first-order condition with respect to R&D effort, we obtain $e_i = q_i$. Inserting this optimal effort into the first-order condition with respect to output, we obtain

$$q_i = \bar{\alpha}_i - \bar{c}_i - \rho \sum_{j=1}^n b_{ij} q_j + \varphi_1 \sum_{j=1}^n a_{ij}^{(1)} q_j + \varphi_2 \sum_{j=1}^n a_{ij}^{(2)} q_j.$$

Denoting by $\mu_i \equiv \bar{\alpha}_i - \bar{c}_i$, we can write this as

$$q_i = \mu_i - \rho \sum_{j=1}^n b_{ij} q_j + \varphi_1 \sum_{j=1}^n a_{ij}^{(1)} q_j + \varphi_2 \sum_{j=1}^n a_{ij}^{(2)} q_j. \tag{41}$$

If the matrix $\mathbf{I}_n + \rho\mathbf{B} - \varphi_1 \mathbf{A}^{(1)} - \varphi_2 \mathbf{A}^{(2)}$ is invertible, this gives us the equilibrium quantities

$$\mathbf{q} = (\mathbf{I}_n + \rho\mathbf{B} - \varphi_1 \mathbf{A}^{(1)} - \varphi_2 \mathbf{A}^{(2)})^{-1} \boldsymbol{\mu}.$$

Let us now write the econometric equivalent of Equation (41). Proceeding as in Section 8.1, using Equations (23) and (24) and introducing time t, we get

$$\mu_{it} = \mathbf{x}_{it}^\top \boldsymbol{\beta} + \eta_i + \kappa_t + \epsilon_{it}.$$

[57]A model of R&D network formation in which the technology portfolios of firms are explicitly considered is analyzed in König [2012].

[58]Let $\mathbf{A}$ and $\mathbf{B}$ be $m \times n$ matrices. The Hadamard product of $\mathbf{A}$ and $\mathbf{B}$ is defined by $[\mathbf{A} \circ \mathbf{B}]_{ij} = [\mathbf{A}]_{ij}[\mathbf{B}]_{ij}$ for all $1 \leq i \leq m$, $1 \leq j \leq n$, i.e. the Hadamard product is simply an element-wise multiplication.

Plugging this value of μ_{it} into Equation (41), we obtain

$$q_{it} = \varphi_1 \sum_{j=1}^{n} a_{ij,t}^{(1)} q_{jt} + \varphi_2 \sum_{j=1}^{n} a_{ij,t}^{(2)} q_{jt} - \rho \sum_{j=1}^{n} b_{ij} q_{jt} + \mathbf{x}_{it}^{\top} \boldsymbol{\beta} + \eta_i + \kappa_t + \epsilon_{it},$$

where $a_{ij,t}^{(1)} = a_{ij,t} b_{ij}$ and $a_{ij,t}^{(2)} = a_{ij,t}(1 - b_{ij})$. This is Equation (31) in Section 9.2.3.

E. Direct and Indirect Technology Spillovers: Theory

We extend our model by allowing for direct (between collaborating firms) and indirect (between non-collaborating firms) technology spillovers. The profit of firm $i \in \mathcal{N}$ is still given by $\pi_i = (p_i - c_i) q_i - \frac{1}{2} e_i^2$, where the inverse demand is $p_i = \bar{\alpha}_i - q_i - \rho \sum_{j=1}^{n} b_{ij} q_j$. The main change is in the marginal cost of production, which is now equal to[59]

$$c_i = \bar{c}_i - e_i - \varphi \sum_{j=1}^{n} a_{ij} e_j - \chi \sum_{j=1}^{n} w_{ij} e_j, \tag{42}$$

where w_{ij} are weights characterizing alternative channels for technology spillovers than R&D collaborations (representing for example a patent cross-citation, a flow of workers, or technological proximity measured by the matrix P_{ij} introduced in Footnote 35). Inserting this marginal cost of production into the profit function gives

$$\pi_i = (\bar{\alpha}_i - \bar{c}_i) q_i - q_i^2 - \rho q_i \sum_{j=1}^{n} b_{ij} q_j + q_i e_i + \varphi q_i \sum_{j=1}^{n} a_{ij} e_j + \chi q_i \sum_{j=1}^{n} w_{ij} e_j - \frac{1}{2} e_i^2.$$

As above, from the first-order condition with respect to R&D effort, we obtain $e_i = q_i$. Inserting this optimal effort into the first-order condition with respect to output, we obtain

$$q_i = \bar{\alpha}_i - \bar{c}_i - \rho \sum_{j=1}^{n} b_{ij} q_j + \varphi \sum_{j=1}^{n} a_{ij} q_j + \chi \sum_{j=1}^{n} w_{ij} q_j.$$

Denoting by $\mu_i \equiv \bar{\alpha}_i - \bar{c}_i$, we can write this as

$$q_i = \mu_i - \rho \sum_{j=1}^{n} b_{ij} q_j + \varphi \sum_{j=1}^{n} a_{ij} q_j + \chi \sum_{j=1}^{n} w_{ij} q_j. \tag{43}$$

If the matrix $\mathbf{I}_n + \rho \mathbf{B} - \varphi \mathbf{A} - \chi \mathbf{W}$ is invertible, this gives us the equilibrium quantities

$$\mathbf{q} = (\mathbf{I}_n + \rho \mathbf{B} - \varphi \mathbf{A} - \chi \mathbf{W})^{-1} \boldsymbol{\mu}.$$

Let us now write the econometric equivalent of Equation (43). Proceeding as in Section 8.1, using Equations (23) and (24) and introducing time t, we get

$$\mu_{it} = \mathbf{x}_{it}^{\top} \boldsymbol{\beta} + \eta_i + \kappa_t + \epsilon_{it}.$$

Plugging this value of μ_{it} into Equation (43), we obtain

$$q_{it} = \varphi \sum_{j=1}^{n} a_{ij,t} q_{jt} + \chi \sum_{j=1}^{n} w_{ij,t} q_{jt} - \rho \sum_{j=1}^{n} b_{ij} q_{jt} + \mathbf{x}_{it}^{\top} \boldsymbol{\beta} + \eta_i + \kappa_t + \epsilon_{it}.$$

[59] See also Eq. (1) in Goyal and Moraga-Gonzalez [2001].

This is Equation (33) in Section 9.2.4.

F. Proofs

Before we proceed with the proof of Proposition 1 we state the following lemma which will be needed in the proof.

Lemma 1. *Let* $\mathbf{A}$ *and* $\mathbf{B}$ *be two symmetric, real matrices and assume that the inverse* $\mathbf{A}^{-1}$ *exists and is non-negative and also that* $\mathbf{B}$ *is non-negative. Provided that* $\lambda_{\max}(\mathbf{A}^{-1}\mathbf{B}) < 1$ *we have that*

(i) the following series expansion exists

$$(\mathbf{A}+\mathbf{B})^{-1} = \sum_{k=0}^{\infty}(-1)^k(\mathbf{A}^{-1}\mathbf{B})^k\mathbf{A}^{-1},$$

(ii) for any $\mathbf{x} \in \mathbf{R}^n_+$ *we have that* $\mathbf{A}^{-1}\mathbf{B}\mathbf{x} < \mathbf{x}$*, and*

(iii) if also $\mathbf{A}^{-1}\mathbf{x} > \mathbf{0}$ *then* $(\mathbf{A}+\mathbf{B})^{-1}\mathbf{x} > \mathbf{0}$*.*

Proof of Lemma 1. We first prove part (i) of the lemma. Notice that

$$\begin{aligned}(\mathbf{A}+\mathbf{B})^{-1} &= (\mathbf{A}(\mathbf{I}_n + \mathbf{A}^{-1}\mathbf{B}))^{-1} \\ &= (\mathbf{I}_n + \mathbf{A}^{-1}\mathbf{B}))^{-1}\mathbf{A}^{-1} \\ &= \sum_{k=0}^{\infty}(-1)^k(\mathbf{A}^{-1}\mathbf{B})^k\mathbf{A}^{-1},\end{aligned}$$

where the Neumann series expansion for $(\mathbf{I}_n + \mathbf{A}^{-1}\mathbf{B}))^{-1}$ can be applied if $\lambda_{\max}(\mathbf{A}^{-1}\mathbf{B}) < 1$.

For the second part (ii), observe that $\lambda_{\max}(\mathbf{A}^{-1}\mathbf{B}) < 1$ is equivalent to $\mathbf{A}^{-1}\mathbf{B}\mathbf{x} < \mathbf{x}$ for any $\mathbf{x} \in \mathbf{R}^n_+$. To see this consider an orthonormal basis of $\mathbf{R}^n$ spanned by the eigenvectors of $\mathbf{A}^{-1}\mathbf{B}$. Then we can write $\mathbf{x} = \sum_{i=1}^n c_i\mathbf{v}_i$ with suitable coefficients $c_i = \mathbf{x}^\top\mathbf{v}_i/(\mathbf{v}_i^\top\mathbf{v}_i)$ and $\mathbf{A}^{-1}\mathbf{B}\mathbf{v}_i = \lambda_i\mathbf{v}_i$. It then follows that

$$\mathbf{A}^{-1}\mathbf{B}\mathbf{x} = \sum_{i=1}^n c_i\lambda_i\mathbf{v}_i \leq \lambda_{\max}(\mathbf{A}^{-1}\mathbf{B})\sum_{i=1}^n c_i\mathbf{v}_i = \lambda_{\max}(\mathbf{A}^{-1}\mathbf{B})\mathbf{x}.$$

Hence, if $\lambda_{\max}(\mathbf{A}^{-1}\mathbf{B}) < 1$ it must hold that $\mathbf{A}^{-1}\mathbf{B}\mathbf{x} < \mathbf{x}$.

For part (iii) of the proof note that we can write the series expansion of the inverse as follows

$$(\mathbf{A}+\mathbf{B})^{-1}\mathbf{x} = \sum_{k=0}^{\infty}(-1)^k(\mathbf{A}^{-1}\mathbf{B})^k\mathbf{A}^{-1}\mathbf{x} = \mathbf{A}^{-1}\mathbf{x} - \mathbf{A}^{-1}\mathbf{B}\mathbf{A}^{-1}\mathbf{x} + \mathbf{A}^{-1}\mathbf{B}\mathbf{A}^{-1}\mathbf{B}\mathbf{A}^{-1}\mathbf{x} - \ldots.$$

By assumption we have that $\mathbf{A}^{-1}\mathbf{x} \geq \mathbf{0}$. Then denote by $\tilde{\mathbf{x}} = \mathbf{A}^{-1}\mathbf{x} \geq \mathbf{0}$. Then the first two terms in the series can be written as

$$(\mathbf{I}_n - \mathbf{A}^{-1}\mathbf{B})\mathbf{A}^{-1}\mathbf{x} = (\mathbf{I}_n - \mathbf{A}^{-1}\mathbf{B})\tilde{\mathbf{x}} > 0$$

where the inequality follows from part (ii) of the lemma. Next, consider the third and fourth terms in the series expansion

$$(\mathbf{A}^{-1}\mathbf{B}\mathbf{A}^{-1}\mathbf{B} - \mathbf{A}^{-1}\mathbf{B}\mathbf{A}^{-1}\mathbf{B}\mathbf{A}^{-1}\mathbf{B})\tilde{\mathbf{x}} = \mathbf{A}^{-1}\mathbf{B}\mathbf{A}^{-1}\mathbf{B}(\mathbf{I}_n - \mathbf{A}^{-1}\mathbf{B})\tilde{\mathbf{x}} \geq 0,$$

where the inequality follows again from the fact that $(\mathbf{I}_n - \mathbf{A}^{-1}\mathbf{B})\tilde{\mathbf{x}} > 0$ from part (ii) of the lemma and the assumption that $\mathbf{A}^{-1}$ and $\mathbf{B}$ are non-negative matrices. We can then iterate by induction to show the desired claim. □

Proof of Proposition 1. Let us start with the most general setup, i.e. case (i). The profit of firm $i \in \mathcal{N}$ is given by

$$\pi_i = (\bar{\alpha}_i - \bar{c}_i)q_i - q_i^2 - \rho\sum_{j\in\mathcal{M}_m, j\neq i} q_iq_j + q_ie_i + \varphi q_i\sum_{j=1}^n a_{ij}e_j - \frac{1}{2}e_i^2,$$

where $b_{ij} \in \{0,1\}$ is the ij-th element of the $n \times n$ matrix $\mathbf{B}$ defined by

$$\mathbf{B} \equiv \sum_{m=1}^{M} (\mathbf{u}_m \mathbf{u}_m^\top - \mathbf{D}_m) = \left(\begin{array}{cccc|cccc|c} 0 & 1 & \cdots & 1 & 0 & \cdots & \cdots & 0 & \cdots \\ 1 & 0 & \cdots & \vdots & \vdots & & & \vdots & \\ \vdots & \vdots & \ddots & 1 & \vdots & & & \vdots & \\ 1 & \cdots & 1 & 0 & 0 & \cdots & \cdots & 0 & \cdots \\ \hline 0 & \cdots & \cdots & 0 & 0 & 1 & \cdots & 1 & \\ \vdots & & & \vdots & 1 & 0 & \cdots & \vdots & \\ \vdots & & & \vdots & \vdots & \vdots & \ddots & 1 & \\ 0 & \cdots & \cdots & 0 & 1 & \cdots & 1 & 0 & \\ \hline \vdots & & & \vdots & & & & & \ddots \end{array} \right)_{n \times n},$$

and $\mathbf{u}_m$ is a $n \times 1$ zero-one vector with elements $u_{mi} = 1$ if $i \in \mathcal{M}_m$ and $u_{mi} = 0$ otherwise for all $i = 1, \ldots, n$. Moreover, $\mathbf{D}_m = \operatorname{diag}(\mathbf{u}_m)$ is the diagonal matrix with diagonal entries given by $\mathbf{u}_m$. The FOC of profits with respect to R&D effort e_i of firm i is given by

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

so that we obtain

$$e_i = q_i.$$

The FOC with respect to quantity is given by

$$\frac{\partial \pi_i}{\partial q_i} = \bar{\alpha}_i - \bar{c}_i - 2q_i - \rho \sum_{j=1}^{n} b_{ij} q_j + e_i + \varphi \sum_{j=1}^{n} a_{ij} e_j.$$

Note that the Hessian is given by

$$\begin{pmatrix} \frac{\partial^2 \pi_i}{\partial q_i^2} & \frac{\partial^2 \pi_i}{\partial q_i \partial e_i} \\ \frac{\partial^2 \pi_i}{\partial e_i \partial q_i} & \frac{\partial^2 \pi_i}{\partial e_i^2} \end{pmatrix} = \begin{pmatrix} -2 & 1 \\ 1 & -1 \end{pmatrix}.$$

The eigenvalues of the Hessian are $\frac{1}{2}(-3 \pm \sqrt{5})$, which are both negative. Hence, the Hessian is negative definite, and the profit function is strictly quasi-concave in both variables q_i and e_i.

Next, inserting equilibrium effort and rearranging terms gives

$$q_i = (\bar{\alpha}_i - \bar{c}_i) - \rho \sum_{j=1}^{n} b_{ij} q_j + \varphi \sum_{j=1}^{n} a_{ij} q_j.$$

In the following we denote by $\mu_i \equiv \bar{\alpha}_i - \bar{c}_i$, so that we obtain for equilibrium quantity

$$q_i = \mu_i - \rho \sum_{j=1}^{n} b_{ij} q_j + \varphi \sum_{j=1}^{n} a_{ij} q_j. \tag{44}$$

In matrix-vector notation it can be written as

$$\mathbf{q} = \boldsymbol{\mu} - \rho \mathbf{B} \mathbf{q} + \varphi \mathbf{A} \mathbf{q}$$

or, equivalently,

$$(\mathbf{I}_n + \rho \mathbf{B} - \varphi \mathbf{A}) \mathbf{q} = \boldsymbol{\mu}.$$

The matrix $\mathbf{I}_n + \rho \mathbf{B} - \varphi \mathbf{A}$ is invertible if its determinant is not zero. This also guarantees the uniqueness and existence of the equilibrium. Following Lee and Liu [2010], the determinant of $\mathbf{I}_n - \sum_{j=1}^{p} \lambda_j \mathbf{W}_j$ is strictly positive if $\sum_{j=1}^{p} |\lambda_j| < 1/\max_{j=1,\ldots,p} \|\mathbf{W}_j\|$, where $\|\mathbf{W}_j\|$ is any matrix norm, including the spectral norm (which is the largest eigenvalue of $\mathbf{W}_j$). We have that the largest eigenvalue of the matrix $\mathbf{B}$ is equal to the size of the largest market $|\mathcal{M}_m|$ minus one (as this is a block-diagonal matrix with all elements being one in each block and zero diagonal), and the largest eigenvalue of $\mathbf{A}$ is the Perron-Frobenius eigenvalue $\lambda_{\mathrm{PF}}(\mathbf{A})$.

A sufficient condition for invertibility is then given by

$$\rho + \varphi < \left(\max \left\{ \lambda_{\mathrm{PF}}(\mathbf{A}), \max_{m=1,\dots,M} \{(|\mathcal{M}_m| - 1)\} \right\} \right)^{-1},$$

which can also be written as

$$1 + \varphi/\rho < \left(\max \left\{ \lambda_{\mathrm{PF}}(\mathbf{A}), \max_{m=1,\dots,M} \{(|\mathcal{M}_m| - 1)\rho_m\} \right\} \right)^{-1}.$$

When the inverse of $\mathbf{I}_n + \rho\mathbf{B} - \varphi\mathbf{A}$ exists, we can write equilibrium quantities as

$$\mathbf{q} = (\mathbf{I}_n + \rho\mathbf{B} - \varphi\mathbf{A})^{-1}\boldsymbol{\mu}.$$

We have shown that there exists a unique equilibrium given by $\mathbf{q} = (\mathbf{I}_n + \rho\mathbf{B} - \varphi\mathbf{A})^{-1}\boldsymbol{\mu}$, but we have not yet shown that it is interior, i.e. $q_i > 0$, $\forall i \in \mathcal{N}$. We will deal with corner solutions below. Profits in equilibrium can be written as

$$\pi_i = (\bar{\alpha}_i - \bar{c}_i)q_i - \rho q_i \sum_{j=1}^n b_{ij} q_j + \varphi q_i \sum_{j=1}^n a_{ij} q_j - \frac{1}{2} q_i^2.$$

From Equation (44) it follows that

$$\begin{aligned}
\rho q_i \sum_{j=1}^n b_{ij} q_j - \varphi q_i \sum_{j=1}^n a_{ij} q_j &= \rho q_i \sum_{j=1}^n b_{ij} q_j - \varphi q_i \sum_{j=1}^n a_{ij} q_j \\
&= q_i((\rho\mathbf{B} - \varphi\mathbf{A})\mathbf{q})_i \\
&= q_i((\mathbf{I}_n + \rho\mathbf{B} - \varphi\mathbf{A})\mathbf{q} - \mathbf{q})_i \\
&= q_i\left((\bar{\alpha}_i - \bar{c}_i) - q_i\right), \qquad (45)
\end{aligned}$$

so that we can write equilibrium profits as

$$\pi_i = (\bar{\alpha}_i - \bar{c}_i)q_i - q_i\left((\bar{\alpha}_i - \bar{c}_i) - q_i\right) - \frac{1}{2} q_i^2 = \frac{1}{2} q_i^2.$$

Let us now deal with case (ii) in the proposition, i.e. we assume that all firms operate in the same market so that $M = 1$. The first-order condition for a firm i is given by Equation (44), which, when $M = 1$, can be written as

$$q_i = \mu_i - \rho \sum_{j \neq i} q_j + \varphi \sum_{j=1}^n a_{ij} q_j$$

Let us have the following notation: $\overline{q}_{-i} \equiv \sum_{j \neq i} q_j$, which is the total ouput of all firms but excluding firm i. The equation above is equivalent to

$$q_i = \mu_i - \rho\overline{q}_{-i} + \varphi \sum_{j=1}^n a_{ij} q_j$$

We can now define $\overline{q} \equiv \sum_{j \neq i} q_j + q_i$, which corresponds to the total output of all firms (including i). The equation above is now equivalent to

$$q_i = \mu_i - \rho\overline{q} + \rho q_i + \varphi \sum_{j=1}^n a_{ij} q_j$$

$$\Leftrightarrow q_i = \frac{1}{1-\rho}\mu_i - \frac{\rho}{1-\rho}\overline{q} + \frac{\varphi}{1-\rho} \sum_{j=1}^n a_{ij} q_j \qquad (46)$$

Observe that even if firms are local monopolies (i.e. $\rho = 0$) this solution is still well-defined. Observe also that $1 - \rho > 0$ if and only if $\rho < 1$, which we assume throughout.

In matrix form, Equation (46) can be written as

$$\left[\mathbf{I}-\frac{\varphi}{1-\rho}\mathbf{A}\right]\mathbf{q}=\frac{1}{1-\rho}\boldsymbol{\mu}-\frac{\rho\overline{q}}{1-\rho}\mathbf{u}$$

where $\boldsymbol{\mu}=(\mu_1,\ldots,\mu_n)^\top$, and $\mathbf{u}=(1,\ldots,1)^\top$. Denote $\phi=\varphi/(1-\rho)$. If $\phi\lambda_{\mathrm{PF}}(\mathbf{A})<1$, this is equivalent to

$$\mathbf{q}=\left(\frac{1}{1-\rho}\right)(\mathbf{I}-\phi\mathbf{A})^{-1}\boldsymbol{\mu}-\frac{\rho\overline{q}}{1-\rho}(\mathbf{I}-\phi\mathbf{A})^{-1}\mathbf{u}$$

This equation is equivalent to:

$$\mathbf{q}=\left(\frac{1}{1-\rho}\right)\left[\mathbf{b}_{\boldsymbol{\mu}}(G,\phi)-\rho\overline{q}\,\mathbf{b}_{\mathbf{u}}(G,\phi)\right] \tag{47}$$

where $\mathbf{b}_{\mathbf{u}}(G,\varphi/(1-\rho))=(\mathbf{I}-\phi\mathbf{A})^{-1}\mathbf{u}$ is the *unweighted* vector of Bonacich centralities and $\mathbf{b}_{\boldsymbol{\mu}}(G,\varphi/(1-\rho))=(\mathbf{I}-\phi\mathbf{A})^{-1}\boldsymbol{\mu}$ is the *weighted* vector of Bonacich centralities where the weights are the μ_i for $i=1,\ldots,n$.[60]

We need now to calculate $\overline{q}$. Multiplying Equation (47) to the left by $\mathbf{u}^\top$, we obtain:

$$(1-\rho)\,\overline{q}=\left\|\mathbf{b}_{\boldsymbol{\mu}}(G,\phi)\right\|_1-\rho\overline{q}\left\|\mathbf{b}_{\mathbf{u}}(G,\phi)\right\|_1$$

where

$$\left\|\mathbf{b}_{\boldsymbol{\mu}}(G,\phi)\right\|_1=\mathbf{u}^{\mathrm{T}}\mathbf{b}_{\boldsymbol{\mu}}(G,\phi)=\sum_{i=1}^{n}b_{\mu_i}(G,\phi)=\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{p=0}^{+\infty}\phi^p a_{ij}^{[p]}\mu_j$$

is the sum of the weighted Bonacich centralities and

$$\left\|\mathbf{b}_{\mathbf{u}}(G,\phi)\right\|_1=\mathbf{u}^{\top}\mathbf{b}_{\mathbf{u}}(G,\phi)=\sum_{i=1}^{n}b_{u,i}(G,\phi)=\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{p=0}^{+\infty}\phi^p a_{ij}^{[p]}$$

is the sum of the unweighted Bonacich centralities. Solving this equation, we get:

$$\overline{q}=\frac{\left\|\mathbf{b}_{\boldsymbol{\mu}}(G,\phi)\right\|_1}{(1-\rho)+\rho\left\|\mathbf{b}_{\mathbf{u}}(G,\phi)\right\|_1}$$

Plugging this value of $\overline{q}$ into Equation (47), we finally obtain:

$$\mathbf{q}=\left(\frac{1}{1-\rho}\right)\left[\mathbf{b}_{\boldsymbol{\mu}}(G,\phi)-\frac{\rho\left\|\mathbf{b}_{\boldsymbol{\mu}}(G,\phi)\right\|_1}{1-\rho+\rho\left\|\mathbf{b}_{\mathbf{u}}(G,\phi)\right\|_1}\mathbf{b}_{\mathbf{u}}(G,\phi)\right] \tag{48}$$

This corresponds to Equation (9) in the proposition. For each firm i, we thus have

$$q_i=\left(\frac{1}{1-\rho}\right)\left[b_{\boldsymbol{\mu},i}(G,\phi)-\frac{\rho\left\|\mathbf{b}_{\boldsymbol{\mu}}(G,\phi)\right\|_1}{1-\rho+\rho\left\|\mathbf{b}_{\mathbf{u}}(G,\phi)\right\|_1}b_{\mathbf{u},i}(G,\phi)\right] \tag{49}$$

Next, we consider corner solutions and provide conditions which guarantee that the equilibrium is always interior. For that, we would like to show that $q_i>0$, $\forall i=1,\ldots,n$. Using Equation (49), this is equivalent to

$$b_{\boldsymbol{\mu},i}(G,\phi)>\frac{\rho\left\|\mathbf{b}_{\boldsymbol{\mu}}(G,\phi)\right\|_1}{1-\rho+\rho\left\|\mathbf{b}_{\mathbf{u}}(G,\phi)\right\|_1}b_{\mathbf{u},i}(G,\phi),\quad \forall i=1,\ldots,n. \tag{50}$$

[60] A definition and further discussion of the Bonacich centrality is given in Appendix A.4.

Denote by $\underline{\mu} = \max_i \{\mu_i \mid i \in N\}$ and $\overline{\mu} = \max_i \{\mu_i \mid i \in N\}$, with $\underline{\mu} < \overline{\mu}$. Then, $\forall i = 1, \ldots, n$, we have

$$\begin{aligned}
\|\mathbf{b}_{\mathbf{u}}(G,\phi)\|_1 &= \sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{p=0}^{\infty}\phi^p a_{ij}^{[p]}\mu_j \\
&\le \overline{\mu}\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{p=0}^{\infty}\phi^p a_{ij}^{[p]} = \overline{\mu}\,\|\mathbf{b}_{\mathbf{u}}(G,\phi)\|_1
\end{aligned}$$

and

$$b_{\boldsymbol{\mu},i}(G,\phi) = \sum_{j=1}^{n}\sum_{p=0}^{\infty}\phi^p a_{ij}^{[p]}\mu_j \ge \underline{\mu}\, b_{\mathbf{u},i}(G,\phi) = \sum_{j=1}^{n}\sum_{p=0}^{\infty}\phi^p a_{ij}^{[p]}\underline{\mu}$$

Thus, a sufficient condition for Equation (50) to hold is:

$$\underline{\mu}\, b_{\mathbf{u},i}(G,\phi) > \frac{\rho\overline{\mu}\,\|\mathbf{b}_{\mathbf{u}}(G,\phi)\|_1}{1-\rho+\rho\,\|\mathbf{b}_{\mathbf{u}}(G,\phi)\|_1} b_{\mathbf{u},i}(G,\phi)$$

or equivalently

$$\underline{\mu} > \frac{\rho\overline{\mu}\,\|\mathbf{b}_{\mathbf{u}}(G,\phi)\|_1}{1-\rho+\rho\,\|\mathbf{b}_{\mathbf{u}}(G,\phi)\|_1}$$

or

$$1-\rho > \rho\,\|\mathbf{b}_{\mathbf{u}}(G,\phi)\|_1\left(\frac{\overline{\mu}}{\underline{\mu}}-1\right) \tag{51}$$

Observe that, by definition,

$$\|\mathbf{b}_{\mathbf{u}}(G,\phi)\|_1 = \sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{p=0}^{\infty}\phi^p a_{ij}^{[p]} = \sum_{p=0}^{\infty}\phi^p\mathbf{u}^\top\mathbf{A}^p\mathbf{u} \tag{52}$$

We know that $\lambda_{\mathrm{PF}}(\mathbf{A}^p) = \lambda_{\mathrm{PF}}(\mathbf{A})^p$, for all $p \ge 0$.[61] Also, $\mathbf{u}^\top\mathbf{A}^p\mathbf{u}/n$ is the average connectivity in the matrix $\mathbf{A}^p$ of paths of length p in the original network $\mathbf{A}$, which is smaller than that its spectral radius $\lambda_{\mathrm{PF}}(\mathbf{A})^p$ [Cvetkovic et al., 1995], i.e. $\mathbf{u}^\top\mathbf{A}^p\mathbf{u}/n \le \lambda_{\mathrm{PF}}(\mathbf{A})^p$. Therefore, Equation (52) leads to the following inequality

$$\|\mathbf{b}_{\mathbf{u}}(G,\phi)\|_1 = \sum_{p=0}^{\infty}\phi^p\mathbf{u}^\top\mathbf{A}^p\mathbf{u} \le n\sum_{p=0}^{\infty}\phi^p\lambda_{\mathrm{PF}}(\mathbf{A})^p = \frac{n}{1-\phi\lambda_{\mathrm{PF}}(\mathbf{A})}.$$

A sufficient condition for Equation (51) to hold is thus

$$\phi\lambda_{\mathrm{PF}}(\mathbf{A}) + \frac{n\rho}{1-\rho}\left(\frac{\overline{\mu}}{\underline{\mu}}-1\right) < 1.$$

Clearly, this interior equilibrium is unique. This is the condition given in the proposition for case (ii).

Let us now go back to case (i) in the proposition and show that we have an interior equilibrium with all firms producing at positive quantity levels, that is $\mathbf{q} = (\mathbf{I}_n + \rho\mathbf{B} - \varphi\mathbf{A})^{-1}\boldsymbol{\mu} > \mathbf{0}$. To do this we would like to apply Lemma 1. Let $\mathbf{I}_n - \varphi\mathbf{A}$ be the matrix $\mathbf{A}$ in the lemma and $\rho\mathbf{B}$ the corresponding matrix $\mathbf{B}$. We have that both are real and symmetric, and that $\mathbf{B}$ is a non-negative matrix. Further, provided that $\varphi < 1/\lambda_{\mathrm{PF}}(\mathbf{A})$, the inverse $\mathbf{A}^{-1}$ exists and is non-negative. Next, we need to show that $\lambda_{\mathrm{PF}}(\mathbf{A}^{-1}\mathbf{B}) < 1$, but this is equivalent to

$$\lambda_{\mathrm{PF}}((\mathbf{I}_n - \varphi\mathbf{A})^{-1}\rho\mathbf{B}) < 1.$$

[61] Observe that $\lambda_{\mathrm{PF}}(\mathbf{A}^p) = \lambda_{\mathrm{PF}}(\mathbf{A})^p$ is true for both a *symmetric* and an *asymmetric* adjacency matrix $\mathbf{A}$ as long as $\mathbf{A}$ has non-negative entries $a_{ij} \ge 0$. This follows from the Perron-Frobenius theorem.

Note that

$$\begin{aligned}\lambda_{\mathrm{PF}}((\mathbf{I}_n-\varphi\mathbf{A})^{-1}\rho\mathbf{B}) &= \rho\lambda_{\mathrm{PF}}((\mathbf{I}_n-\varphi\mathbf{A})^{-1}\mathbf{B})\\ &\le \rho\lambda_{\mathrm{PF}}((\mathbf{I}_n-\varphi\mathbf{A})^{-1})\lambda_{\mathrm{PF}}(\mathbf{B})\\ &= \frac{\rho\lambda_{\mathrm{PF}}(\mathbf{B})}{1-\varphi\lambda_{\mathrm{PF}}(\mathbf{A})},\end{aligned}$$

so that a sufficient condition is given by

$$\frac{\rho\lambda_{\mathrm{PF}}(\mathbf{B})}{1-\varphi\lambda_{\mathrm{PF}}(\mathbf{A})} < 1,$$

which is implied by

$$\rho\lambda_{\mathrm{PF}}(\mathbf{B}) = \rho\max_{m=1,\ldots,M}\{(|\mathcal{M}_m|-1)\} < 1-\varphi\lambda_{\mathrm{PF}}(\mathbf{A}).$$

The lemma then implies that $(\mathbf{A}+\mathbf{B})^{-1}\mathbf{x} > \mathbf{0}$ for any vector $\mathbf{x} > \mathbf{0}$, and in particular for the vector $\boldsymbol{\mu}$, which is positive by assumption.

Consider now case (iii) where not only $M=1$ but also $\mu_i=\mu$ for all $i=1,\ldots,n$. If $\phi\lambda_{\mathrm{PF}}(\mathbf{A})<1$, the equilibrium condition in Equation (48) can be further simplified to

$$\mathbf{q} = \frac{\mu}{1-\rho+\rho\|\mathbf{b_u}(G,\phi)\|_1}\mathbf{b_u}(G,\phi). \tag{53}$$

It should be clear that the output is now always strictly positive.

Let us now consider case (iv) where markets are independent and goods are non-substitutable (i.e., $\rho=0$). If $\varphi<\lambda_{\mathrm{PF}}(\mathbf{A})^{-1}$, the equilibrium quantity further simplifies to $\mathbf{q}=\mu\mathbf{b_u}(G,\phi)$, which is always strictly positive. Equilibrium profit follows from Equation (11). □

Proof of Proposition 2. We first give a proof of part (ii) of the proposition. Assuming that $\mu_i=\mu$ for all $i=1,\ldots,n$, at the Nash equilibrium, we have that $\mathbf{q}=\mu\mathbf{M}(G,\varphi)\mathbf{u}$, where we have denoted by $\mathbf{M}(G,\varphi)\equiv(\mathbf{I}_n-\varphi\mathbf{A})^{-1}$.[62] We then obtain

$$W(G) = \mathbf{q}^\top\mathbf{q} = \mu^2\mathbf{u}^\top\mathbf{M}(G,\varphi)^2\mathbf{u}.$$

Observe that the quantity $\mathbf{u}^\top\mathbf{M}(G,\varphi)\mathbf{u}$ is the *walk generating function* $N_G(\varphi)$ of G that we defined in detail in Appendix A.2. Using the results of Appendix A.2, we obtain:

$$\begin{aligned}\mathbf{u}^\top\mathbf{M}(G,\varphi)^2\mathbf{u} &= \mathbf{u}^\top\left(\sum_{k=0}^{\infty}\varphi^k\mathbf{A}^k\right)^2\mathbf{u}\\ &= \mathbf{u}^\top\left(\sum_{k=0}^{\infty}\sum_{l=0}^{k}\varphi^l\mathbf{A}^l\varphi^{k-l}\mathbf{A}^{k-l}\right)\mathbf{u}\\ &= \sum_{k=0}^{\infty}(k+1)\varphi^k\mathbf{u}^\top\mathbf{A}^k\mathbf{u}\\ &= N_G(\varphi)+\sum_{k=0}^{\infty}k\varphi^k\mathbf{u}^\top\mathbf{A}^k\mathbf{u}.\end{aligned}$$

[62] Note that there exists a relationship between the matrix $\mathbf{M}(G,\varphi)$ with elements $m_{ij}(G,\varphi)$ and the length of the shortest path $\ell_{ij}(G)$ between nodes i and j in the network G, which have been used e.g. in Jackson and Wolinsky [1996]. Namely

$$\ell_{ij}(G) = \lim_{\varphi\to 0}\frac{\partial\ln m_{ij}(G,\varphi)}{\partial\ln\varphi} = \lim_{\varphi\to 0}\frac{\varphi}{m_{ij}(G,\varphi)}\frac{\partial m_{ij}(G,\varphi)}{\partial\varphi}.$$

See also Newman [2010, Chap. 6]. This means that the length of the shortest path between i and j is given by the relative percentage change in the weighted number of walks between nodes i and j in G with respect to a relative percentage change in φ in the limit of $\varphi\to 0$.

Alternatively, we can write

$$\sum_{k=0}^{\infty}(k+1)\varphi^k \mathbf{u}^\top \mathbf{A}^k \mathbf{u} = \sum_{k=0}^{\infty}(k+1)N_k\varphi^k = \frac{d}{d\varphi}(\varphi N_G(\varphi)),$$

so that

$$\mathbf{u}^\top \mathbf{M}(G,\varphi)^2\mathbf{u} = \frac{d}{d\varphi}(\varphi N_G(\varphi)) = N_G(\varphi) + \varphi\frac{d}{d\varphi}N_G(\varphi).$$

Using Rayleigh's inequality, one can show that [Van Mieghem, 2011, p. 51]

$$\frac{d}{d\varphi}(\varphi N_G(\varphi)) \geq \frac{1}{\lambda_1}\frac{d}{d\varphi}(N_G(\varphi)).$$

From this we can obtain a lower bound on welfare given by

$$W(G) \geq \mu^2 \frac{1}{\lambda_1}\frac{d}{d\varphi}(N_G(\varphi)).$$

Further, using the fact that

$$\mathbf{u}^\top \mathbf{A}^k \mathbf{u} = \sum_{i=1}^{n}(\mathbf{u}^\top \mathbf{v}_i)^2 \lambda_i^k,$$

$$N_G(\varphi) = \sum_{i=1}^{n}\frac{(\mathbf{v}_i^\top \mathbf{u})^2}{1-\lambda_i\varphi},$$

we can write

$$\begin{aligned}
\mathbf{u}^\top \mathbf{M}(G,\varphi)^2\mathbf{u} &= \sum_{i=1}^{n}\frac{(\mathbf{v}_i^\top \mathbf{u})^2}{1-\lambda_i\varphi} + \sum_{i=1}^{n}(\mathbf{u}^\top \mathbf{v}_i)^2\sum_{k=0}^{\infty}k\varphi^k\lambda_i^k \\
&= \sum_{i=1}^{n}\frac{(\mathbf{v}_i^\top \mathbf{u})^2}{1-\lambda_i\varphi} + \sum_{i=1}^{n}\frac{(\mathbf{u}^\top \mathbf{v}_i)^2\varphi\lambda_i}{(1-\varphi\lambda_i)^2} \\
&= \sum_{i=1}^{n}\frac{(\mathbf{u}^\top \mathbf{v}_i)^2}{1-\varphi\lambda_i}\left(1+\frac{\varphi\lambda_i}{1-\varphi\lambda_i}\right) \\
&= \sum_{i=1}^{n}\frac{(\mathbf{u}^\top \mathbf{v}_i)^2}{(1-\varphi\lambda_i)^2}.
\end{aligned}$$

From the above it follows that welfare can also be written as

$$W(G) = \mu^2\frac{d}{d\varphi}(\varphi N_G(\varphi)) = \mu^2\sum_{i=1}^{n}\frac{(\mathbf{u}^\top \mathbf{v}_i)^2}{(1-\varphi\lambda_i)^2}.$$

This expression shows that gross welfare is highest in the graph where λ_1 approaches $1/\varphi$. Since, in the k-regular graph G_k it holds that $N_G(\varphi) = \frac{n}{1-k\varphi}$ and $\frac{d}{d\varphi}(\varphi N_G(\varphi)) = N_G(\varphi) + \varphi\frac{d}{d\varphi} = N_G(\varphi) = \frac{n}{1-k\varphi} + \frac{nk\varphi}{(1-k\varphi)^2} = \frac{n}{1-k\varphi}\left(1+\frac{k\varphi}{1-k\varphi}\right) = \frac{n}{(1-k\varphi)^2}$, which gives us a lower bound on welfare in the efficient graph $\frac{n}{(1-\frac{2m}{n}\varphi)^2} \leq W(G^*)$, where we have used the fact that the number of links in a k-regular graph is given by $m = \frac{nk}{2}$.

In order to derive an upper bound, observe that

$$\mathbf{u}^\top \mathbf{M}(G,\varphi)^2\mathbf{u} = \sum_{i=1}^{n}\frac{(\mathbf{u}^\top \mathbf{v}_i)^2}{(1-\varphi\lambda_i)^2},$$

and we can write welfare as follows

$$
\begin{aligned}
W(G) &= \mu^2 \sum_{i=1}^{n} \frac{(\mathbf{u}^\top \mathbf{v}_i)^2}{(1-\varphi\lambda_i)^2} \\
&\leq \mu^2 \frac{\sum_{i=1}^{n} (\mathbf{u}^\top \mathbf{v}_i)^2}{(1-\varphi\lambda_1)^2} \\
&\leq \mu^2 \frac{n}{(1-\varphi\lambda_1)^2},
\end{aligned}
$$

where we have used the fact that $N_G(0) = \sum_{i=1}^{n} (\mathbf{u}^\top \mathbf{v}_i)^2 = n$ so that $(\mathbf{u}^\top \mathbf{v}_1)^2 < n$. Moreover, the largest eigenvalue in a graph G with m links and n nodes is bounded from above by $\lambda_1 \leq \sqrt{\frac{2m(n-1)}{n}} \leq n-1$.[63] This gives us an upper bound on welfare according to

$$
W(G^*) \leq \mu^2 \frac{n}{\left(1-\varphi\sqrt{2m(n-1)/n}\right)^2},
$$

which completes part (ii) of the proposition. Part (iii) follows immediately, if the number of links m can be chosen freely, because the largest eigenvalue λ_1 is upper bounded by the largest eigenvalue of the complete graph K_n, which is the $(n-1)$-regular graph. In this case, upper and lower bounds coincide, and the efficient graph is therefore complete, that is $K_n = \text{argmax}_{G \in \mathcal{G}(n)} W(G)$.

Finally, a similar calculation as in part (ii) shows that

$$
\boldsymbol{\mu}^\top \mathbf{M} \boldsymbol{\mu} = \sum_{i=1}^{n} \frac{(\boldsymbol{\mu}^\top \mathbf{v}_i)^2}{1-\varphi\lambda_i},
$$

and similarly

$$
\boldsymbol{\mu}^\top \mathbf{M}^2 \boldsymbol{\mu} = \sum_{i=1}^{n} \frac{(\boldsymbol{\mu}^\top \mathbf{v}_i)^2}{(1-\varphi\lambda_i)^2},
$$

so that welfare can be written as

$$
W(G) = \mu^2 \boldsymbol{\mu}^\top \mathbf{M}^2 \boldsymbol{\mu} = \mu^2 \sum_{i=1}^{n} \frac{(\boldsymbol{\mu}^\top \mathbf{v}_i)^2}{(1-\varphi\lambda_i)^2},
$$

which completes part (i) of the proposition. □

Proof of Proposition 3. In the case of imperfectly substitutable goods, welfare can be written as

$$
W(G) = \mathbf{q}^\top \mathbf{q} + \frac{\rho}{2} \mathbf{q}^\top \mathbf{B} \mathbf{q}.
$$

Further, denoting by $\widetilde{\mathbf{M}} = (\mathbf{I}_n + \rho\mathbf{B} - \varphi\mathbf{A})^{-1}$ we can write equilibrium output as $\mathbf{q} = \widetilde{\mathbf{M}}\boldsymbol{\mu}$, and welfare can be written as

$$
W(G) = \boldsymbol{\mu}^\top \widetilde{\mathbf{M}}^2 \boldsymbol{\mu} + \frac{\rho}{2} \boldsymbol{\mu}^\top \widetilde{\mathbf{M}} \mathbf{B} \widetilde{\mathbf{M}} \boldsymbol{\mu}.
$$

Observe that $\widetilde{\mathbf{M}} = (\mathbf{I}_n - \varphi\mathbf{C})^{-1}$, where we have denoted by $\mathbf{C} = \mathbf{A} - \frac{\rho}{\varphi}\mathbf{B}$, so that we can write $\widetilde{\mathbf{M}} = \sum_{k=0}^{\infty} \varphi^k \mathbf{C}^k$. Let $\{\nu_i\}_{i=1}^{n}$ be the eigenvalues of $\mathbf{C}$ and $\mathbf{v_i}$ the associated eigenvectors. Further, let $\Lambda = \text{diag}\{\nu_1, \ldots, \nu_n\}$ and $\mathbf{S}$ the matrix whose columns are the eigenvectors $\mathbf{v}_i$. Then we have that $\mathbf{C} = \mathbf{S}\Lambda\mathbf{S}^\top$, and we can write $\widetilde{\mathbf{M}} = \sum_{k=0}^{\infty} \varphi^k \mathbf{S}\Lambda^k\mathbf{S}^\top$. From this one can show that

$$
\boldsymbol{\mu}^\top \widetilde{\mathbf{M}} \boldsymbol{\mu} = \sum_{i=1}^{n} \frac{(\boldsymbol{\mu}^\top \mathbf{v}_i)^2}{1-\varphi\nu_i},
$$

[63] If we assume that G is connected then we can also use the bound $\lambda_1 \leq \sqrt{2m-n+1} \leq n-1$.

and similarly

$$\boldsymbol{\mu}^\top \widetilde{\mathbf{M}}^2 \boldsymbol{\mu} = \sum_{i=1}^n \frac{(\boldsymbol{\mu}^\top \mathbf{v}_i)^2}{(1-\varphi\nu_i)^2}.$$

Moreover, we have that

$$\begin{aligned}
\boldsymbol{\mu}^\top \widetilde{\mathbf{M}}\mathbf{B}\widetilde{\mathbf{M}}\boldsymbol{\mu} &= \boldsymbol{\mu}^\top \widetilde{\mathbf{M}} \sum_{m=1}^M \left(\mathbf{u}_m \mathbf{u}_m^\top - \mathbf{D}_m\right) \widetilde{\mathbf{M}}\boldsymbol{\mu} \\
&= \sum_{m=1}^M (\boldsymbol{\mu}^\top \widetilde{\mathbf{M}}\mathbf{u}_m)(\mathbf{u}_m^\top \widetilde{\mathbf{M}}\boldsymbol{\mu}) - \boldsymbol{\mu}^\top \widetilde{\mathbf{M}}\boldsymbol{\mu} \\
&= \sum_{m=1}^M \left(\sum_{i=1}^n \frac{(\boldsymbol{\mu}^\top \mathbf{v}_i)(\mathbf{u}_m^\top \mathbf{v}_i)}{1-\varphi\nu_i} \right)^2 - \sum_{i=1}^n \frac{(\boldsymbol{\mu}^\top \mathbf{v}_i)^2}{(1-\varphi\nu_i)^2} \\
&= \left(\sum_{i=1}^n \frac{(\boldsymbol{\mu}^\top \mathbf{v}_i)(\mathbf{v}_i \mathbf{B} \mathbf{v}_i)}{1-\varphi\nu_i} \right)^2 - \sum_{i=1}^n \frac{(\boldsymbol{\mu}^\top \mathbf{v}_i)^2}{(1-\varphi\nu_i)^2}.
\end{aligned}$$

It then follows that welfare can be written as

$$W(G) = \frac{2-\rho}{2} \sum_{i=1}^n \frac{\boldsymbol{\mu}^\top \mathbf{v}_i}{1-\varphi\nu_i} \left(\frac{\boldsymbol{\mu}^\top \mathbf{v}_i}{1-\varphi\nu_i} \left(1 + \frac{\rho}{2-\rho} \mathbf{v}_i^\top \mathbf{B} \mathbf{v}_i \right) + \frac{\rho}{2-\rho} \sum_{j\neq i}^n \frac{(\boldsymbol{\mu}^\top \mathbf{v}_i)(\mathbf{v}_i^\top \mathbf{B} \mathbf{v}_j)}{1-\varphi\nu_i} \right).$$

□

Proof of Proposition 4. We start with the proof of part (i) of the proposition. Assuming that $\mu_i = \mu$ for all $i = 1, \ldots, n$, we have that

$$\mathbf{q} = \frac{\mu}{1+\rho(\mathbf{u}^\top \mathbf{M}(G,\phi)\mathbf{u} - 1)} \mathbf{M}(G,\phi)\mathbf{u},$$

with $\mathbf{M}(G,\phi) \equiv (\mathbf{I}_n - \phi\mathbf{A})^{-1}$, and we can write

$$W(G) = \frac{2-\rho}{2} \frac{\mu^2}{(1+\rho(\mathbf{u}^\top \mathbf{M}(G,\phi)\mathbf{u}-1))^2} \left(\mathbf{u}^\top \mathbf{M}(G,\phi)^2 \mathbf{u} + \frac{\rho}{2-\rho} (\mathbf{u}^\top \mathbf{M}(G,\phi)\mathbf{u})^2 \right).$$

Using the fact that $\mathbf{u}^\top \mathbf{M}(G,\phi)\mathbf{u} = N_G(\phi)$ and $\mathbf{u}^\top \mathbf{M}(G,\phi)^2\mathbf{u} = \frac{d}{d\phi}(\phi N_G(\phi))$, we then can write welfare in terms of the walk generating function $N_G(\phi)$ as

$$W(G) = \frac{2-\rho}{2} \frac{\mu^2}{(1+\rho(N_G(\phi)-1))^2} \left(\frac{d}{d\phi}(\phi N_G(\phi)) + \frac{\rho}{2-\rho} N_G(\phi)^2 \right).$$

Next, observe that

$$N_G(\phi) = N_0 + N_1\phi + N_2\phi^2 + O(\phi^3),$$

and consequently

$$\frac{d}{d\phi}(\phi N_G(\phi)) = N_0 + 2N_1\phi + 3N_2\phi^2 + O(\phi^3).$$

Inserting into welfare gives

$$\begin{aligned}
W(G) &= \frac{\mu^2(-\rho + N_0\rho + 2)}{2N_0(1-\rho+N_0\rho)^2} - \frac{N_1\mu^2\rho(-\rho+N_0\rho+2)}{N_0(1-\rho+N_0\rho)^3}\phi \\
&+ \frac{1}{2}\mu^2(2-\rho) \left(\frac{-N_1^2 + N_0N_2}{N_0^3(1-\rho+N_0\rho)^2} + \left(\frac{3N_1^2\rho^2}{(1-\rho+N_0\rho)^4} - \frac{2N_2\rho}{(1-\rho+N_0\rho)^3} \right) \left(\frac{1}{N_0} + \frac{\rho}{2-\rho} \right) \right) \phi^2 \\
&+ O(\phi)^3.
\end{aligned}$$

Using the fact that

$$\begin{aligned}
N_0 &= n, \\
N_1 &= 2m = n\bar{d}, \\
N_2 &= \mathbf{d}^\top \mathbf{d} = n(\bar{d}^2 + \sigma_d^2),
\end{aligned}$$

we get

$$\begin{aligned}
W(G) = {} & \frac{\mu^2(2-\rho+n\rho)}{2n(1-\rho+n\rho)^2} - \frac{2\left(m\mu^2\rho(2-\rho+n\rho)\right)}{n(1-\rho+n\rho)^3}\phi \\
& + \frac{1}{2}\mu^2(2-\rho)\left(\frac{-4m^2+2mn+n^2\sigma^2}{n^3(1-\rho+n\rho)^2} + \left(\frac{12m^2\rho^2}{(1-\rho+n\rho)^4}\right.\right. \\
& \left.\left. - \frac{4m\rho}{(1-\rho+n\rho)^3} - \frac{2n\rho\sigma^2}{(1-\rho+n\rho)^3}\right)\left(\frac{1}{n} + \frac{\rho}{2-\rho}\right)\right)\phi^2 \\
& + O(\phi)^3.
\end{aligned}$$

Taking the derivative with respect to σ_d^2 yields

$$\frac{\partial W(G)}{\partial \sigma_d^2} = \frac{\mu^2\phi^2\left(\left(\left(1+n-2n^2\right)\rho - 1\right)\rho + 2 - 2(1+n)\rho\right)}{2n(1+(n-1)\rho)^3} + O(\phi)^3,$$

and in the limit of large n we obtain

$$\lim_{n\to\infty} n^2 \frac{\partial W(G)}{\partial \sigma_d^2} = -\frac{\mu^2\rho\phi^2}{\rho^2} + O(\phi)^3,$$

which is negative, indicating that welfare is decreasing in the degree variance σ_d^2 for large n up to the second order of ϕ.

Let us now deal with part (ii) of the proposition. Up to the third order in ϕ we have that

$$N_G(\phi) = N_0 + N_1\phi + N_2\phi^2 + N_3\phi^3 + O(\phi^4),$$

and consequently

$$\frac{d}{d\phi}\left(\phi N_G(\phi)\right) = N_0 + 2N_1\phi + 3N_2\phi^2 + 4N_2\phi^3 + O(\phi^4).$$

Using the fact that

$$\begin{aligned}
N_0 &= n, \\
N_1 &= 2m = n\bar{d}, \\
N_2 &= \mathbf{d}^\top \mathbf{d} = n(\bar{d}^2 + \sigma_d^2),
\end{aligned}$$

and inserting into welfare gives

$$\begin{aligned}
W(G) = {} & \frac{\mu^2(2-\rho+n\rho)}{2n(1-\rho+n\rho)^2} - \frac{2\left(m\mu^2\rho(2-\rho+n\rho)\right)}{n(1-\rho+n\rho)^3}\phi \\
& + \frac{1}{2}\mu^2(2-\rho)\left(\frac{-4m^2+2mn+n^2\sigma^2}{n^3(1-\rho+n\rho)^2} + \left(\frac{12m^2\rho^2}{(1-\rho+n\rho)^4}\right.\right. \\
& \left.\left. - \frac{4m\rho}{(1-\rho+n\rho)^3} - \frac{2n\rho\sigma^2}{(1-\rho+n\rho)^3}\right)\left(\frac{1}{n} + \frac{\rho}{2-\rho}\right)\right)\phi^2 \\
& + \frac{1}{2}\mu^2(2-\rho)\left(-\frac{4m\rho\left(-4m^2+2mn+n^2\sigma^2\right)}{n^3(1-\rho+n\rho)^3} + \frac{2\left(8m^3-8m^2n+n^2N_3-4mn^2\sigma^2\right)}{n^4(1-\rho+n\rho)^2}\right. \\
& \left. + \left(-\frac{32m^3\rho^3}{(1-\rho+n\rho)^5} + \frac{24m^2\rho^2}{(1-\rho+n\rho)^4} - \frac{2N_3\rho}{(1-\rho+n\rho)^3} + \frac{12mn\rho^2\sigma^2}{(1-\rho+n\rho)^4}\right)\left(\frac{1}{n} + \frac{\rho}{2-\rho}\right)\right)\phi^3 \\
& + O(\phi)^4.
\end{aligned}$$

Taking the derivative with respect to N_3 yields

$$\frac{\partial W(G)}{\partial N_3} = -\frac{\mu^2\phi^3\left(\left(1+\left(n^2-1\right)\rho\right)\rho - 2(1-\rho)\right)}{n^2(1+(n-1)\rho)^3} + O(\phi)^4,$$

and in the limit of large n we obtain

$$\lim_{n\to\infty} n^3\frac{\partial W(G)}{\partial N_3} = -\frac{\mu^2\rho\phi^3}{\rho^2} + O(\phi)^4.$$

It follows that welfare in the limit of large n is decreasing in N_3. We further have that [Van Mieghem, 2011, p. 183]

$$\begin{aligned}
N_3 &= \sum_{i=1}^n d_i^3 - \frac{1}{2}\sum_{i=1}^n\sum_{j=1}^n a_{ij}(d_i-d_j)^2 \\
&= \frac{N_2^2}{N_1} + \rho_d(G)\underbrace{\left(\sum_{i=1}^n d_i^3 - \frac{N_2^2}{N_1}\right)}_{\geq 0},
\end{aligned}$$

where $\rho_d(G)$ is the degree assortativity coefficient of G [Newman, 2003]. Since welfare is decreasing in N_3, and N_3 is increasing in the assortativity $\rho_d(G)$, we have that welfare in the limit of large n is decreasing in $\rho_d(G)$.

We now deal with parts (iii) and (iv) of the proposition. We first provide a lower bound on welfare for the efficient graph by considering the complete graph K_n. Welfare can be written as

$$W(G) = \frac{2-\rho}{2}\frac{\mu^2}{\rho^2}\frac{\mathbf{u}^\top\mathbf{M}^2\mathbf{u} + \frac{\rho}{2-\rho}(\mathbf{u}^\top\mathbf{M}\mathbf{u})^2}{\left(\frac{1-\rho}{\rho} + \mathbf{u}^\top\mathbf{M}\mathbf{u}\right)^2}.$$

For the k-regular graph G_k we have that

$$\begin{aligned}
\mathbf{u}^\top\mathbf{M}\mathbf{u} &= \frac{n}{1-(k-1)\phi}, \\
\mathbf{u}^\top\mathbf{M}^2\mathbf{u} &= \frac{n}{(1-(k-1)\phi)^2},
\end{aligned}$$

and welfare is given by

$$W(G_k) = \frac{\mu^2 n((n-1)\rho+2)}{2(\rho(k\phi+n-1)-k\phi+1)^2}.$$

As $k = 2m/n$ this is

$$W(G_k) = \frac{\mu^2 n^3((n-1)\rho+2)}{2(2m(\rho-1)\phi+(n-1)n\rho+n)^2}.$$

Together with the definition of the average degree $\bar{d} = \frac{2m}{n}$ this gives us the lower bound on welfare in part (iii) of the proposition. In particular, for the complete graph K_n we get

$$\begin{aligned}
\mathbf{u}^\top\mathbf{M}\mathbf{u} &= \frac{n}{1-(n-1)\phi}, \\
\mathbf{u}^\top\mathbf{M}^2\mathbf{u} &= \frac{n}{(1-(n-1)\phi)^2},
\end{aligned}$$

so that we obtain for welfare in the complete graph

$$W(K_n) = \frac{\mu^2 n((n-1)\rho+2)}{2((n-1)\rho(\phi+1)-n\phi+\phi+1)^2}.$$

Using the fact that $\phi = \frac{\varphi}{1-\rho}$ we can write this as follows

$$W(K_n) = \frac{\mu^2 n((n-1)\rho + 2)}{2((n-1)\rho - n\varphi + \varphi + 1)^2}.$$

This gives us the lower bound on welfare in part (iv) of the proposition. To obtain an upper bound, note that welfare can be written as

$$W(G) = \frac{2-\rho}{2}\frac{\mu^2}{\rho^2}\frac{\frac{\mathbf{u}^\top \mathbf{M}^2\mathbf{u}}{(\mathbf{u}^\top \mathbf{M}\mathbf{u})^2} + \frac{\rho}{2-\rho}}{\frac{\left(\frac{1-\rho}{\rho} + \mathbf{u}^\top \mathbf{M}\mathbf{u}\right)^2}{(\mathbf{u}^\top \mathbf{M}\mathbf{u})^2}}.$$

Next, observe that

$$\frac{\left(\frac{1-\rho}{\rho} + \mathbf{u}^\top \mathbf{M}\mathbf{u}\right)^2}{(\mathbf{u}^\top \mathbf{M}\mathbf{u})^2} = 1 + \frac{2(1-\rho)}{\rho}\frac{1}{\mathbf{u}^\top \mathbf{M}\mathbf{u}} + \frac{1-\rho}{\rho}\frac{1}{(\mathbf{u}^\top \mathbf{M}\mathbf{u})^2} \geq 1,$$

for $0 < \rho \leq 1$. This implies that

$$W(G) \leq \frac{2-\rho}{2}\frac{\mu^2}{\rho^2}\left(\frac{\rho}{2-\rho} + \frac{\mathbf{u}^\top \mathbf{M}^2\mathbf{u}}{(\mathbf{u}^\top \mathbf{M}\mathbf{u})^2}\right). \tag{54}$$

Moreover, we have that

$$\begin{aligned}
\frac{\mathbf{u}^\top \mathbf{M}^2\mathbf{u}}{(\mathbf{u}^\top \mathbf{M}\mathbf{u})^2} &= \frac{\frac{d}{d\phi}\left(\phi N_G(\phi)\right)}{N_G(\phi)^2} \\
&= \frac{\sum_{i=1}^n \frac{(\mathbf{u}^\top \mathbf{v}_i)^2}{(1-\phi\lambda_i)^2}}{\left(\sum_{i=1}^n \frac{(\mathbf{u}^\top \mathbf{v}_i)^2}{1-\phi\lambda_i}\right)^2} \\
&\leq \frac{\frac{1}{1-\phi\lambda_1}\sum_{i=1}^n \frac{(\mathbf{u}^\top \mathbf{v}_i)^2}{1-\phi\lambda_i}}{\left(\sum_{i=1}^n \frac{(\mathbf{u}^\top \mathbf{v}_i)^2}{1-\phi\lambda_i}\right)^2} \\
&= \frac{1}{(1-\phi\lambda_1)N_G(\phi)} \\
&\leq \frac{1}{n(1-\phi\lambda_1)},
\end{aligned}$$

where we have used the fact that $N_G(\phi) \geq N_G(0) = N_0 = n$. Hence, we obtain an upper bound on welfare in the efficient graph G^* for large n given by

$$W(G^*) \leq \frac{2-\rho}{2}\frac{\mu^2}{\rho^2}\left(\frac{1}{n(1-\phi\lambda_1)} + \frac{\rho}{2-\rho}\right).$$

Using the upper bound $\lambda_1 \leq \sqrt{\frac{2m(n-1)}{n}} \leq n-1$ we get

$$W(G^*) \leq \frac{2-\rho}{2}\frac{\mu^2}{\rho^2}\left(\frac{\rho}{2-\rho} + \frac{1}{n\left(1-\phi\sqrt{\frac{2m(n-1)}{n}}\right)}\right).$$

This allows us to state an upper and lower bound (from the explicit solution for welfare in the complete graph K_n) for the efficient graph $G^* = \text{argmax}_{G\in\mathcal{H}(n,m)} W(G)$. As the largest eigenvalue is bounded from above by $\lambda_1 \leq n-1$ we obtain

$$W(G^*) \leq \frac{2-\rho}{2}\frac{\mu^2}{\rho^2}\left(\frac{\rho}{2-\rho} + \frac{1}{n(1-\phi(n-1))}\right), \tag{55}$$

which gives an upper bound for the efficient graph $G^* = \text{argmax}_{G\in\mathcal{G}(n)} W(G)$. Using the fact that $\phi = \frac{\varphi}{1-\rho}$

we can write this as

$$W(G^*) \leq \frac{\mu^2((n-1)n\rho\varphi + (\rho-1)((n-1)\rho+2))}{2n\rho^2((n-1)\varphi+\rho-1)}.$$

In the following let us denote by $\overline{W}$ the upper bound on welfare in Equation (55). Then, for part (iv) of the proposition, note that in the limit of large n the upper bound $\overline{W}$ converges to

$$\lim_{n\to\infty} \overline{W} = \frac{\mu^2}{2\rho},$$

while for the complete graph K_n we get

$$\lim_{n\to\infty} W(K_n) = \frac{\mu^2\rho}{2(\rho-(1-\rho)\phi)^2}.$$

Hence, we have that

$$\lim_{n\to\infty} \frac{W(K_n)}{\overline{W}} = \frac{\rho^2}{(\rho-(1-\rho)\phi)^2}.$$

Thus, we get

$$\lim_{\varphi\to 0}\lim_{n\to\infty} \frac{W(K_n)}{\overline{W}} = 1,$$

which proves part (iv) of the proposition. □

Proof of Proposition 5. (i) In the case of independent markets ($\rho = 0$), welfare can be written as

$$W(G) = \mathbf{q}^\top\mathbf{q} = \boldsymbol{\mu}^\top\mathbf{M}(G,\phi)^2\boldsymbol{\mu},$$

where $\mathbf{M}(G,\phi) = (\mathbf{I}_n - \phi\mathbf{A})^{-1}$. Using the fact that (see the proof of Proposition 2):

$$\boldsymbol{\mu}^\top\mathbf{M}(G,\phi)^2\boldsymbol{\mu} = \frac{d}{d\phi}\left(\phi N_G(\phi)\right) = \boldsymbol{\mu}^\top\frac{d}{d\phi}\left(\phi\mathbf{M}(G,\phi)\right)\boldsymbol{\mu},$$

welfare can be written as

$$W(G) = \boldsymbol{\mu}^\top\frac{d}{d\phi}\left(\phi\mathbf{M}(G,\phi)\right)\boldsymbol{\mu},$$

we can write the change in welfare due to the exit of firm i as follows

$$\begin{aligned} W(G) - W(G^{-i},\phi) &= \left(\boldsymbol{\mu}^\top\mathbf{M}(G,\phi)^2\boldsymbol{\mu} - \boldsymbol{\mu}^\top\mathbf{M}(G^{-i},\phi)^2\boldsymbol{\mu}\right) \\ &= \left(\frac{d}{d\phi}\phi\left(\boldsymbol{\mu}^\top\mathbf{M}(G,\phi)\boldsymbol{\mu} - \boldsymbol{\mu}^\top\mathbf{M}(G^{-i},\phi)\boldsymbol{\mu}\right)\right). \end{aligned}$$

Denoting by

$$\Delta_i(G,\phi) \equiv \boldsymbol{\mu}^\top\mathbf{M}(G,\phi)\boldsymbol{\mu} - \boldsymbol{\mu}^\top\mathbf{M}(G^{-i},\phi)\boldsymbol{\mu},$$

we can write the change in welfare as follows

$$W(G) - W(G^{-i},\phi) = \left(\frac{d}{d\phi}\phi\left(\Delta_i(G,\phi)\right)\right).$$

We next turn to the analysis of the quantity $\Delta_i(G,\phi)$. We first make the following observation (see Lemma 1 in Ballester et al. [2006])

$$m_{jk}(G^{-i},\phi) = m_{jk}(G,\phi) - \frac{m_{ij}(G,\phi)m_{ik}(G,\phi)}{m_{ii}(G,\phi)}.$$

We then can write

$$
\begin{aligned}
\boldsymbol{\mu}^\top \mathbf{M}(G^{-i},\phi)\boldsymbol{\mu} &= \sum_{j,k} \mu_j m_{jk}(G^{-i},\phi)\mu_k \\
&= \boldsymbol{\mu}^\top \mathbf{M}(G,\phi)\boldsymbol{\mu} - \frac{\sum_{j,k} \mu_j m_{ij}(G,\phi) m_{ik}(G,\phi)\mu_k}{m_{ii}(G,\phi)} \\
&= \boldsymbol{\mu}^\top \mathbf{M}(G,\phi)\boldsymbol{\mu} - \frac{b_{\boldsymbol{\mu},i}(G,\phi)^2}{m_{ii}(G,\phi)},
\end{aligned}
$$

and we obtain

$$
\Delta_i(G,\phi) = \frac{b_{\boldsymbol{\mu},i}(G,\phi)^2}{m_{ii}(G,\phi)}.
$$

We then define the centrality measure

$$
c_i \equiv \frac{1}{2}\frac{d}{d\phi}\left(\phi \Delta_i(G,\phi)\right) = \frac{1}{2}\frac{d}{d\phi}\left(\frac{\phi b_{\boldsymbol{\mu},i}(G,\phi)^2}{m_{ii}(G,\phi)}\right). \tag{56}
$$

The centrality c_i corresponds to the welfare loss incurred from to the removal of firm i. Observe that

$$
m_{ii}(G,\phi) = N_G(\phi,i) \equiv \sum_{k=0}^{\infty} a_{ii}^{[k]}\phi^k,
$$

is the generating function of the number of closed walks that start and terminate at node i. It can be written as [Van Mieghem, 2011]

$$
N_G(\phi,i) = \sum_{k=1}^{n} \frac{(\mathbf{v}_k\mathbf{v}_k^\top)_{ii}}{1-\lambda_k\phi} = -\frac{c_{\mathbf{A}^{-i}}\left(\frac{1}{\phi}\right)}{\phi c_{\mathbf{A}}\left(\frac{1}{\phi}\right)},
$$

where $c_{\mathbf{A}}(\phi) \equiv \det(\mathbf{A} - \phi\mathbf{I}_n)$ is the characteristic polynomial of the matrix $\mathbf{A}$, and $\mathbf{A}^{-i}$ is the matrix obtained from $\mathbf{A}$ by removing the i-th column and row. We can then write the centrality index as follows.

$$
c_i = \frac{1}{2}\frac{d}{d\phi}\left(\frac{\phi b_{\boldsymbol{\mu},i}(G,\phi)^2}{N_G(\phi,i)}\right). \tag{57}
$$

This shows that the centrality index $\mathbf{c}(G,\phi)$ is determined by the Bonacich centrality $b_i(G,\phi)$ of firm i and the walk generating function $N_G(\phi,i)$. Further note that[64]

$$
\begin{aligned}
\frac{d}{d\phi}\left(\frac{\phi b_{\boldsymbol{\mu},i}(G,\phi)^2}{N_G(\phi,i)}\right) &= \frac{b_{\boldsymbol{\mu},i}(G,\phi)^2}{N_G(\phi,i)} + \frac{\phi b_{\boldsymbol{\mu},i}(G,\phi)}{N_G(\phi,i)}\Big[2(\mathbf{M}(G,\phi)\mathbf{A}\mathbf{b}_{\boldsymbol{\mu}}(G,\phi))_i \\
&\quad - \frac{b_{\boldsymbol{\mu},i}(G,\phi)}{N_G(\phi,i)}(\mathbf{M}(G,\phi)\mathbf{A}\mathbf{M}(G,\phi))_{ii}\Big] \\
&= \frac{b_{\boldsymbol{\mu},i}(G,\phi)}{N_G(\phi,i)}\left[2(\mathbf{M}(G,\phi)\mathbf{b}_{\boldsymbol{\mu}}(G,\phi))_i - \frac{b_{\boldsymbol{\mu},i}(G,\phi)}{N_G(\phi,i)}(\mathbf{M}(G,\phi)^2)_{ii}\right] \\
&= \boldsymbol{\mu}^\top \mathbf{M}(G,\phi)\boldsymbol{\mu} - \boldsymbol{\mu}^\top \mathbf{M}(G^{-i},\phi)\boldsymbol{\mu} \\
&= \Delta_i(G,\phi).
\end{aligned} \tag{58}
$$

We then can write the centrality as follows

$$
c_i = \frac{b_{\boldsymbol{\mu},i}(G,\phi)}{N_G(\phi,i)}\left[(\mathbf{M}(G,\phi)\mathbf{b}_{\boldsymbol{\mu}}(G,\phi))_i - \frac{1}{2}\frac{b_{\boldsymbol{\mu},i}(G,\phi)}{N_G(\phi,i)}(\mathbf{M}(G,\phi)^2)_{ii}\right].
$$

[64] We have used the fact that $\frac{d\mathbf{M}(G,\phi)}{d\phi} = \mathbf{M}(G,\phi)\mathbf{A}\mathbf{M}(G,\phi)$, which follows from $\frac{d\mathbf{X}^{-1}}{d\phi} = -\mathbf{X}^{-1}\frac{d\mathbf{X}}{d\phi}\mathbf{X}^{-1}$ for any invertible matrix $\mathbf{X}$.

(ii) Welfare for $\rho > 0$ is given by

$$W(G) = \frac{1}{2}\sum_{i=1}^{n} q_i^2 + \frac{\rho}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} b_{ij} q_i q_j + \sum_{i=1}^{n} \pi_i.$$

Using the fact that $\pi_i = \frac{1}{2}q_i^2$, we obtain for welfare

$$W(G) = \sum_{i=1}^{n} q_i^2 + \frac{\rho}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} b_{ij} q_i q_j.$$

In vector-matrix notation this can be written as

$$W(G) = \mathbf{q}^\top(G)\mathbf{q}(G) + \frac{\rho}{2}\mathbf{q}^\top(G)\mathbf{B}\mathbf{q}(G).$$

We denote by G^{-i} the network obtained from G by removing firm i. Applying Lemma 1 in Ballester et al. [2006] to the weighted symmetric matrix $\mathbf{M}(G,\rho,\varphi)$, we get

$$m_{jk}(G^{-i},\rho,\varphi) = m_{jk}(G,\rho,\varphi) - \frac{m_{ij}(G,\rho,\varphi)m_{ik}(G,\rho,\varphi)}{m_{ii}(G,\rho,\varphi)}$$

For equilibrium output we have that $\mathbf{q} = \mathbf{b}_{\boldsymbol{\mu}} = \mathbf{M}\boldsymbol{\mu}$, so that we obtain for the output of firm j after the removal of firm i

$$\begin{aligned}
q_j(G^{-i}) &= \sum_{l=1}^{n-1} m_{jl}(G^{-i})\mu_l^{-i}(G) \\
&= \sum_{l=1}^{n-1}\left(m_{jl}(G) - \frac{m_{ij}(G)m_{il}(G)}{m_{ii}(G)}\right)\mu_l^{-i}(G) \\
&= \sum_{l=1}^{n} m_{jl}(G)\mu_l - \frac{m_{ij}(G)}{m_{ii}(G)}\sum_{l=1}^{n} m_{il}(G)\mu_l \\
&= (\mathbf{M}(G)\boldsymbol{\mu})_j - \frac{m_{ij}(G)b_{\boldsymbol{\mu},i}(G)}{m_{ii}(G)}.
\end{aligned}$$

Moreover, we have that

$$\sum_{j=1}^{n-1} q_j(G^{-i}) = \mathbf{u}^\top\mathbf{M}(G)\boldsymbol{\mu}(G) - \frac{b_{\mathbf{u},i}(G)b_{\boldsymbol{\mu},i}(G)}{m_{ii}(G)}$$

We then have that

$$\begin{aligned}
\mathbf{q}(G^{-i})^\top\mathbf{B}^{-i}\mathbf{q}(G^{-i}) &= \sum_{j,k=1}^{n} q_j(G^{-i})b_{jk}^{-i}q_k(G^{-i}) \\
&= \sum_{j,k\neq i}^{n} b_{jk}\left(q_j(G) - \frac{m_{ij}(G)q_i(G)}{m_{ii}(G)}\right)\left(q_k(G) - \frac{m_{ik}(G)q_i(G)}{m_{ii}(G)}\right) \\
&= \sum_{j,k\neq i}^{n} b_{jk}q_j(G)q_k(G) - \frac{b_{\boldsymbol{\mu},i}(G)}{m_{ii}(G)}\sum_{j,k\neq i}^{n} b_{jk}(q_j(G)m_{ik}(G) + q_k(G)m_{ij}(G)) \\
&+ \frac{b_{\boldsymbol{\mu},i}(G)^2}{m_{ii}(G)^2}\sum_{j,k\neq i}^{n} b_{jk}m_{ij}(G)m_{ik}(G).
\end{aligned}$$

This can be simplified to

$$\mathbf{q}(G^{-i})^\top\mathbf{B}^{-i}\mathbf{q}(G^{-i}) = \mathbf{q}(G)^\top\mathbf{B}\mathbf{q}(G) - \frac{q_i(G)}{m_{ii}(G)}\left(2(\mathbf{M}(G)\mathbf{B}\mathbf{q}(G))_i - \frac{q_i(G)}{m_{ii}(G)}(\mathbf{M}(G)\mathbf{B}\mathbf{M}(G))_{ii}\right)$$

In the special case of $\mathbf{B} = \mathbf{I}_n$ this is

$$\mathbf{q}(G^{-i})^\top \mathbf{q}(G^{-i}) = \mathbf{q}(G)^\top \mathbf{q}(G) - \frac{q_i(G)}{m_{ii}(G)} \left(2(\mathbf{M}(G)\mathbf{q}(G))_i - \frac{q_i(G)}{m_{ii}(G)} (\mathbf{M}(G)^2)_{ii} \right).$$

We then obtain

$$\begin{aligned}
W(G) - W(G^{-i}) &= (\mathbf{q}^\top(G)\mathbf{q}(G) - \mathbf{q}^\top(G^{-i})\mathbf{q}(G^{-i})) + \frac{\rho}{2}(\mathbf{q}^\top(G)\mathbf{B}\mathbf{q}(G) - \mathbf{q}^\top(G^{-i})\mathbf{B}\mathbf{q}(G^{-i})) \\
&= \frac{q_i(G)}{m_{ii}(G)} \left((\mathbf{M}(G)(2\mathbf{I}_n + \rho\mathbf{B})\mathbf{q}(G))_i - \frac{1}{2}\frac{q_i(G)}{m_{ii}(G)} (\mathbf{M}(G)(2\mathbf{I}_n + \rho\mathbf{B})\mathbf{M}(G))_{ii} \right) \\
&= \frac{b_{\boldsymbol{\mu},i}(G)}{m_{ii}(G)} \left((\mathbf{M}(G)(2\mathbf{I}_n + \rho\mathbf{B})\mathbf{b}_{\boldsymbol{\mu}}(G))_i - \frac{1}{2}\frac{b_{\boldsymbol{\mu},i}(G)}{m_{ii}(G)} (\mathbf{M}(G)(2\mathbf{I}_n + \rho\mathbf{B})\mathbf{M}(G))_{ii} \right).
\end{aligned}$$

□

Proof of Proposition 6. (i) The FOC of profits in Equation (17) with respect to effort is

$$\frac{\partial \pi_i}{\partial e_i} = q_i - e_i + s = 0,$$

so that equilibrium effort is

$$e_i = q_i + s.$$

The FOC with respect to output is given by

$$\frac{\partial \pi_i}{\partial q_i} = (\bar{\alpha} - \bar{c}_i) - 2q_i - \rho \sum_{j \neq i} b_{ij} q_j + e_i + \varphi \sum_{j=1}^n a_{ij} e_j = 0.$$

Inserting equilibrium efforts, rearranging terms and introducing the reduced from variable $\mu_i \equiv \bar{\alpha} - \bar{c}_i$ gives

$$q_i = \mu_i - \rho \sum_{j \neq i} b_{ij} q_j + \varphi \sum_{j=1}^n a_{ij} q_j + s + \varphi d_i s.$$

where $d_i = \sum_{j=1}^n a_{ij}$ is the degree (or total number of links) of firm i. In vector-matrix notation this is

$$(\mathbf{I}_n + \rho\mathbf{B} - \varphi\mathbf{A})\mathbf{q} = \boldsymbol{\mu} + s\mathbf{u} + \varphi s\mathbf{A}\mathbf{u}.$$

We then can write equilibrium quantities as follows

$$\mathbf{q} = \bar{\mathbf{q}} + s\mathbf{r},$$

where we have denoted by

$$\begin{aligned}
\bar{\mathbf{q}} &\equiv (\mathbf{I}_n + \rho\mathbf{B} - \varphi\mathbf{A})^{-1}\boldsymbol{\mu} = \mathbf{M}\boldsymbol{\mu} \\
\mathbf{r} &\equiv \varphi(\mathbf{I}_n + \rho\mathbf{B} - \varphi\mathbf{A})^{-1} \left(\frac{1}{\varphi}\mathbf{I}_n + \mathbf{A} \right) \mathbf{u} = \mathbf{M}\mathbf{u} + \varphi\mathbf{M}\mathbf{d},
\end{aligned}$$

where $\mathbf{M} = (\mathbf{I}_n + \rho\mathbf{B} - \varphi\mathbf{A})^{-1}$. The vector $\bar{\mathbf{q}}$ gives equilibrium quantities in the absence of the subsidy and is derived in Section 3. The vector $\mathbf{r}$ has elements r_i for $i = 1, \ldots, n$. Furthermore, equilibrium profits are given by

$$\pi_i = \frac{1}{2}q_i^2 + \frac{1}{2}s^2,$$

(ii) Net social welfare is given by

$$\overline{W}(G, s) = W(G, s) - s\sum_{i=1}^n e_i = \sum_{i=1}^n \left(q_i^2 + \pi_i - s e_i \right) = \sum_{i=1}^n q_i^2 - s\sum_{i=1}^n q_i - \frac{n}{2}s^2.$$

Using the fact that $q_i = \bar{q}_i + sr_i$, where

$$\begin{aligned}\bar{\mathbf{q}} &= (\mathbf{I}_n - \varphi\mathbf{A})^{-1}\boldsymbol{\mu} = \mathbf{M}\boldsymbol{\mu} \\ \mathbf{r} &= \varphi(\mathbf{I}_n - \varphi\mathbf{A})^{-1}\left(\frac{1}{\varphi}\mathbf{I}_n + \mathbf{A}\right)\mathbf{u} = \boldsymbol{\mu} + \varphi\mathbf{d},\end{aligned}$$

we can write net welfare as follows

$$\overline{W}(G,s) = \sum_{i=1}^{n}(\bar{q}_i + r_i s)^2 - \sum_{i=1}^{n}(\bar{q}_i + r_i s) - \frac{n}{2}s^2.$$

The FOC of net welfare $\overline{W}(G,s)$ is given by

$$\frac{\partial\overline{W}(G,s)}{\partial s} = 2\sum_{i=1}^{n}\bar{q}_i\,(2r_i - 1) + s\sum_{i=1}^{n}\left(2r_i^2 - 2r_i - 1\right) = 0,$$

from which we obtain the optimal subsidy level

$$s^* = \frac{\sum_{i=1}^{n}\bar{q}_i\,(1 - 2r_i)}{\sum_{i=1}^{n}\left(r_i\,(2r_i - 2) - 1\right)},$$

where the equilibrium quantities are given by Equation (18). For the second-order derivative we obtain

$$\frac{\partial^2\overline{W}(G,s)}{\partial s^2} = -\sum_{i=1}^{n}\left(-2r_i^2 + 2r_i + 1\right),$$

and we have an interior solution if the condition $\sum_{i=1}^{n}\left(-2r_i^2 + 2r_i + 1\right) \geq 0$ is satisfied.

(iii) Net welfare can be written as

$$\begin{aligned}\overline{W}(G,s) &= \frac{1}{2}\sum_{i=1}^{n}q_i^2 + \frac{\rho}{2}\sum_{i=1}^{n}\sum_{j\neq i}^{n}b_{ij}q_iq_j + \sum_{i=1}^{n}\pi_i - s\sum_{i=1}^{n}e_i \\ &= \sum_{i=1}^{n}q_i^2 + \frac{n}{2}s^2 + \frac{\rho}{2}\sum_{i=1}^{n}\sum_{j\neq i}^{n}b_{ij}q_iq_j - \sum_{i=1}^{n}(q_i + s)s.\end{aligned}$$

Using the fact that $q_i = \bar{q}_i + sr_i$, where

$$\begin{aligned}\bar{\mathbf{q}} &\equiv (\mathbf{I}_n + \rho\mathbf{B} - \varphi\mathbf{A})^{-1}\boldsymbol{\mu} \\ \mathbf{r} &\equiv \varphi(\mathbf{I}_n + \rho\mathbf{B} - \varphi\mathbf{A})^{-1}\left(\frac{1}{\varphi}\mathbf{I}_n + \mathbf{A}\right)\mathbf{u},\end{aligned}$$

we can write net welfare as follows

$$\overline{W}(G,s) = \sum_{i=1}^{n}(\bar{q}_i + r_i s)^2 - ns^2 + \frac{\rho}{2}\sum_{i=1}^{n}\sum_{j\neq i}^{n}b_{ij}(\bar{q}_i + sr_i)(\bar{q}_j + sr_j) - \sum_{i=1}^{n}(\bar{q}_i s + r_i s^2).$$

The FOC of net welfare $\overline{W}(G,s)$ is given by

$$\frac{\partial\overline{W}(G,s)}{\partial s} = \sum_{i=1}^{n}\left(2\bar{q}_ir_i - \bar{q}_i + \frac{\rho}{2}b_{ij}(\bar{q}_ir_j + \bar{q}_jr_i)\right) + s\sum_{i=1}^{n}\left(2r_i^2 - 2r_i - 1 + \rho\sum_{j=1}^{n}b_{ij}r_ir_j\right) = 0,$$

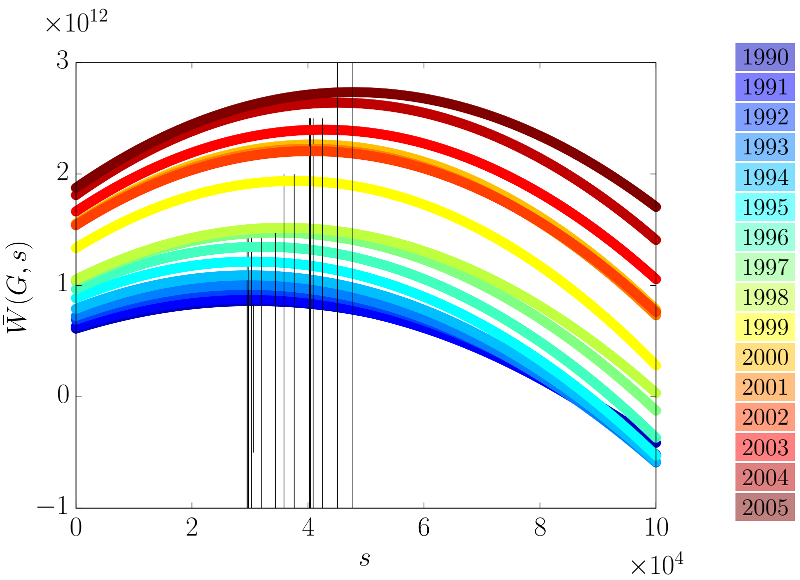


Figure F.1: The concave welfare function $\bar{W}(G,s)$ for different years and different subsidy levels s. The location of the maximum s^* for each year is indicated with a vertical line.

from which we obtain the optimal subsidy level

$$s^* = \frac{\sum_{i=1}^n \left(\bar{q}_i(2r_i+1) + \frac{\rho}{2}\sum_{j=1}^n b_{ij}(\bar{q}_i r_j + \bar{q}_j r_i)\right)}{\sum_{i=1}^n \left(1 + r_i\left(2 - 2r_i - \rho\sum_{j=1}^n b_{ij} r_j\right)\right)},$$

where the equilibrium quantities are given by Equation (18). The second-order derivative is given by

$$\frac{\partial^2 \overline{W}(G,s)}{\partial s^2} = -\sum_{i=1}^n \left(-2r_i^2 + 2r_i + 1 - \rho\sum_{j=1}^n b_{ij} r_i r_j.\right).$$

Hence, the solution is interior if $\sum_{i=1}^n \left(-2r_i^2 + 2r_i + 1 - \rho\sum_{j=1}^n b_{ij} r_i r_j\right) \geq 0$.

The concave welfare function $\bar{W}(G,s)$ for different years and different subsidy levels s is shown in Figure F.1. The location of the maximum s^* for each year is indicated with a vertical line.

□

Proof of Proposition 7. (i) The FOC of profits from Equation (20) with respect to effort is

$$\frac{\partial \pi_i}{\partial e_i} = q_i - e_i + s_i = 0,$$

so that equilibrium effort is

$$e_i = q_i + s_i.$$

The FOC with respect to output is given by

$$\frac{\partial \pi_i}{\partial q_i} = (\bar{\alpha} - \bar{c}_i) - 2q_i - \rho\sum_{j\neq i} b_{ij} q_j + e_i + \varphi\sum_{j=1}^n a_{ij} e_j = 0.$$

Inserting equilibrium efforts and rearranging terms gives

$$q_i = \mu_i - \rho\sum_{j\neq i} b_{ij} q_j + \varphi\sum_{j=1}^n a_{ij} q_j + s_i + \varphi\sum_{j=1}^n a_{ij} s_j.$$

In vector-matrix notation this is

$$(\mathbf{I}_n + \rho\mathbf{B} - \varphi\mathbf{A})\mathbf{q} = \boldsymbol{\mu} + \mathbf{s} + \varphi\mathbf{A}\mathbf{s}.$$

We then can write equilibrium quantities as follows

$$\mathbf{q} = \bar{\mathbf{q}} + \mathbf{R}\mathbf{s},$$

where we have denoted by

$$\begin{aligned}\bar{\mathbf{q}} &\equiv (\mathbf{I}_n + \rho\mathbf{B} - \varphi\mathbf{A})^{-1}\boldsymbol{\mu} = \mathbf{M}\boldsymbol{\mu} \\ \mathbf{R} &\equiv \varphi(\mathbf{I}_n + \rho\mathbf{B} - \varphi\mathbf{A})^{-1}\left(\frac{1}{\varphi}\mathbf{I}_n + \mathbf{A}\right) = \mathbf{M} + \varphi\mathbf{M}\mathbf{A},\end{aligned}$$

with $\mathbf{M} = (\mathbf{I}_n + \rho\mathbf{B} - \varphi\mathbf{A})^{-1}$. The matrix $\mathbf{R}$ has elements r_{ij} for $1 \leq i, j \leq n$. Furthermore, one can show that equilibrium profits are given by

$$\pi_i = \frac{1}{2}q_i^2 + \frac{1}{2}s_i^2.$$

(ii) Net welfare can be written as follows

$$\begin{aligned}\overline{W}(G, \mathbf{s}) &= \sum_{i=1}^{n}\left(\frac{q_i^2}{2} + \pi_i - s_i e_i\right) \\ &= \sum_{i=1}^{n} q_i^2 - \sum_{i=1}^{n} q_i s_i - \frac{1}{2}\sum_{i=1}^{n} s_i^2.\end{aligned}$$

Using the fact that $q_i = \bar{q}_i + r_{ij}s_j$, with

$$\begin{aligned}\bar{\mathbf{q}} &= (\mathbf{I}_n - \varphi\mathbf{A})^{-1}\boldsymbol{\mu} = \mathbf{M}\boldsymbol{\mu} \\ \mathbf{R} &= \varphi(\mathbf{I}_n - \varphi\mathbf{A})^{-1}\left(\frac{1}{\varphi}\mathbf{I}_n + \mathbf{A}\right) = \boldsymbol{\mu} + \varphi\mathbf{d},\end{aligned}$$

where $\mathbf{R}$ is symmetric, i.e. $r_{ij} = r_{ji}$, we can write net welfare as follows

$$\overline{W}(G, \mathbf{s}) = \sum_{i=1}^{n}\bar{q}_i^2 - \sum_{i=1}^{n}\bar{q}_i s_i - \frac{1}{2}\sum_{i=1}^{n}s_i^2 + \sum_{i=1}^{n}\left(\sum_{j=1}^{n}r_{ij}s_j\right)\left(2\bar{q}_i + \sum_{j=1}^{n}r_{ij}s_j - s_i\right). \tag{59}$$

Equation (59) can be written in vector-matrix notation as follows

$$\overline{W}(G, \mathbf{s}) = \bar{\mathbf{q}}^\top\bar{\mathbf{q}} - \frac{1}{2}\mathbf{s}^\top\left(\mathbf{I}_n - 2\mathbf{R}(\mathbf{R} - \mathbf{I}_n)\right)\mathbf{s} - \bar{\mathbf{q}}^\top(\mathbf{I}_n - 2\mathbf{R})\mathbf{s}.$$

Denoting by $\mathbf{Q} \equiv \mathbf{I}_n - 2\mathbf{R}(\mathbf{R} - \mathbf{I}_n)$ and $\mathbf{c}^\top \equiv \bar{\mathbf{q}}^\top(\mathbf{I}_n - 2\mathbf{R})$ we find that maximizing net welfare is equivalent to solving the following quadratic programming problem: $\min_{\mathbf{s}\in\mathbb{R}_+^n}\left\{\mathbf{c}^\top\mathbf{s} + \frac{1}{2}\mathbf{s}^\top\mathbf{Q}\mathbf{s}\right\}$ [cf. Boyd and Vandenberghe, 2004]. The FOC for net welfare $\overline{W}(G, \mathbf{s})$ of Equation (59) yields the following system of linear equations

$$\begin{aligned}\frac{\partial\overline{W}(G, \mathbf{s})}{\partial s_i} &= -\bar{q}_i - s_i + \sum_{k=1}^{n}r_{ki}\left(2\bar{q}_k + \sum_{j=1}^{n}r_{kj}s_j - s_k\right) \\ &\quad + \sum_{k=1}^{n}\left(\sum_{j=1}^{n}r_{kj}s_j\right)\left(\frac{1}{2}r_{ki} - \delta_{ki}\right) = 0.\end{aligned}$$

In vector-matrix notation this can be written as

$$(\mathbf{I}_n + 2\mathbf{R} - 2\mathbf{R}^2)\mathbf{s} = (2\mathbf{R} - \mathbf{I}_n)\bar{\mathbf{q}}.$$

When the conditions for invertibility are satisfied, it then follows that the optimal subsidy levels can be written as

$$\mathbf{s}^* = (\mathbf{I}_n + 2\mathbf{R} - 2\mathbf{R}^2)^{-1}(2\mathbf{R} - \mathbf{I}_n)\bar{\mathbf{q}}, \tag{60}$$

with $\bar{\mathbf{q}} = (\mathbf{I}_n - \varphi\mathbf{A})^{-1}\boldsymbol{\mu} = \mathbf{b}_{\boldsymbol{\mu}}$. The second-order derivative is given by

$$\frac{\partial^2 \overline{W}(G,\mathbf{s})}{\partial s_i \partial s_j} = -\delta_{ij} - 2r_{ij} + 2\sum_{k=1}^{n} r_{ki} r_{kj}.$$

In vector-matrix notation this can be written as

$$\frac{\partial^2 \overline{W}(G,\mathbf{s})}{\partial \mathbf{s} \partial \mathbf{s}^\top} = -\mathbf{I}_n + 2\mathbf{R} - 2\mathbf{R}^2.$$

Hence, we obtain a global maximum for the concave quadratic optimization problem if the matrix $\mathbf{I}_n + 2\mathbf{R} - 2\mathbf{R}^2 = \mathbf{I}_n - 2\mathbf{R}^2 + 2\mathbf{R}$ is positive definite, which means that it is also invertible and its inverse is also positive definite.

(iii) In the case of interdependent markets, when goods are substitutable, net welfare can be written as

$$\begin{aligned}
\overline{W}(G,\mathbf{s}) &= \frac{1}{2}\left(\sum_{i=1}^{n} q_i^2 + \rho\sum_{i=1}^{n}\sum_{j\neq i}^{n} b_{ij} q_i q_j\right) + \sum_{i=1}^{n} \pi_i - \sum_{i=1}^{n} s_i e_i \\
&= \sum_{i=1}^{n} q_i^2 - \sum_{i=1}^{n} q_i s_i - \frac{1}{2}\sum_{i=1}^{n} s_i^2 + \frac{\rho}{2}\sum_{i=1}^{n}\sum_{j\neq i}^{n} b_{ij} q_i q_j.
\end{aligned}$$

Using the fact that $q_i = \bar{q}_i + r_{ij}s_j$, with

$$\begin{aligned}
\bar{\mathbf{q}} &\equiv (\mathbf{I}_n + \rho\mathbf{B} - \varphi\mathbf{A})^{-1}\boldsymbol{\mu} \\
\mathbf{R} &\equiv \varphi(\mathbf{I}_n + \rho\mathbf{B} - \varphi\mathbf{A})^{-1}\left(\frac{1}{\varphi}\mathbf{I}_n + \mathbf{A}\right),
\end{aligned}$$

where $\mathbf{R}$ is in general not symmetric, unless $\mathbf{AB} = \mathbf{BA}$,[65] we can write net welfare as follows

$$\begin{aligned}
\overline{W}(G,\mathbf{s}) = &\sum_{i=1}^{n}\left(\bar{q}_i + \sum_{j=1}^{n} r_{ij}s_j\right)^2 - \sum_{i=1}^{n}\left(\bar{q}_i + \sum_{j=1}^{n} r_{ij}s_j\right)s_i - \frac{1}{2}\sum_{i=1}^{n} s_i^2 \\
&+ \frac{\rho}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} b_{ij}\left(\bar{q}_i + \sum_{k=1}^{n} r_{ik}s_k\right)\left(\bar{q}_j + \sum_{l=1}^{n} r_{jl}s_l\right).
\end{aligned} \tag{61}$$

In vector-matrix notation we can write Equation (61) as follows

$$\overline{W}(G,\mathbf{s}) = \bar{\mathbf{q}}^\top\bar{\mathbf{q}} + \frac{\rho}{2}\bar{\mathbf{q}}^\top\mathbf{B}\bar{\mathbf{q}} - \frac{1}{2}\mathbf{s}^\top\left(\mathbf{I}_n + 2\mathbf{R}^\top(\mathbf{I}_n - \mathbf{R} - \frac{\rho}{2}\mathbf{BR})\right)\mathbf{s} - \bar{\mathbf{q}}^\top\left(\mathbf{I}_n - 2\mathbf{R} - \rho\mathbf{BR}\right)\mathbf{s}.$$

If we denote by $\mathbf{Q} \equiv \mathbf{I}_n + 2\mathbf{R}^\top(\mathbf{I}_n - \mathbf{R} - \frac{\rho}{2}\mathbf{BR})$ and $\mathbf{c}^\top \equiv \bar{\mathbf{q}}^\top\left(\mathbf{I}_n - 2\mathbf{R} - \rho\mathbf{BR}\right)$ we find that maximizing net welfare is equivalent to solving the following quadratic programming problem: $\min_{\mathbf{s}\in\mathbb{R}_+^n}\left\{\mathbf{c}^\top\mathbf{s} + \frac{1}{2}\mathbf{s}^\top\mathbf{Q}\mathbf{s}\right\}$ [cf. Boyd and Vandenberghe, 2004], where we can replace $\mathbf{Q}$ with the symmetric matrix $\frac{1}{2}\left(\mathbf{Q}^\top + \mathbf{Q}\right)$

[65] While the inverse of a symmetric matrix is symmetric, the product of symmetric matrices is not necessarily symmetric.

to obtain an equivalent problem. The FOC from Equation (61) is given by

$$\frac{\partial \overline{W}(G,\mathbf{s})}{\partial s_i} = -\bar{q}_i + 2\sum_{k=1}^{n} r_{ki}\bar{q}_k - s_i - 2\sum_{k=1}^{n} r_{ki}s_k + 2\sum_{k=1}^{n}\sum_{j=1}^{n} r_{ki}r_{kj}s_j$$
$$+\frac{\rho}{2}\sum_{l=1}^{n}\sum_{j=1}^{n} b_{li}\bar{q}_l r_{ji} + \frac{\rho}{2}\sum_{l=1}^{n}\sum_{j=1}^{n} b_{lj}\bar{q}_j r_{li} + \frac{\rho}{2}\sum_{l=1}^{n}\sum_{j=1}^{n} b_{lj}\left(r_{li}\sum_{k=1}^{n} r_{jk}s_k + r_{ji}\sum_{k=1}^{n} r_{lk}s_k\right) = 0.$$

In vector-matrix notation this can be written as follows

$$\frac{\partial \overline{W}(G,\mathbf{s})}{\partial \mathbf{s}} = -\bar{\mathbf{q}} + \bar{\mathbf{q}}^\top(2\mathbf{R}+\rho\mathbf{B}\mathbf{R}) - \mathbf{s} - 2\mathbf{R}^\top\left(\mathbf{I}_n - \frac{1}{2}(2\mathbf{I}_n+\rho\mathbf{B})\mathbf{R}\right)\mathbf{s}.$$

When the matrix $\mathbf{I}_n - 2\mathbf{R}^\top\left(\frac{1}{2}(2\mathbf{I}_n+\rho\mathbf{B}\right)\mathbf{R} - \mathbf{I}_n)$ is invertible, the optimal subsidy levels can then be written as

$$\mathbf{s}^* = \left(\mathbf{I}_n - 2\mathbf{R}^\top\left(\frac{1}{2}(2\mathbf{I}_n+\rho\mathbf{B}\right)\mathbf{R} - \mathbf{I}_n)\right)^{-1}\left(\mathbf{R}^\top(2\mathbf{I}_n+\rho\mathbf{B}) - \mathbf{I}_n\right)\bar{\mathbf{q}}, \tag{62}$$

where the equilibrium quantities in the absence of the subsidy are given by

$$\bar{\mathbf{q}} = (\mathbf{I}_n + \rho\mathbf{B} - \varphi\mathbf{A})^{-1}\boldsymbol{\mu}.$$

The second-order derivative is given by

$$\frac{\partial^2 \overline{W}(G,\mathbf{s})}{\partial \mathbf{s}\partial \mathbf{s}^\top} = -\mathbf{I}_n + 2\mathbf{R}^\top(\mathbf{I}_n - \frac{1}{2}(2\mathbf{I}_n+\rho\mathbf{B})\mathbf{R}).$$

Hence, we obtain a global maximum for the concave quadratic optimization problem if the matrix $\mathbf{I}_n + 2\mathbf{R}^\top(\mathbf{I}_n - \frac{1}{2}(2\mathbf{I}_n+\rho\mathbf{B})\mathbf{R})$ is positive definite. Note that if this matrix is positive definite then it is also invertible and its inverse is also positive definite.

Note that when the condition for concavity is not satisfied then we can sill use Equations (60) or (62), respectively, as a candidate for a welfare improving subsidy program. However, there might exist other subsidy programs that yield even higher welfare gains.

□

Proof of Proposition 8. In terms of the walk generating function welfare can be written as

$$W(G) = \frac{2-\rho}{2}\frac{\mu^2}{\rho^2}\frac{N_G(\phi)^2}{\left(\frac{1-\rho}{\rho}+N_G(\phi)\right)^2}\left(\frac{\rho}{2-\rho} + \frac{\frac{d}{d\phi}(\phi N_G(\phi))}{N_G(\phi)^2}\right)$$
$$= \frac{2-\rho}{2}\frac{\mu^2}{\rho^2}\frac{(1-\phi\lambda_1)^2 N_G(\phi)^2}{\left(\frac{1-\rho}{\rho}(1-\phi\lambda_1)+(1-\phi\lambda_1)N_G(\phi)\right)^2}\left(\frac{\rho}{2-\rho} + \frac{(1-\phi\lambda_1)^2\frac{d}{d\phi}(\phi N_G(\phi))}{(1-\phi\lambda_1)^2 N_G(\phi)^2}\right).$$

Then the following limits for the walk generating function hold

$$\lim_{\phi\uparrow 1/\lambda_1}(1-\phi\lambda_1)N_G(\phi) = \lim_{\phi\uparrow 1/\lambda_1}(1-\phi\lambda_1)\sum_{i=1}^{n}\frac{(\mathbf{u}^\top\mathbf{v}_i)^2}{1-\phi\lambda_i} = (\mathbf{u}^\top\mathbf{v}_1)^2$$
$$\lim_{\phi\uparrow 1/\lambda_1}(1-\phi\lambda_1)^2 N_G(\phi)^2 = \lim_{\phi\uparrow 1/\lambda_1}(1-\phi\lambda_1)^2\left(\sum_{i=1}^{n}\frac{(\mathbf{u}^\top\mathbf{v}_i)^4}{(1-\phi\lambda_i)^2} + \sum_{i=1}^{n}\sum_{j\neq i}^{n}\frac{(\mathbf{u}^\top\mathbf{v}_i)^2(\mathbf{u}^\top\mathbf{v}_j)^2}{(1-\phi\lambda_i)(1-\phi\lambda_j)}\right) = (\mathbf{u}^\top\mathbf{v}_1)^4$$
$$\lim_{\phi\uparrow 1/\lambda_1}(1-\phi\lambda_1)^2\frac{d}{d\phi}(\phi N_G(\phi)) = \lim_{\phi\uparrow 1/\lambda_1}(1-\phi\lambda_1)^2\sum_{i=1}^{n}\frac{(\mathbf{u}^\top\mathbf{v}_i)^2}{(1-\phi\lambda_i)^2} = (\mathbf{u}^\top\mathbf{v}_1)^2.$$

In the limit of $\phi \uparrow \lambda_1$ we then get for welfare

$$\begin{aligned}
\lim_{\phi \uparrow 1/\lambda_1} W(G) &= \frac{2-\rho}{2}\frac{\mu^2}{\rho^2}\frac{(\mathbf{u}^\top \mathbf{v}_1)^4}{(\mathbf{u}^\top \mathbf{v}_1)^4}\left(\frac{\rho}{2-\rho} + \frac{(\mathbf{u}^\top \mathbf{v}_1)^2}{(\mathbf{u}^\top \mathbf{v}_1)^4}\right) \\
&= \frac{2-\rho}{2}\frac{\mu^2}{\rho^2}\left(\frac{\rho}{2-\rho} + \frac{1}{(\mathbf{u}^\top \mathbf{v}_1)^2}\right).
\end{aligned}$$

This expression is increasing with decreasing values of $(\mathbf{u}^\top \mathbf{v}_1)^2 = \|\mathbf{v}_1\|_1^2$. We thus find that the welfare maximizing graph G^* is the one that minimizes the ℓ^1-norm $\|\mathbf{v}_1\|_1$ of the principal eigenvector $\mathbf{v}_1$ associated with the largest eigenvalue λ_1. $\square$