# Strategic Network Decisions and Knowledge Spillovers: Evidence from R&D Collaborations of the U.S. firms

Kippeum Lee*

Pennsylvania State University

Click here for the latest version

October, 2023

**Abstract**

This paper examines the effect of private R&D investment on productivity, considering R&D collaborations and knowledge spillovers. While existing literature emphasizes R&D's direct effects on innovation and cost reduction, it often neglects R&D's role in shaping collaborative networks. Investing in R&D enhances a firm's learning capacity and augments the firm's appeal as a collaboration partner. Consequently, the effect of R&D is underestimated without accounting for its role in fostering collaborations. To bridge the gap, I develop a dynamic model of a firm that internalizes its decision on whom to collaborate with and following spillovers. This framework allows R&D to improve productivity and affect the collaboration network, with varying propensities for collaborations across firms. Using the data on firm-to-firm R&D collaborations among U.S. firms from 1980 to 2001, I find the long-term effect of R&D is 16% underestimated if we ignore its subsidiary role in expanding the collaboration network.

Keywords: R&D collaborations, R&D investments, knowledge spillovers.
JEL Classification: D24, D25, D85, L14, L24

*Ph.D. candidate, Department of Economics, Pennsylvania State University, Kern Building, University Park, PA 16802, kul598@psu.edu. I am immensely grateful to Joris Pinkse, Sung Jae Jun, Karl Schurter for their advice and support on this project. All remaining errors are my own.

# 1 Introduction

This paper revisits one of the classic questions in the industrial organization field: How does Research and Development (R&D) affect productivity and welfare? This reexamination is motivated by a notable trend of increasing R&D collaborations among firms to enhance productivity, such as Microsoft and Intel's collaborative work on an operating system (Hagedoorn 2002).[1] Prior studies posit that R&D investment is one of the key drivers to forming collaborations because of a firm's forward-looking behavior (Cohen and Levinthal 1989; Hernán, Marín, and Siotis 2003; Badillo and Moreno 2016). Intuitively, in the prospect of future spillovers, large R&D investments make a firm attractive as a collaboration partner and expand a firm's capacity to communicate with productive firms. This raises an empirically important yet unanswered question: Through fostering collaborations, does R&D have an additional effect on productivity and welfare? Disregarding this effect leads to underestimating the R&D's in-house and social impacts and misguides firms and policymakers for R&D subsidy. However, the existing methods for endogenous network formation focus on agents' static actions.

This paper's contribution is to provide a methodology that internalizes a firm's forward-looking behaviors in forming a network, thereby enabling a comprehensive analysis of the R&D's effect in the presence of collaborations and spillovers. In doing so, I first develop an estimable dynamic structural model that incorporates two crucial features—strategic R&D network formation and R&D spillovers. This model allows firms' R&D investments not only to improve productivity but also to boost collaboration networks, and each firm has different propensities for forming R&D collaborations. It introduces an additional channel for R&D to improve productivity via network formation. Next, the model fits the panel of R&D collaborations of U.S. R&D-intensive firms from 1980 to 2001. Lastly, I quantify the effect of R&D through networking and give R&D subsidy policy implications.

The framework in this paper extends a dynamic model of a firm employed by Doraszelski and Jaumandreu (2013) and Peters et al. (2017), in which private R&D augments firm-specific productivity unobserved by the econometricians. In the extended framework, I introduce the following key features: (i) a firm achieves efficiency gain through its private R&D investment and R&D spillovers from collaborative efforts with other firms; and (ii) a firm makes dynamic decisions on whom to collaborate with in the next period based on the expected gains from collaborations. The anticipated benefits

[1] According to PwC's 17th Annual Global CEO Survey conducted in 2014, 20% of 1,344 CEOs in 68 countries identified mergers and acquisitions, R&D collaborations, or strategic alliances as significant strategies. Source: https://www.pwc.com/gx/en/ceo-survey/2014/assets/pwc-17th-annual-global-ceo-survey-jan-2014.pdf

are contingent on several firm-specific characteristics, especially their R&D investments, as well as the decisions made by other firms, which makes networking strategic. This model allows me to disentangle the private R&D's effect on productivity into the *direct* effect and *indirect* effect through endogenous network formation.

However, the strategic game of the network formation model inherently embeds two estimation challenges. First, when firms consider collaborations with all peers, the potential network portfolios grow exponentially, with each collaboration requiring mutual agreement. To circumvent this, I borrow the idea from Aguirregabiria and Ho (2012) and introduce the concept of a *collaboration delegate*–each firm delegates its collaboration decision to these agents. At the cost of the delegate's restricted information on others' private shock to network formation, it simplifies an extensive network formation game of firms into a manageable incomplete information game of delegate pairs deciding on binary collaboration status.

Second, the dimension of state variables escalates because a delegate's information set includes the information of all other firms. To mitigate the high dimensionality problem and streamline the strategic game among delegates, I adopt the *inclusive-values* approach (Hendel and Nevo 2006; Aguirregabiria and Ho 2012) and assume bounded rationality on delegates. Instead of considering all firms' information and choices in the economy, a delegate pair's decision relies only on key aggregated information and some moments of the beliefs on other delegate pairs' decisions. These assumptions are plausible in this setting but simplify the complicated strategic game with a huge dimension of variables, making it feasible to estimate.

The model's estimation unfolds in two stages. First, I estimate the production function, inclusive of unobserved productivity, extending Ackerberg, Caves, and Frazer (2015). I additionally provide an extended method to consider productivity spillovers from collaborations that are not directly observed from the data. Second, the strategic network formation model is estimated using Leung (2015)'s two-step estimator. The beliefs on forming collaborations are estimated under the assumption of the equilibrium symmetric on observables in the first step, and these beliefs then inform the full model's estimation.

The theoretical framework suggested in this paper bridges the gap between the two pieces of literature on R&D and knowledge spillovers by incorporating the decision on networks into a firm's dynamic model. Recently beyond the individual returns to R&D, the social impacts of R&D, particularly the role of spillovers among firms, have obtained significant attention (Bloom, Schankerman, and Van Reenen 2013; Lychagin et al. 2016; Konig, Liu, and Zenou 2019; Iyoha 2020; Zacchia 2020; Malikov and Zhao 2021). Yet, the majority of these studies either assume networks as given or address potential endogeneity via instrumental variables without internalizing the decision-making process. A recent pertinent study, Hsieh, König, and Liu (2022), aligns closely with my work, given its modeling of the joint endogeneity of R&D and R&D networks. The underlying structure, however, is different from my framework in that I allow R&D and spillovers to have a direct bearing on unobserved productivity and delve into the dynamic interplay between R&D and collaborations. This allows me to shed light on the unique dimension of R&D in enhancing network capabilities, which is the first attempt in the literature to the best of my knowledge.

Furthermore, from an econometric perspective, my framework diverges from the existing models with endogenous network formation by considering an agent's dynamic incentive to form networks. While there have been substantial advancements in identifying and estimating network formation models and in models with endogenous network formation, the focus has predominantly been on static models (Johnsson and Moon 2021; de Paula 2020; Auerbach 2022). My paper enriches this literature by providing a micro-foundation for a forward-looking firm's network formation and offering an approach to simplifying complex strategic games under incomplete information.

Then I bring this model to panel data of mostly cross-market or buyer-supplier R&D alliances, as well as annual company reports of U.S. R&D-intensive publicly listed firms from 1980 to 2001. To construct firm-level collaboration networks, I combine two datasets: (1) the official announcements of R&D alliances and (2) co-patents or co-licensing between inventors or scientists from different firms (Hagedoorn 2002; Zacchia 2020).

The empirical results suggest that a 10 percent increase in R&D investment leads to a 6 percent higher likelihood of engaging in any collaboration for firms that haven't previously collaborated and a 3 percent increase in the expected number of collaborations for firms already in R&D collaboration networks. Thus, in the long run, a 10 percent increase in private R&D investment directly improves productivity by 0.31 percent but indirectly increases productivity by 0.05 percent through expanding a collaboration network and increased R&D spillovers. (I.e. for a median firm, by investing $147 million more in R&D, the direct effect on productivity leads to the increase of $165 million in its output, and the indirect effect through the network formation leads to additional $28 million in its output) It suggests that the effect of R&D is underestimated by 16 percent if we disregard R&D's additional role in network formation. These findings indicate while collaborating firms may lean on R&D partners's R&D, there is a strategic incentive to invest in R&D to attract future collaborators.

In addition to the private productivity gains, this study suggests that firms engaged in collaborations yield additional social benefits through not only spillovers but also broadening their networks. When I consider endogenous network formation, the wedge between marginal social and private returns to R&D increased by 22.8 percentage points. This implies market failure in R&D investment, underscoring the importance of targeted R&D subsidies for firms involved in collaborations. The findings further reveal a significant increase in the wedge, particularly among smaller firms with few collaborators. This is because they are more prone to expand their R&D networks by collaborating with larger firms, thus leading to higher marginal social returns. In the conventional approach with exogenous network formation, large firms with many collaborations are naturally favored for targeted R&D subsidies due to the anticipated social benefits from their extensive collaborations. However, this paper also highlights that firms with fewer collaborations should be considered for subsidies, given their propensity to expand their R&D networks by partnering with larger, more productive firms.

I organize the rest of this paper as follows: The next section describes the data and features of the sample of the U.S. R&D collaboration network that I observe. Section 3 presents the empirical framework and discusses the challenges and solutions. In section 4, I provide the estimation strategy of the model. Section 5 reports the empirical results of the model and discusses the effect of R&D and policy implications on R&D subsidies. Section 6 concludes.

## 2 Data: U.S. R&D Collaboration Network

I begin by describing the data employed to illustrate the firm-level R&D collaboration network within the U.S. to highlight features important for my empirical methodology. R&D collaborations refer to partnerships between firms sharing resources, expertise, and costs to conduct R&D activities. For example, Intel and HP collaborated on developing a new microprocessor architecture in the 1990s. While R&D collaborations span from equity-based joint ventures to contractual agreements, the general procedure for R&D collaboration can be summarized as follows. Forming a collaboration with a potential collaborator takes time and money. It includes examining the credibility and suitability of potential partners, setting specific aims, scope, and effective dates, and drafting a legal contract. Once effective, firms exert and share R&D efforts. The duration of collaboration might or might not be set, but firms typically review the ongoing processes annually to decide whether to continue, amend, or terminate.[2]

For the study, I rely on a panel of publicly listed U.S. firms with at least one patent from 1980 to 2001 in the S&P COMPUSTAT database, as initially assembled by Bloom, Schankerman, and Van Reenen (2013). COMPUSTAT gathers companies' financial statements from form 10-K reports submitted to the U.S. Securities and Exchange Commission (SEC), providing comprehensive information on firms' sales, capital stock, R&D investments, expenses, and employees. Specifically, the restriction on firms that have obtained patents is for focusing on R&D-intensive firms and their collaborations. This process yields an unbalanced panel dataset of 715 mostly manufacturing firms, each with at least four observations between 1980 and 2001, and therefore 13,720 firm-year observations.

Two data sources capture R&D collaborations among those firms. The first source is cross-licensing activities between firms obtained from Zacchia (2020). The United States Patents and Trademark Office (USPTO) data from the NBER archive provides detailed information on U.S. patents. Each patent is associated with unique identifiers for individual inventors and the firms to which the patent is assigned. It reveals collaborations through patents jointly filed with multiple firms, enabling the construction of an R&D alliance network based on co-patenting. It captures both official and unofficial collaborations effectively.

To complement the first data source and identify collaborations that did not produce a co-patent, I combined the second data source, the Cooperative Agreements and Technology Indicators (CATI) database. It documents technology agreements between U.S.

[2]The collaboration process can be inferred from the examples of R&D collaboration agreements or contracts from the U.S. Securities and Exchange Commission (SEC).

firms collected from various sources, including newspaper and journal articles, books, specialized journals, and company annual reports (See Hagedoorn (2002) and Schilling (2009))[3]. These agreements encompass alliances that entail technology transfer or joint research initiatives. Combining these two data sources, the analysis involves 19,667 R&D collaborations among the sample of firms.

**Descriptive patterns in collaborations** Out of 715 firms studied, 453 have engaged in at least one collaboration within the analyzed time frame, with a median of 3 collaborations per collaborating firm. But there is an increasing trend in the sample. Figure 1 left graph depicts the yearly count of total and 'connected' firms—those with at least one collaboration. Despite sample attrition in recent years, the number of connected firms escalated between 1981 and 1998.[4] The right graph in Figure 1 demonstrates a similar trend for the total number of collaborations and averaged collaborations per firm. This suggests a potential growth in inter-firm networks.[5]

The collaboration data reveals that 33% of firm-year observations have a single collaboration, while 80% have under 10 as described in Figure 10. Multiple collaborations might indicate different collaborations or a large R&D collaboration consortium. Although exact forms of collaborations cannot be perfectly retrieved from co-patenting data, official announcements from the CATI database indicate that over 90% involve just two firms and less than 1% constitute large consortiums. Hence, this paper assumes one-on-one collaborations without additional group collaboration impact.

In this data set, most collaborating firms are across-market or buyer-supplier pairs. Figure 2 shows the kernel density distributions of the Mahalanobis market and technology similarity measures between collaborators.[6]. The mean market and technological similarities are 0.18 and 0.87, with medians at 0.03 and 0.76, respectively. Specifically, 67% of firm pairs operate in different industries, and the median market similarity is still 0.17 for within-industry pairs (See Appendix A.2 for the summary of the industries collaborating firms operate). This suggests that while most collaborating firms exhibit technological congruity, they generally operate in distinct markets. Collaborations be-

[3]I would like to express my gratitude to Professor Hagedoorn for providing this dataset.

[4]For detailed information and possible explanations of the attrition of the sample, see Bloom, Schankerman, and Van Reenen (2013).

[5]However, this increase might reflect the surge in patent counts over the past two decades, as the collaboration data is principally drawn from co-patent activities, although a consistent rising trend is also observed in officially announced R&D collaborations from the CATI database (Hagedoorn 2002).

[6]For market similarity, I define $n_{iq}$ as the number of a sales agent in product market $q$ for a firm $i$ for the proxy of sales in product market $q$. Then the Mahalanobis market similarity measure is given as $\sum_{q=1}^{\mathscr{Y}}(\frac{n_{i\tau}}{n_i}\frac{n_{jq}}{n_j})n_i n_j$. The measure for technological similarity can be similarly defined using the information on technological fields of patents firms have.

Table 1: Summary statistics

| | No network | # of networks | |
|---|---|---|---|
| | | below median | above median |
| $Y_{it}$: sales (value-added) | 460.3 | 710.7 | 3,752 |
| $L_{it}$: employees (thousand) | 8.954 | 14.39 | 52.69 |
| $K_{it}$: capital stocks | 867.0 | 1,163 | 7,078 |
| $Y_{it}/L_{it}$ | 49.87 | 55.59 | 80.56 |
| $R_{it}$: R&D expenditures | 9.368 | 40.35 | 606.5 |
| Observations | 8,628 | 2,803 | 2,289 |

*This table reports average characteristics of firms in the sample based on the number of networks. All monetary values are in 2009 million USD.

tween market rivals, such as Apple and IBM[7], are rare due to potential business stealing effects unless anticipated collaboration benefits like market expansion, surpass the disadvantages. Therefore, this paper's theoretical model will concentrate on the positive spillovers from collaborations rather than the effects of market rivalry.

Table 1 displays summary statistics for firm accounting data based on the number of networks. It suggests that firms with more collaborations tend to be larger regarding sales, input variables, and R&D expenditures. Especially Figure 3 shows the positive relationship between R&D investments and the number of collaborations. This correlation is confounded with other firm characteristics. However, I conduct the reduced-form regression (See Appendix A.3), controlling for other firm variables, and I find there is still a significant positive effect of R&D on the number of collaborations.

One concern with the data is the scarcity of information on the duration of collaborations.[8] To address this, a standard 3-year duration for R&D partnerships is assumed, following precedent studies (Hanaki, Nakajima, and Ogura 2010; Konig, Liu, and Zenou 2019) and a survey[9]. Co-licensing data assumes that a co-patent between two firms at time $t$ implies a connection lasting from $t-1$ to $t+1$. Official R&D collaborations are presumed to last 3 years from the announcement date unless the exact duration is known.

Another limitation is that I focus on patent-holding U.S. firms that are publicly listed, omitting smaller and non-patent-holding entities. Given the paper's emphasis on examining spillovers among R&D-intensive firms, this narrowed focus is deemed justified, as such firms are usually large and publicly listed.[10]

[7] Apple and IBM were rivals in the personal computer market but collaborated on PowerPC architecture to compete with Intel's processors.

[8] Around 15% of the sample from the CATI database have duration information.

[9] Using a survey of the top managers of 52 companies in the biotechnology industry, Deeds and Hill (1998) found that the average duration of R&D collaborations is 3.47 years

[10] Other potential sample selection issues, including this exclusion criterion, are discussed in detail in

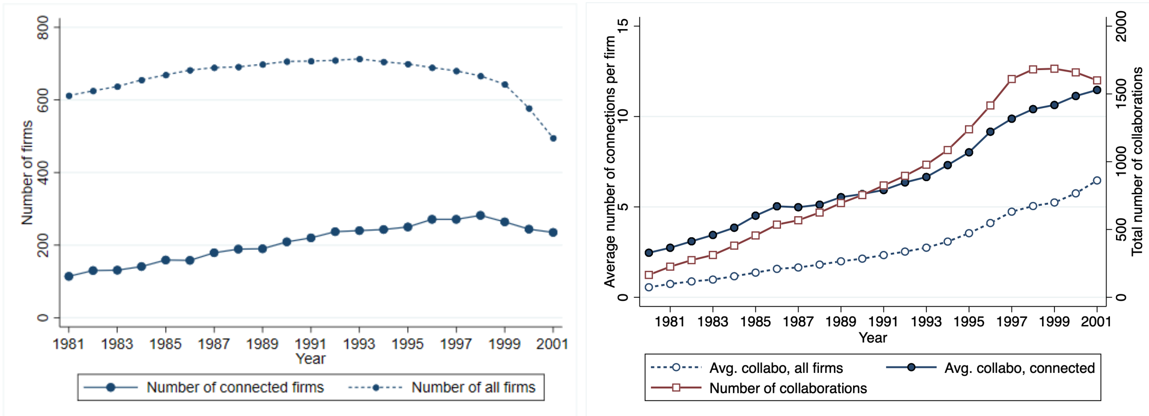



Figure 1: The number of connected firms and average links per firm

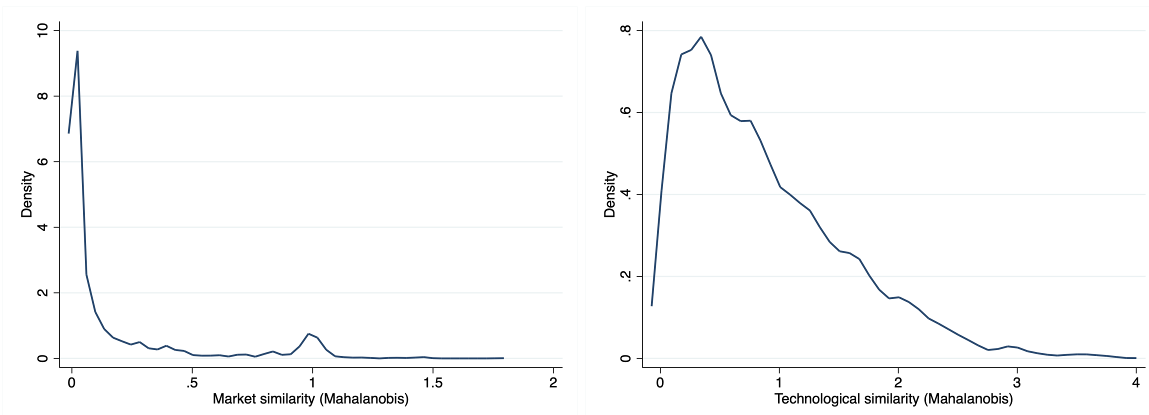



Figure 2: Kernel density distributions of Mahalanobis
market and technological similarity measures

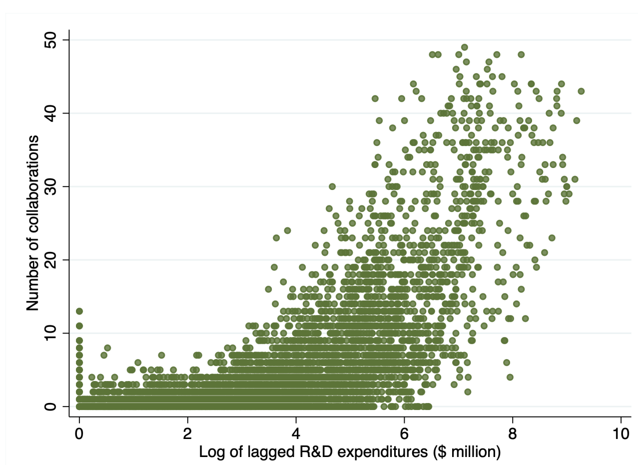



Figure 3: Log of lagged R&D expenditures and the number of collaborations

# 3 Theoretical Framework

This section introduces a dynamic structural model that captures spillovers and a firm's decision-making process regarding R&D collaborations. First, in the following subsection 3.1, I introduce a firm's production function with firm-specific productivity that benefits from both private R&D investment and R&D spillover from collaborating firms. Then to account for the *indirect* effect of private R&D through network formation, the section 3.2 provides a micro-foundation for R&D collaboration decision, offering a comprehensive analysis of the underlying mechanism that drives the decision-making process. In particular, I discuss the potential challenges in estimating the model and propose an approach that is plausible but significantly streamlines the estimation for the network formation model. Lastly, using the suggested theoretical framework, the section 3.3 disentangles the effect of R&D and defines the *direct* effect and *indirect* effects through network formation.

## 3.1 Production and endogenous productivity

**Production function** There are $N$ firms in the economy, $i = 1, ..., N$ that live infinitely many discrete periods, $t = 1, ..., \infty$. At each period $t$, firm $i$ has a Cobb-Douglas production function with Hicks neutral production technology:

$$y_{it} = \alpha_0 + \alpha_k k_{it} + \alpha_l l_{it} + \omega_{it} + \epsilon_{it} \tag{1}$$

where $y_{it}$, $l_{it}$, and $k_{it}$ are value-added(output minus intermediate inputs), labor, and capital in logs, respectively. The firm-specific productivity $\omega_{it}$ captures efficiency in production, such as technological innovation or absorption, changes in process and labor composition, and managerial abilities, unobserved by the econometricians but known to the firm when making decisions. However, a stochastic shock to production, $\epsilon_{it}$, is not known to both the firm and econometrician until the actual short-run profit is realized at the end of the period; thus, it is independent of all variables known to the firm. In the short run, the capital stock, $k_{it}$, is regarded as a fixed factor.

Assuming a firm operates in a monopolistically competitive market and input prices are exogenously given, the profit-maximizing problem of a firm determines all the flexible input variables and output prices. In this context, for a given capital stock, the firm's (expected) short-run profit, denoted by $\pi_{it}$, is characterized by its productivity, i.e., $\pi_{it} =$

Appendix B.4. of Bloom, Schankerman, and Van Reenen (2013).

$\pi(\omega_{it})$.[11]

**Productivity evolution process** The firm-specific productivity $\omega_{it}$ endogenously evolves within the model and follows a controlled first-order Markov process:

$$\omega_{it+1} = \beta_0 + \beta_1 \omega_{it} + \beta_2 r_{it} + \beta_3 S_{it} + u_{it+1} \tag{2}$$

where $u_{it+1}$ is productivity shock *i.i.d.* across time and firms. Private R&D investment flow in logs is denoted as $r_{it}$, and spillover is denoted as $S_{it}$. Spillover $S_{it}$ indicates R&D spillovers, $\Sigma_{j\neq i} g_{ijt} t_{ij} r_{jt}$ where $r_{jt}$ is firm $j$'s R&D investment flow and $t_{ij}$ is time-invariant Jaffe technological similarity measure between firm $i$ and $j$.[12] The binary R&D collaboration status between firm $i$ and $j$ at period $t$ is denoted as $g_{ijt} \in \{0, 1\}$. The network is undirected, i.e., $g_{ijt} = g_{jit}$, and self-tie is excluded, i.e., $g_{iit} = 0$. Thus R&D spillovers, $\Sigma_{j\neq i} g_{ijt} t_{ij} r_{jt}$, is the technology-weighted sum of collaborating firms' R&D inputs, and it captures the effect of firms in R&D collaborations sharing each other's R&D resources or experts/scientists. Spillovers $S_{it}$ might also include productivity spillovers, $\Sigma_{j\neq i} g_{ijt} t_{ij} \omega_{jt}$ to account for learning from collaborators' endowed knowledge or technology summarized by productivity. This will be discussed more in section 4.1, but I simply set $S_{it}$ as only R&D spillovers for the benchmark. The remaining uncertainty, $\nu_{it+1}$, reflects the randomness inherently embedded in the R&D process, including success in innovation or discovery. The model captures important aspects of productivity evolution by assuming that $\omega_{it+1}$ not only persists over time but is also allowed to be shifted by other dynamic choices, including R&D investment and spillovers from R&D collaborations.

The identification of improvement effects from private R&D and spillovers is based on a structural assumption about timing. The evolution process implicitly presumes that learning occurs with a delay which is why private R&D and spillovers are lagged, implying that the improvements in firm productivity take a period to materialize (Malikov and Zhao 2021). Consequently, the identification of the equation (2) is obtained once we know productivity levels.

This production model does not specify a negative business-stealing effect from collaborations. As discussed in section 2, most R&D collaborating firms are across-market or buyer-supplier pairs, so they do not suffer from market-rival effects from collaborations. However, in section 3.2, to allow firms' tendency to avoid market rivals when

[11]For simplicity, I would suppress the capital stock throughout the theoretical framework.

[12]Jaffe technological similarity measure is defined as $\frac{F_i' F_j}{\sqrt{F_i}\sqrt{F_j}}$ where $F_i$ denotes the vector of the shares of firm $i$'s patenting in different technology field.

choosing collaborators in anticipation of potential business-stealing effects, the measure for market similarity between two firms will be considered as a form of networking cost.

## 3.2 R&D collaboration decision

This framework considers R&D collaborations, $\mathbf{g}_{it} = (g_{i1t}, ..., g_{iNt})$, as firms' decisions. Therefore, R&D investment $r_{it}$ not only enhances productivity in the next period but also influences the formation of next-period collaborations, $\mathbf{g}_{it+1}$, leading to higher spillovers in the future. This subsection establishes the micro-foundation for internally developed R&D collaboration networks.

Throughout this subsection, time subscript $t$ is suppressed, and superscript $'$ will refer to the future period in order to simplify the exposition.

**Timeline** Before introducing the model, I clarify a firm's timeline for dynamic decisions in a period. A firm's timeline within a period $t$ is given as follows:

1. At the beginning of each period, firm-specific productivity $\omega_i$ is realized and short-term profit $\pi(\omega_i)$ is determined.
2. A firm decides on the level of R&D investment $r_i$ under the given productivity level $\omega_i$ and the current collaboration status $\mathbf{g}_i = (g_{i1}, ..., g_{iN})$.
3. For the remainder of the period, a firm determines its R&D collaborations for the next period, $\mathbf{g}'_i = (g'_{i1}, ..., g'_{iN})$.

Our main interest is in the third stage, the R&D collaboration decision stage. The first stage, where the short-term profit is determined, is discussed in the previous section 3.1. Subsequently, in the second stage, the firm decides on the level of R&D investment to improve productivity under the given collaboration status. This can be a strategic game between firms. Since we would focus on the third stage, I would rely on the assumption that the equilibrium of firms' R&D investments exists and that R&D levels we observe are determined by one of the equilibria if there are many. A specific model for R&D investment can be found in Appendix B. For the rest of the period, a firm communicates with potential R&D collaboration partners and determines its one-period R&D collaborations for the next period. In the case of the existing collaborations, firms would update and determine whether they continue or quit the next period. This timeline reflects the description of the process of R&D collaboration in section 2 and implies that forming R&D collaboration takes time to be effective due to the process of communication and making contracts or evaluating the process.

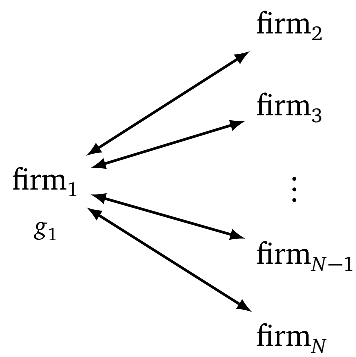



Figure 4: Firm choosing $N-1$ collaborations

**Information set** In the R&D collaboration decision stage, a firm can potentially collaborate with any other firms in the economy. First, the firm draws a vector of unobserved (to researchers) *i.i.d* networking cost shocks $\nu_i = (\nu_{i1}, ..., \nu_{iN})$. It captures innate uncertainty in a communication process or matching niche needs for technology. It is pair-specific and symmetric between two firms (i.e., $\nu_{ij} = \nu_{ji}$). Also, it can observe all firms' productivity $\omega = (\omega_1, ..., \omega_N)$, R&D input levels $\mathbf{r} = (r_1, ..., r_N)$, a current collaboration network $\mathbf{g} = [g_{ij}]_{i,j}$ and networking cost variables $\mathbf{x} = [\mathbf{x}_{ij}]_{i,j}$ in the economy. A vector of networking cost variables $\mathbf{x}_{ij}$ is used to capture heterogeneity in networking costs, such as similarities and past collaboration status. I will denote the observable information as a matrix of state variables $\mathbf{s} = (\omega, \mathbf{r}, \mathbf{g}, \mathbf{x})$.

**Collaboration decision problem** After a firm observing $(\mathbf{s}, \nu_i)$, a firm determines or updates its one-period collaboration status with $N-1$ other firms as in Figure 4. Every collaboration is one-on-one, and there is no complementarity or substitution from forming a collaboration with more than one firm. This assumption relies on the pattern in the data described in section 2, in which 90% of official collaborations involve only 2 firms. A value function for a firm that represents this procedure is given as follows[13]:

$$V_i(\mathbf{s}, \nu_i) = \pi(\omega_i) + \max_{\mathbf{g}_i^{*\prime} = (g_{i1}^{*}{}', ..., g_{iN}^{*}{}')} \left\{ \rho E_t \Big[ V_i(\mathbf{s}', \nu_i') \Big| \mathbf{s}, \mathbf{g}_i \Big] - \Sigma_{j \neq i} g_{ij}' (C(\mathbf{x}_{ij}) + \nu_{ij}) \right\} \tag{3}$$
$$\text{s.t. } g_{ij}' = 1 \text{ if } g_{ij}^{*}{}' = 1 \text{ and } g_{ji}^{*}{}' = 1 \text{ for all } j$$

where $C(\cdot)$ is a pair-specific networking cost function and $\rho$ is a discount factor. A firm chooses its latent collaboration status $\mathbf{g}_i^*$ by comparing the potential benefits expected to obtain from collaborations and the networking cost. The networking cost is incurred

[13]If there is any R&D cost incurred in the R&D investment decision stage, it would be included in the model, but I suppress it for simplicity.

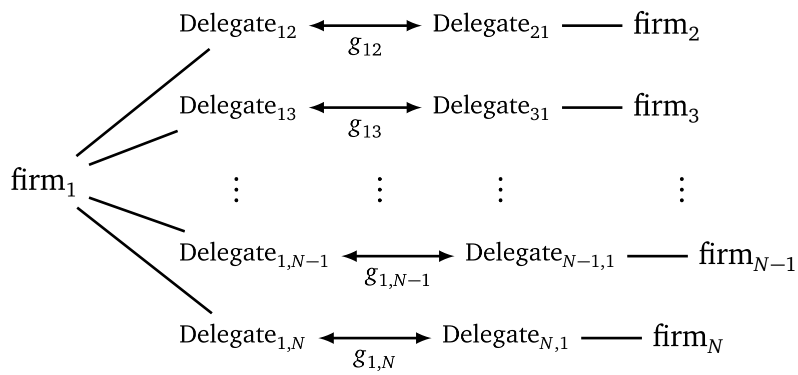



Figure 5: Firm delegating $N-1$ collaboration decisions

once the collaboration is determined and involves the network formation and networking costs, such as the costs for drafting a contract and communication.[14] The cost function $C$ is given as:

$$C(\mathbf{x}_{ij}) = \gamma_0 + \gamma_1 g_{ij} + \gamma_2 a_{ij} + \gamma_3 a_{ij}^2 + \mathbf{d}_{ij}^{\intercal}\gamma_d \tag{4}$$

where $\mathbf{x}_{ij} \equiv (g_{ij}, a_{ij}, a_{ij}^2, \mathbf{d}_{ij}^{\intercal})^{\intercal}$ is a $(x_c \times 1)$ vector of networking cost-related variables, $a_{ijt}$ is the duration of collaboration, and $\mathbf{d}_{ijt}$ is a vector of technology, market, geographical, and R&D-size similarities. It implies networking costs depend on previous relationships and various similarities between the two firms. Then, if corresponding firms agree with their collaboration, the actual R&D collaboration between them, $g_{ij}$, will be effective in the next period.

### 3.2.1 Alternative approach and reducing the dimensionality

The described framework, however, entails some problems that complicate the estimation of the model: (1) The action set of the firm's choice on collaborations, $\mathbf{g}_i$, is $\{0,1\}^{n-1}$, therefore the size of the choice set is $2^{n-1}$ which is huge even with a moderate number of firms. (2) Forming collaborations requires agreements from corresponding firms and is a strategic game between firms. (3) The dimension of the state variables $\mathbf{s} = (\omega, \mathbf{r}, \mathbf{g}, \mathbf{x})$ is enormous because it includes information on the current collaboration network in the industry and all other firms' productivity and R&D levels. I address those difficulties by introducing (1) collaboration delegates and (2) inclusive values and bounded rationality of delegates.

[14]The actual networking cost is more likely to be incurred after the next period collaborations become effective. However, the timing of incurring networking cost—before or after the collaboration is effective—does not affect the model if the idiosyncratic part of cost $\nu_i$ is known at period $t$.

**Collaboration delegates** Instead of a firm itself deciding $n-1$ collaborations simultaneously, I assume that it delegates its $n-1$ decisions to agents for collaborations—I call them *collaboration delegates*. Now every firm has $n-1$ collaboration delegates, one for each possible collaboration. It sends collaboration delegates to other firms and makes them discuss the collaboration decisions with other firms' collaboration delegates, as described in Figure 5. Therefore, for a collaboration $ij$, a delegate-pair $ij$ (equivalently, a delegate-pair $ji$) from firms $i$ and $j$ together decides on the collaboration status for the next period, $g'_{ij}$, by maximizing the sum of the two firms' expected profit. A distinguishing aspect of the collaboration delegate model is that a delegate-pair $ij$ can only observe their own private cost shock, $\nu_{ij}$, but not other delegate-pairs' private cost shocks, including firm $i$ or $j$'s inner relationships' shocks, $\nu_{ik}$ or $\nu_{jk}$ for $k \neq i, j$.

A similar idea has been used in the game-theoretic analysis, including the relationship between upstream and downstream firms (Bjornerstedt and Stennek 2007; Collard-Wexler, Gowrisankaran, and Lee 2019) and the problem of firms with many markets such as airline industries (Aguirregabiria and Ho 2012), but it has not been employed in a collaboration setting. Especially in the context of R&D collaborations, this assumption could be plausible. An R&D collaboration is usually described as conducting an R&D project together by a group of inventors or scientists. The detailed aspects of the project tend to be recognized fully only by the involved people. It corresponds to the collaboration delegates model where the delegate of collaboration cannot perfectly observe all the information of other collaborations. However, it excludes the cases when delegates in a firm can communicate and share their private information.

This model also excludes the case when more than 2 firms form an R&D consortium together. In that case, they might choose collaboration status by maximizing the sum of all engaged firms' expected values, not every firm pair's. However, as described in section 2, more than 90% of official collaborations in the U.S. involve only 2 firms. Thus, the problem caused by limiting the formation to one-on-one would not be severe in this empirical setting.

**Assumption 1.**

1. *A firm assigns a delegate for each collaboration. Firm $i$ and $j$'s delegates for the collaboration $ij$ choose $g'_{ij} \in \{0,1\}$ to maximize the expected and discounted value of the stream of the sum of two firms' profits.*

2. *The shock for the network formation $\{\nu_{ij}\}$ is private information of a delegate-pair $ij$. This shock is unknown to the delegates $ik$, $jk$, and $\ell k$, $\ell, k \neq i, j$ and identically and independently distributed.*

By the assumption 1, at the cost of limiting delegate-pairs' information on other delegate-pairs' cost shocks, the game of $n$ firms simultaneously choosing its $n-1$ possible collaborations $\mathbf{g}_i' = (g_{i1}', ..., g_{in}')' \in \{0,1\}^{n-1}$ becomes the incomplete information game of $n(n-1)/2$ delegates deciding only one connection $g_{ij}' \in \{0,1\}$, respectively.

To write down this framework into a model, let $V_{ij}$ represent firm $i$'s value function *from the perspective of delegate-pair* $ij$ given as follows:

$$V_{ij}(\mathbf{s}, \nu_{ij}) = \pi(\omega_i) + \left\{ E_t\left[ \rho V_{ij}(\mathbf{s}', \nu_{ij}') - \mathrm{TC}_{-ij} \middle| \mathbf{s}, g_{ij}' \right] - g_{ij}'(C(\mathbf{x}_{ij}) + \nu_{ij}) \right\} \tag{5}$$

where $\mathrm{TC}_{-ij} = \Sigma_{k \neq i,j} g_{ik}'(C(\mathbf{x}_{ik}) + \nu_{ik})$ is the total networking cost except a link $ij$. A firm $j$'s value function from the perspective of delegate-pair $ij$, $V_{ji}$, is similarly defined. Using the value function, the expected marginal benefits of firm $i$ and $j$ from the collaboration $ij$ are respectively given by:

$$\begin{aligned} MU_{ij}(\mathbf{s}) &= EV_{ij}^1 - EV_{ij}^0 - C(\mathbf{x}_{ij}) \\ MU_{ji}(\mathbf{s}) &= EV_{ji}^1 - EV_{ji}^0 - C(\mathbf{x}_{ji}) \end{aligned} \tag{6}$$

where $EV_{ij}^1 = E[\rho V_{ij}(\mathbf{s}', \nu_{ij}') - \mathrm{TC}_{-ij} | \mathbf{s}, g_{ij}' = 1]$ and $EV_{ij}^0 = E_t[\rho V_{ij}(\mathbf{s}', \nu_{ij}') - \mathrm{TC}_{-ij} | \mathbf{s}, g_{ij}' = 0]$ are expected firm values when collaborating and not collaborating, respectively. Then, the equilibrium concept of R&D collaboration networks considered in this model is *pairwise-stable with transfer* (Bloch and Jackson 2007). It implies that a delegate pair $ij$ would agree to collaborate if the sum of two firms $i$ and $j$'s expected marginal utilities is positive:

$$g_{ij}' = 1 \iff MU_{ij}(\mathbf{s}) + MU_{ji}(\mathbf{s}) - (\nu_{ij} + \nu_{ji}) > 0 \tag{7}$$

It allows one side of the link to present a negative marginal payoff but only up to both firms enjoying positive utility with a transfer. Based on this R&D network decision rule, the network formation process follows stochastic best-response dynamics.

**Inclusive values and bounded rationality** Yet there remains the large dimensionality problem of the state variables because it is a strategic game between delegate-pairs. The estimation of the expected firm values in the delegate's best response (7) hinges on the state variables conditioned on, $\mathbf{s} = (\omega, \mathbf{r}, \mathbf{g}, \mathbf{x}) \in \mathbb{R}^N \times \mathbb{R}^N \times \mathbb{R}^{N \times N} \times \mathbb{R}^{dN \times N}$, and the matrix of beliefs on other delegate-pairs' decisions, which makes computation intractable. I solve this problem by proposing a vector of *inclusive values* and imposing the bounded rationality of delegates. Inclusive values are the variables that aggregate the essential or payoff-relevant information in the state variables, $\mathbf{s} = (\omega, \mathbf{r}, \mathbf{g}, \mathbf{x})$ (Hendel and Nevo, 2006; Nevo and Rossi, 2008; Aguirregabiria and Ho, 2012; Kalouptsidi, 2018). Then delegate-pairs rely only on the information of inclusive values, not the entire state variables, which implies bounded rationality.

Let $\mathbf{w}_{ij}$ be a vector of inclusive values for firm $i$ from delegate $ij$'s perspective, which I will examine each value after defining the bounded rationality.

**Assumption 2.** *(Bounded rationality)*
*The expectation of firm $i$'s value from the delegate $ij$'s perspective depends on state variables $\mathbf{s} = (\omega, \mathbf{r}, \mathbf{g}, \mathbf{x})$ only through a vector of inclusive values, $\mathbf{w}_{ij}$.*

$$E[\rho V_{ij}^{i}(\mathbf{s}', \nu_{ij}') - TC_{-ij} | \mathbf{s}, g_{ij}'] = E[\rho V_{ij}^{i}(\mathbf{s}', \nu_{ij}') - TC_{-ij} | \mathbf{w}_{ij}, g_{ij}'] \tag{8}$$

*The value function of firm $j$ from the delegate $ij$'s perspective can be similarly assumed.*

This assumption states that collaboration delegates are boundedly rational in their perceptions of how the information in the industry would influence their future values. In other words, the inclusive values summarize the public information well enough to represent the state variables. It could be rather unrealistic to presume that delegate-pairs take every individual firm's detailed information into consideration. Hence, it might not be too restrictive to impose bounded rationality that delegate-pairs extract only important or aggregated information.

A vector of inclusive values that are boiled down from the state variables $\mathbf{s} = (\omega, \mathbf{r}, \mathbf{g}, \mathbf{x})$ is given by:

$$\mathbf{w}_{ij} = (\omega_i, r_i, \mathbf{x}_{ij}, S_i, \widehat{\mathrm{TC}}_{-ij}^{\intercal}(\mathbf{P}), \phi_{ij}(\mathbf{P})^{\intercal})^{\intercal} \in \mathscr{W} \tag{9}$$

where $S_i$ is R&D spillover defined above; $\mathrm{P}_{ij}(\mathbf{w}, \tilde{\mathbf{w}}) \equiv \Pr(\sigma_{ij}(\mathbf{w}_{ij}, \mathbf{w}_{ij}, \nu_{ij}, \nu_{ji}) = 1 | \mathbf{w}_{ij} = \mathbf{w}, \mathbf{w}_{ji} = \tilde{\mathbf{w}})$ for $\mathbf{w}, \tilde{\mathbf{w}} \in \mathscr{W}$ is a belief (conditional choice probability) for collaboration $ij$ associated with $\sigma$ and $\sigma = \{\sigma_{ij}(\mathbf{w}_{ij}, \mathbf{w}_{ji}, \nu_{ij}, \nu_{ji}) : i, j = 1, \ldots, n, i \neq j\}$ is a matrix of strategy functions, one for each delegate-pair; $\mathbf{P} = \{\mathrm{P}_{ij}(\mathbf{w}, \tilde{\mathbf{w}}) :$ for every $ij$ and $\mathbf{w}, \tilde{\mathbf{w}} \in \mathscr{W}\}$ is a matrix of beliefs; $\widehat{\mathrm{TC}}_{-ij}(\mathbf{P}) = \Sigma_{k \neq i,j} P_{ik} \mathbf{x}_{ik}$ is a vector of expected total networking cost variables with the belief $\mathbf{P}$; and $\phi_{ij}(\mathbf{P})$ is a vector of the moments of (expected) networks that summarize the networks such as the expected number of collaborations.

The first three variables, a firm's own variables $(\omega_i, r_i, \mathbf{x}_{ij})$, are the most important variables to estimate the expected firm $i$'s value. Spillover term, $S_i$, aggregates other firms' R&D information. The total expected networking cost $\widehat{\mathrm{TC}}_{-ij}$ summarizes the information from the whole matrix of networking cost variables $\mathbf{x}$ except a link $ijt$ using the belief. These variables are directly relevant to the payoff. Lastly, to avoid considering all other delegate-pairs' information due to the strategic game, a delegate-pair only considers some moments of the beliefs on other delegate-pairs' choices, $\phi_{ij}(\mathbf{P})$, which summarize the whole information. I would denote $\mathbf{w}_{ij}^1 = (\omega_i, r_i, \mathbf{x}_{ij}, S_i)$ as an exogenous part of inclusive values and $\mathbf{w}_{ij}^2 = (\widehat{\text{TC}}_{-ij}(\mathbf{P}), \phi_{ij}(\mathbf{P}))$ as an endogenous part of inclusive values.

At the cost of delegate-pairs' rationality, the best response of a delegate-pair in the equation (7) becomes:

$$g'_{ij} = 1 \iff MU(\mathbf{w}_{ij}) + MU(\mathbf{w}_{ji}) - (\nu_{ij} + \nu_{ji}) > 0 \tag{10}$$

where $MU(\mathbf{w}_{ij}) = MU_{ij}(\mathbf{s})$ and $MU(\mathbf{w}_{ji}) = MU_{ji}(\mathbf{s})$. The marginal utilities draw on vectors of inclusive values $\mathbf{w}_{ij}$ and $\mathbf{w}_{ji}$, which have the feasible number of variables rather than all state variables.

Based on the inclusive values, I define a Moment-based Markov Equilibrium (MME) as follows (Ifrach and Weintraub 2017).

**Definition 1.** *(Moment-based Markov Equilibrium)*
*Each delegate-pair's strategy depends only on the state variables through the inclusive values. A Moment-based Markov Equilibrium (MME) constitutes delegate-pairs' strategy profile and beliefs* $(\sigma^*, \mathbf{P}^*)$ *such that*

1. *Given beliefs* $\mathbf{P}^*$*, each delegate-pair's strategy* $\sigma_{ij}^*$ *maximizes the sum of the two firms' values for each possible state* $(\mathbf{w}_{ij}, \mathbf{w}_{ji}, \nu_{ij}, \nu_{ji})$.

2. *Beliefs* $\mathbf{P}^*$ *satisfies the self-consistency requirement:*

$$\mathbf{P}^* = \Big[ \Pr(g'_{ij} = 1 | \mathbf{w}_{ij}(\mathbf{P}^*), \mathbf{w}_{ji}(\mathbf{P}^*)) \Big]_{i,j \in \mathcal{N}} \tag{11}$$

   *Or, equivalently it satisfies* $P_{ij}^*(\mathbf{W}^1) = \Pr(g_{ij} = 1 | \mathbf{W}^1, \mathbf{P}^*)$ *where* $\mathbf{W}^1 = [\mathbf{w}_{ij}^1]_{i,j \in \mathcal{N}}$ *is a matrix of all delegate-pairs' exogenous part of inclusive values.*

The existence of an MME of the game follows from Ifrach and Weintraub (2017).

## 3.3 Direct and indirect effect of R&D

From this theoretical framework, we can define the one-time *direct* effect of private R&D on productivity as follows:

$$\text{DE}_{it} = \frac{\partial E[\omega_{it+1} | \cdot]}{\partial r_{it}} = \beta_2 \tag{12}$$

where $\beta_2$ is the effect of R&D on next period productivity in equation (2). The direct effect is what canonical models define as the effect of private R&D. However, in this framework, private R&D allows firms to expand collaboration networks and eventually enjoy higher spillovers. I define it as the *indirect* effect of private R&D through network formation. The one-time *indirect* effect through network formation and spillovers can be written as:

$$\mathrm{IE}_{it} = \frac{\partial E[\omega_{it+2}|\cdot]}{\partial r_{it}} - \beta_1\beta_2 = \beta_3 \frac{\partial E[S_{it+1}|\cdot]}{\partial r_{it}} \tag{13}$$

where $\beta_1\beta_2$ represents R&D's effect through time-persistency in productivity and $\beta_3$ is the effect of R&D spillovers on productivity in equation (2). Thus the total one-time effect of R&D on productivity would be $\mathrm{DE}_{it} + \mathrm{IE}_{it}$. In the empirical estimation, I would compare the direct and indirect effects of R&D and how large it becomes when considering the total effect rather than just the direct effect, as we did in the canonical model. Also, since its effect cumulates over time as productivity and collaborations are usually time-persistent, the long-term effect will also be discussed.

# 4 Estimation of the Structural Model

This section demonstrates the identification and estimation strategies for the model in section 3. My approach to estimating the structural model proceeds in two steps. First, the production function estimation is conducted, accounting for the endogenous productivity evolution, following the modified Ackerberg, Caves, and Frazer (2015) approach. Second, I estimate the network formation function by matching delegates' best response and the R&D collaboration data. Given the estimates from this section, the comparison between the direct and indirect effects of R&D will be discussed.

## 4.1 Production function estimation

The main obstacle in the empirical estimation of production function is that the unobserved (by the econometrician) firm-specific productivity $\omega_{it}$ could be correlated with flexible input variables such as labor. To address this problem, I adopt the control function approach that is proposed by Olley and Pakes (1996) with the correction advocated by Ackerberg, Caves, and Frazer (2015) (henceforth ACF). This method leverages the information contained in the capital investment level $i_{it}$, as productivity is known to each firm when it decides on capital investment.

Obtaining consistent estimates of productivity $\omega_{it}$ and the parameters in equation (1) and (2) requires three sets of assumptions. The first set relates to the timing of firms' decisions. Capital is a state variable, determined in the preceding period as a deterministic function of the firm's previous capital stock and its investment decision: $k_{it} = \kappa(k_{it-1}, i_{it-1})$. Labor, on the other hand, may or may not have dynamic implications. It may be fully adjustable and chosen after productivity is realized or partly (or wholly) determined in the previous period. It, however, needs to be chosen prior to the capital investment $i_{it}$ decision. R&D levels and spillover levels are also chosen or revealed before the capital investment decision.

Based on those variables and productivity, the firm chooses capital investment according to the following function:

$$i_{it} = \mathbb{I}(k_{it}, l_{it}, \omega_{it}, r_{it}, S_{it}) \tag{14}$$

where $\mathbb{I}(\cdot)$ is the demand function of capital investment. Lastly, the demand for capital investment $\mathbb{I}(\cdot)$ is strictly monotonic in productivity conditional on other variables. Compared to the original ACF assumptions, this assumption additionally considers R&D investment and spillover terms because the evolution of productivity depends on them. Thus this assumption on the demand for capital investment extends the ACF assumption but is employed in recent papers considering spillover effects as well (De Loecker 2013; Arque-Castells and Spulber 2022). It requires me to check the validity of this assumption in the context of spillovers and forming collaborations. It could be violated if capital negatively affects network formation because a firm might not want to increase capital investment even though productivity increases to avoid a potential decrease in spillovers in the future. However, I found that there is no significant effect of capital on forming collaborations, as shown in Table 3. Unless it affects network formation, there is no clear incentive for a firm not to increase capital investment as productivity increases at a certain level of spillovers and R&D level. In addition, I check the *ex-post* validity of the strict monotonicity assumption shown in Appendix B.

This guarantees that productivity can be expressed solely as a function of observables $\omega_{it} = \mathbb{I}^{-1}(k_{it}, l_{it}, i_{it}, r_{it}, S_{it})$. Substituting into the production function (2) yields:

$$\begin{aligned} y_{it} &= \alpha_0 + \alpha_k k_{it} + \alpha_l l_{it} + \mathbb{I}^{-1}(k_{it}, l_{it}, i_{it}, r_{it}, S_{it}) + \epsilon_{it} \\ &= \varphi(k_{it}, l_{it}, i_{it}, r_{it}, S_{it}) + \epsilon_{it} \end{aligned} \tag{15}$$

where $\varphi(\cdot)$ is an unknown function of $E[y_{it}|k_{it}, l_{it}, i_{it}, r_{it}, S_{it}]$. I am able to nonparametrically estimate $\varphi$ and productivity $\omega_{it}$ can be expressed as:

$$\omega_{it} = \varphi_{it} - \alpha_0 - \alpha_k k_{it} - \alpha_l l_{it} \tag{16}$$

where $\varphi_{it} = \varphi(k_{it}, l_{it}, i_{it}, r_{it}, S_{it})$. Then in the productivity evolution equation (2), substituting $\omega_{it-1}$ with the equation (15) yields:

$$\begin{aligned} \omega_{it} &= \beta_0 + \beta_1 \omega_{it-1} + \beta_2 r_{it-1} + \beta_3 S_{it-1} + u_{it} \\ &= \beta_0 + \beta_1(\varphi_{it-1} - \alpha_0 - \alpha_k k_{it-1} - \beta_l l_{it-1}) + \beta_2 r_{it-1} + \beta_3 S_{it-1} + u_{it} \end{aligned} \tag{17}$$

Then the suitable set of moments is given as:

$$E[(u_{it} + \epsilon_{it})[k_{it}, l_{it-1}, \varphi_{it-1}, r_{it-1}, S_{it-1}]] = 0 \tag{18}$$

With this set of moments, production elasticities and parameters in the production evolution process can be consistently estimated by GMM.

**Productivity spillovers** One might think that spillovers $S_{it}$ could include not only R&D spillovers $\Sigma_{j\neq i} g_{ijt} t_{ij} r_{jt}$, but also productivity spillovers $\Sigma_{j\neq i} g_{ijt} t_{ij} \omega_{jt}$. Thus firms can learn from collaborating firms' overall knowledge level as well. I propose an extended approach to consider both R&D and productivity spillovers, $S_{it} = [\Sigma_{j\neq i} g_{ijt} t_{ij} r_{jt},\ \Sigma_{j\neq i} g_{ijt} t_{ij} \omega_{jt}]$, where productivity spillovers are not observed. With productivity spillovers, the typical ACF method suggested above cannot be directly employed because I cannot recover a function $\varphi(\cdot)$ in equation (15) as now $S_{it} = [\Sigma_{j\neq i} g_{ijt} t_{ij} r_{jt},\ \Sigma_{j\neq i} g_{ijt} t_{ij} \omega_{jt}]$ requires the information on $\omega_{jt}$ which we don't observe at this point.

To solve the unobservability of productivity spillover, I additionally assume (1) the additive separability of productivity spillover and (2) $|\lambda| < 1$ in the following equation:

$$\omega_{it} = \lambda \Sigma_{j\neq i} g_{ijt} t_{ij} \omega_{jt} + \tilde{\mathbb{I}}^{-1}(k_{it}, l_{it}, i_{it}, r_{it}, \Sigma_{j\neq i} g_{ijt} t_{ij} r_{jt}) \tag{19}$$

The second assumption $|\lambda| < 1$ implies that the effect of a firm's own productivity on the capital investment decision is larger than that of the collaborating firm's productivity. This is typically true because a firm's own productivity is more important than the collaborators'. For the simplicity of notation, I denote $z_{it} = (k_{it}, l_{it}, i_{it}, r_{it})$ as a vector of firm-specific variables except spillover terms. Then using a vector form yields:

$$\omega_t = (\mathbf{I} - \lambda G_t)^{-1} \tilde{\mathbb{I}}^{-1}(\mathbf{z}_t, G_t \mathbf{r}_t) \tag{20}$$

where $\mathbf{z}_t = (z_{1t}, ..., z_{Nt})$ and the $i$th element of $\tilde{\mathbb{I}}^{-1}(\mathbf{z}_t, G_t \mathbf{r}_t)$ is $\tilde{\mathbb{I}}^{-1}(z_{it}, \mathbf{g}_{it}^{\intercal} \mathbf{r}_t)$ and $G_t$ is technological similarity weighted network matrix with an abuse of notation. Then $|\lambda| < 1$

implies that I can approximate $(\mathbf{I}+\lambda G_t)^{-1}$ by a geometric series expansion as follows:

$$\omega_t = \Sigma_{s=0}^{\infty}\lambda^s G_t^s \tilde{\mathbb{I}}^{-1}(\mathbf{z}_t, G_t\mathbf{r}_t) \tag{21}$$

Then I can recover productivity and follow the remaining procedure. To illustrate, If I simply assume linearity of the unknown function $\mathbb{I}^{-1}(\cdot)$ and use a geometric series with $s=0,1$, then $E[y_{it}|\cdot]$ is obtained by regressing $y_{it}$ on $(z_{it}, g_{it}\mathbb{1}, g_{it}z_{it}, g_{it}^2 r_t)$ where $g_{it}^2$ is $i$th row of $G_t^2$. The intuition of this is that the information on productivity is obtained from the capital investment $i_{it}$'s decision and the information on productivity spillover is obtained from its friends' other information or the information of the friends of its friends.

## 4.2 Network formation function estimation

The primary objective of this estimation is to determine the effect of private R&D on the likelihood of forming a collaboration between two firms. This is achieved by matching the delegate-pair's best response function in equation (10) with the collaboration data using a binary-response estimation method. In the best response function, the marginal benefits of collaboration are approximated with the vector of inclusive values $\mathbf{w}_{ijt}$ and some basis functions as follows:

$$\begin{aligned} MU(\mathbf{w}_{ijt}) &\approx \Phi(\mathbf{w}_{ijt})\mathbf{c} \\ MU(\mathbf{w}_{jit}) &\approx \Phi(\mathbf{w}_{jit})\mathbf{c} \end{aligned} \tag{22}$$

where $\Phi(\cdot)$ are the basis functions and $\mathbf{c}$ is a vector of approximation parameters. The brief summarization of the estimation procedure is given as follows: (1) Estimate the equilibrium beliefs on collaboration $\mathbf{P}^*$ and obtain inclusive values $\mathbf{w}_{ijt}(\mathbf{P}^*)$. (2) Estimate the approximation parameters $\mathbf{c}$ using a probit regression. The subsequent paragraphs delve into potential issues and provide a more comprehensive overview of the estimation procedure.

**Multiple equilibria** This network-formation model fundamentally admits multiple equilibria since there could be multiple pairs of beliefs and parameters that result in the same outcome (11). One estimation strategy for the games of incomplete information is estimating equilibrium beliefs directly from the data. However, since I observe one single network, the equilibrium beliefs cannot be directly obtained from the data. Also, estimating $P^*(\mathbf{W}^1)$ as the matrix of exogenous inclusive values suffers from the curse of dimensionality as the matrix of exogenous inclusive values $\mathbf{W}^1$ grows quickly as $N$ grows. To address the problem, I follow the approach suggested in Leung (2015) by assuming symmetry of equilibrium beliefs.

In order to define symmetry in equilibrium beliefs, I need to define permutation functions. Fix any two delegate-pairs $ij, kl$ in $N \times N$. Define $\pi_{(ij)(kl)} : N \times N \rightarrow N \times N$ as it maps the index $ij$ to the index $kl$, $kl$ to $ij$. Similarly, define $\pi^w_{kl}$ as a function that swaps a delegate pair $ij$'s inclusive values and $kl$'s inclusive values. To simplify the notation, I will let $\pi(\cdot)$ denote an arbitrary element in $\Pi \equiv \{(\pi_{kl}, \pi^w_{kl}); ij, kl \in N \times N\}$, the set of permutations. Further, I will abuse notation and use $\pi$ with an element in the set of permutations and with each of its two components as arguments to the function clarify the relevant component.

**Assumption 3.** *(Symmetric equilibrium beliefs)*
*A belief function in the equilibrium, $P^*$, is symmetric, i.e., for any delegate-pair $ij \in N \times N$, and permutation $\pi \in \Pi$, $P^*_{ij}(\mathbf{W}^1) = P^*_{\pi(ij)}(\pi(\mathbf{W}^1))$*

It implies that two delegate pairs with the same attributes have the same conditional linking strategies *ex ante*, prior to their private information draws. It still admits multiple equilibria but this assumption enables the direct estimation of equilibrium beliefs from the data by simply taking the empirical frequency of observationally equivalent delegate pairs, which solves a curse of dimensionality.

**Estimation** Leveraging symmetry, this naturally suggests a consistent two-step estimator following (Leung 2015; Comola and Dekel 2022). The estimation procedure is as follows:

1. Nonparametrically estimate the equilibrium beliefs $\mathbf{P}^*$ using only the observable exogenous part of inclusive values $\mathbf{w}^1$ and obtain inclusive values $\mathbf{w}(\mathbf{P}^*)$.

2. With the obtained inclusive values, estimate the approximation parameters $\mathbf{c}$ in the equation (22) using probit regression under the assumption that networking cost shock follows a normal distribution.

**Endogeneity and R&D tax credit** Another concern that arises from the data is that R&D investment can correlate with the error term. The future collaboration might have been known to a firm, and it could affect R&D decisions in the current period. To address the endogeneity problem, I use R&D user cost induced by tax credit to instrument R&D investment levels, as first suggested in Bloom, Schankerman, and Van Reenen (2013) and exploited in many papers (Konig, Liu, and Zenou 2019; Zacchia 2020; Arque-Castells and Spulber 2022). R&D tax credit is a government-sponsored stimulus in tax to promote innovation by incentivizing R&D investments. It is a supply-side shock from tax-induced changes to individual costs of R&D and, therefore, free from the current period demand shock. The calculation of R&D depends on various firm features and has changed multiple times. Generally, it depends on the dates of incorporation and initiation of research, gross revenue, and research expenses for the past few years. It also depends on state R&D tax credits so that a firm's tax credit relies on multiple states' tax credits based on where its inventors and scientists are. Thus, firm-specific R&D user cost is constructed using tax credits of R&D.[15]. The rationale for this identification strategy is that lagged R&D tax credit directly affects R&D investments but only indirectly affects the formation of collaborations. I use a control function approach following Wooldridge (2015). The residuals from a regression of R&D user cost on all other variables, including R&D investments, are added to a probit regression as additional variables.

# 5 Empirical Results

In this section, I provide the empirical estimation results of the framework and examine the R&D's *direct* and *indirect* effect on productivity through the R&D collaboration network. In the production function estimation, I don't separate industry-level estimations because many collaborations are across various industries. Instead, I add industry-fixed effects. A first-degree polynomial is used to approximate $\varphi(\cdot)$ in the first-stage estimation of ACF.[16] The estimation of the network formation model uses 6,859,270 observations ($n \times (n-1) \times T$). For a collaboration $ij$, since I cannot distinguish firm $i$ and firm $j$'s influences on the formation separately (i.e., the approximation parameter $c$ for a variable is the same for firm $i$ and $j$'s), I report one of them, which is the averaged effect. The belief estimation in the first stage utilizes the second polynomial sieve estimation. The second stage also uses the second polynomial bases functions for the approximation. In addition, for the variables of the moments of the beliefs, $\phi_{ij}(\mathbf{P}_t)$, I use the expected sum and mean of collaboration beliefs $\phi_{ij}(\mathbf{P}_t) = [\Sigma_{j \neq i} P_{ijt} \quad (1/N)\Sigma_{j \neq i} P_{ijt}]^{\intercal}$.

[15] See Appendix B.3 in the Supplementary Material of Bloom, Schankerman, and Van Reenen (2013) for details on the specification of R&D tax credit

[16] Since the data set is not large enough but covers many different industries, errors from using more flexible basis functions seem to be more than benefits.

Table 2: Production function estimation

| | | (1) | (2) | (3) |
|---|---|---|---|---|
| $r_{it-1}$ | | 0.0146** | 0.0097** | 0.0097** |
| | | (0.0005) | (0.0005) | (0.0005) |
| $S_{it-1}$ | R&D spillovers | | 0.0024** | 0.0032** |
| | | | (0.0001) | (0.0008) |
| | $\omega$ spillovers | | | −0.0005 |
| | | | | (0.0005) |
| $\omega_{it-1}$ | | 0.6873** | 0.6870** | 0.6869** |
| | | (0.0067) | (0.0067) | (0.0067) |
| $l_{it}$ | | 0.7878** | 0.7913** | 0.7915** |
| | | (0.0487) | (0.0478) | (0.0480) |
| $k_{it}$ | | 0.2080** | 0.2010** | 0.2008** |
| | | (0.0625) | (0.0618) | (0.0620) |
| Constant | | 2.6295** | 2.6706** | 2.6722** |
| | | (0.0566) | (0.0574) | (0.0574) |
| Obs | | | 13,525 | |

All specifications are controlled for year- and industry-fixed effects.
*Significant at 0.05% level, **Significant at 0.01% level.

## 5.1 Estimates of input elasticities and productivity process

Table 2 outlines the estimated coefficients of the production function and the productivity evolution process. The estimated coefficients that differ significantly from zero at the 0.01 and 0.05 significance levels are denoted by double and single asterisks, respectively. The positive coefficient estimates for $r_{it-1}$ and $S_{it-1}$ suggest that firms that either invest in their own R&D or collaborate on R&D projects with other firms have higher future productivity levels compared to those that do not on average. In column (1), not accounting for spillover effects, the direct marginal effect of R&D on output is 0.015, corroborating findings from previous research.[17] Accounting for spillover effects, however, reduces the average impact of R&D to 0.0097 in columns (2) and (3). In those columns, while the coefficients for R&D spillovers are positively significant with the marginal effect of 0.0024-0.0032, the coefficient for productivity spillovers is insignificant. It implies that the spillover effect from collaborating firms is captured by collaborators' R&D spillovers rather than productivity spillovers.

Based on the result in column (2), a 10 percent increase in private R&D investment in a single period results in a 0.097 percent efficiency gain in the following period, while

[17] Hall, Mairesse, and Mohnen (2010) report variations in the effects of R&D on revenue ranging from 0.01 to 0.25, centered around 0.08. Doraszelski and Jaumandreu (2013) present estimates of elasticity on output for 10 Spanish manufacturing industries, with the average value for all firms standing at 0.015.

a 10 percent increase in the most technologically similar collaborating firm's R&D input results in a 0.024 percent increase in the next period productivity. Also, the effect of past productivity on the current productivity level is captured by the coefficient of $\omega_{it-1}$. Past productivity demonstrates time persistence with an estimated coefficient of 0.687, indicating a substantial long-run effect of one-time shock. Leveraging the persistent nature of productivity, investing 10 percent more in R&D in a single period results in a *direct* long-term efficiency gain of 0.3 percent.

The remaining variables in the profit function—capital and labor—also exert significant effects, with respective coefficients of 0.2010 and 0.7913. The sum of these estimated coefficients is approximately 1, indicative of a constant return to scale.

## 5.2 Estimates of network formation parameters

Table 3 displays the estimated average marginal effects of network formation variables except for total cost variables. The numbers in the Table represent the average percentage point increase in the probability of forming a link from the marginal increase of each variable, and standard errors are in parenthesis. The average marginal effects are usually small because the collaboration network is sparse; around 0.5% of all possible firm pairs are connected. The second column is the Probit regression result, and the third column is the IV-Probit regression result using R&D user cost induced by the tax credit as an instrument. $v_{it}$ indicates the residual from the first estimation stage for the IV-Probit regression.

The average marginal effect of lagged R&D investment $r_{it-1}$ is 0.024 percent point in Probit regression but increases to 0.091 percent point in IV-Probit regression. The endogeneity problem embedded in lagged R&D is significant as the residual from the first stage estimation $v_{ijt}$ is negatively significant. This indicates there could be a negative simultaneity problem in collaboration $g_{ijt}$ and lagged R&D investment $r_{it-1}$. Intuitively, if a next-period collaboration is already determined or known to a firm, it tends to reduce R&D investment.

Based on the IV-Probit result, a 10 percent increase in lagged R&D investment increases the expected number of collaborators by 3 percent for firms engaged in any collaboration and increases the probability of having at least one collaboration by 6 percent for firms in no collaboration. In addition, lagged R&D's effect on increasing the number of collaborators is heterogeneous depending on how many collaborators a firm already has. I first divide firms into the groups with 1 collaborator, 1 $<$ collaborators $\leq$ 3, 3 $<$ collaborators $\leq$ 9, and more than 9 collaborators, considering that 1st, 2nd, and 3rd quartiles

Table 3: Network formation estimation

| Second stage estimation | | Probit | IV-Probit |
|---|---|---|---|
| $r_{it-1}$ | | 0.024** (0.005) | 0.091** (0.028) |
| $\omega_{it-1}$ | | 0.228** (0.027) | 0.039 (0.092) |
| $k_{it-1}$ | | 0.007 (0.004) | −0.016 (0.010) |
| $S_{it-1}$ | | 0.007** (0.001) | 0.009** (0.001) |
| $\text{Previous}_{ijt-1}$ | | 0.601** (0.005) | 0.613** (0.004) |
| $\text{Duration}_{ijt-1}$ | | 0.173** (0.002) | 0.176** (0.002) |
| $\text{R\&D similarity}_{ijt-1}$ | | 0.042** (0.008) | −0.027 (0.022) |
| $\text{Geographical similarity}_{ij}$ | | 0.072** (0.005) | 0.073** (0.005) |
| $\text{Tech similarity}_{ij}$ | | 0.159** (0.004) | 0.140** (0.007) |
| $\text{Market similarity}_{ij}$ | | 0.117** (0.009) | 0.117** (0.009) |
| $\varphi(\mathbf{P}_{it})$ | $\Sigma_{j\neq i}P_{ijt}$ | −0.019** (0.007) | −0.059** (0.018) |
| | $(1/N)\Sigma_{j\neq i}P_{ijt}$ | 0.104** (0.028) | 0.117** (0.030) |
| $v_{ijt}$ | | | −0.067** (0.027) |
| Pseudo-$R^2$ | | 0.753 | 0.753 |
| **First stage estimation** | | | |
| $\text{R\&D user cost}_{it-1}$ | | | −0.867** (0.006) |
| Observations | | 6,859,270 | |

The numbers in the table are the average percentage point increase in the probability of forming a link from the marginal increase of each variable, and standard errors are in parenthesis. The estimates for the expected total cost variables, $\widehat{\text{TC}}_{-ij}$, are excluded in this table. All specifications are controlled for year- and industry-fixed effects. $v_{ijt}$ is the residual from the first stage estimation for the IV-Probit regression. *Significant at 0.05% level, **Significant at 0.01% level.

of the number of collaborators are 1,3, and 9, respectively. For the first group with only one collaborator, the number of next-period collaborations is expected to increase by 5 percent from a 10 percent increase in R&D. Then its effect decreases as the number of collaborators increases; 3 percent, 2 percent, and 1 percent in each following group. It is intuitively plausible since the marginal benefit from adding more collaborators decreases as a firm already has many collaborators.

Lagged productivity $\omega_{it-1}$ marginally increases probability by 0.228 percent point in Probit regression, but it loses its explanatory power to R&D investment in IV-Probit result. There are intuitively two conflicting reasons for highly productive firms to form or not to form collaborations. Highly productive firms are easier to have collaborations as attractive collaboration partners, but they might not have enough incentive to work with others since they are already more productive than others. Also, collaborations are time-persistent. If two firms have ever been in a collaboration before, it increases the probability of forming a collaboration again by 0.613 percent point. More specifically, one more year of the duration of collaboration raises the likelihood additionally by 0.176 percent point. In addition, firms are more likely to form a collaboration as they are more geographically, technologically, or market-wise similar. The Jaffe technology and market similarity measures are used for technology and market similarities between firms $i$ and $j$. The geographical similarity is binary, having a value of 1 if the two firms are located in the same state.

## 5.3 Direct and indirect effect of R&D

In the previous subsections, we find evidence that R&D investment not only directly improves productivity but also increases the likelihood of forming R&D collaborations, $\Pr(g_{ijt} = 1)$. As discussed in section 3.3, the endogenous network formation framework enables a firm to enjoy additional efficiency gains from investing in R&D through better network formation and higher spillover levels. This subsection discusses the total R&D effect, including both *direct* and *indirect* effects given in equation (12) and (13). For counterfactual work, I assume that equilibrium beliefs are continuous on variables.

First, we will analyze the one-time effect of a 10 percent R&D increase on productivity. As given in Table 2, a 10 percent increase in R&D *directly* results in a 0.097 percent efficiency gain on average. To obtain the *indirect* effect of a 10 percent increase in R&D through network formation, I consider the following:

$$\mathrm{IE}_{it} \equiv \beta_3 \Big\{ E[\Sigma_{j\neq i} g_{ijt+1} t_{ij} r_{jt+1} | r^h_{it}, \cdot] - E[\Sigma_{j\neq i} g_{ijt+1} t_{ij} r_{jt+1} | r^o_{it}, \cdot] \Big\} \tag{23}$$

Table 4: The direct and indirect effects of a 10% in R&D

| | Direct effect | Indirect effect | |
|---|---|---|---|
| | | All firms | Firms in collabo |
| Mean | 0.097 | 0.012 | 0.022 |
| $1^{st}$ Quartile | 0.097 | 0.000 | 0.004 |
| $2^{nd}$ Quartile | 0.097 | 0.003 | 0.014 |
| $3^{rd}$ Quartile | 0.097 | 0.015 | 0.033 |
| Obs. | 8,903 | 8,903 | 4,302 |

*Note*: 'All firms' include firms with positive R&D investments and 'Firms in collabo' include firms with positive R&D and at least one collaboration.

where $\beta_3$ is the coefficient for the spillover effect on productivity and $r_{it}^h$ and $r_{it}^o$ are 10 percent increased R&D and the initial R&D, respectively. In equation (23), the first term is the expected spillover level under 10 percent increased R&D, and the second term is the expected spillover level under the initial R&D. Thus, it calculates the expected productivity gain from expanded R&D collaboration network and higher knowledge spillover levels.

In addition, based on each realization of the R&D collaboration network and previous R&D investment level, firms' next period R&D investment $r_{jt+1}$ could be adjusted as well. There are two conflicting incentives to adjust R&D investment upon forming collaborations. Collaborating firms might increase their R&D from conducting a new R&D project with collaborators, but at the same time, they might reduce R&D by enjoying their collaborative partners' resources. To account for those incentives, I use the approximated adjusted R&D to consider potential changes. To obtain the adjusted $r_{jt+1}$, I do not go through the precise dynamic decision rule for R&D investments. Instead, I approximate it from the data by deriving $E[r_{jt+1}|n_{jt+1}, \mathbf{X}_{jt+1}]$ where $n_{jt+1}$ is the number of collaborations firm $j$ have at $t+1$ and $\mathbf{X}_{jt+1}$ includes firm characteristics such as previous R&D investment, capital, labor, and sales and R&D user costs induced by tax credits to reflect the R&D decision rule. It implicitly presumes that the R&D investment level only depends on the number of collaborations, not on the characteristics of collaborators. I ran a linear regression with those variables and year-, industry-, and firm-fixed effects and used estimated parameters to obtain the adjusted R&D (See Appendix C).

The simulation algorithm for calculating the indirect effect is as follows. I first define $P_{ijt+1}^h$ and $P_{ijt+1}^o$ as conditional choice probabilities of forming a collaboration between $i$ and $j$ at $t+1$ with 10 percent increased R&D and the initial R&D, respectively.

1. Draw two R&D networks, $\mathbf{g}_{t+1}^h = \{g_{ijt+1}^h\}_{i,j}$ and $\mathbf{g}_{t+1}^o = \{g_{ijt+1}^o\}_{i,j}$, from $\mathbf{P}^h = \{P_{ijt+1}^h\}_{i,j}$

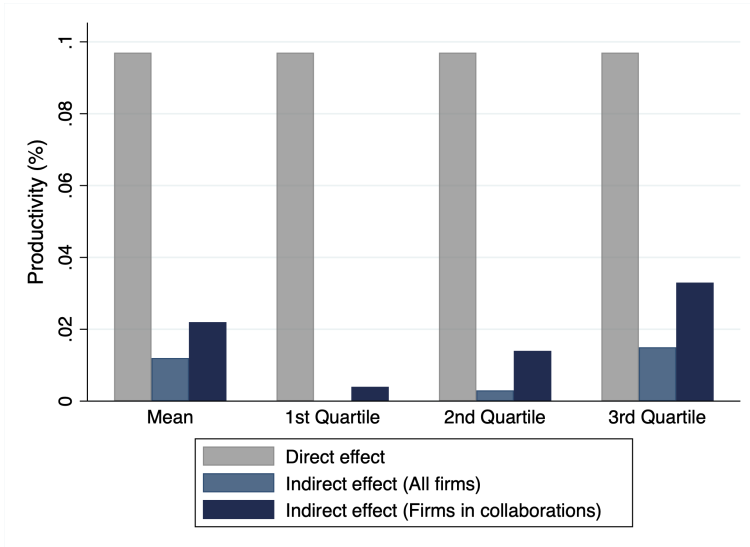



Figure 6: Impulse function of a 10% R&D increase shock

and $\mathbf{P}^o = \{P^o_{ijt+1}\}_{i,j}$, respectively

2. Under the drawn R&D networks, obtain the adjusted R&D investment levels, $r^h_{it+1}$ and $r^o_{it+1}$ for all $i$.

3. Calculate the indirect effect, $\beta_3 \left\{ \Sigma_{j \neq i} g^h_{ijt+1} t_{ij} r^h_{jt+1} - \Sigma_{j \neq i} g^o_{ijt+1} t_{ij} r^o_{jt+1} \right\}$

4. Iterate 1-3 $l$ times and obtain the average of the simulated indirect effects.

I use 5000 simulations and Table 4 summarizes the one-time effects of a 10 percent increase in R&D investment. Figure 6 describes the results in Table 4 to show its relative magnitude evidently. The indirect efficiency gain through collaborations varies from 0 to 0.015 percent within the 1st and 3rd quantiles of the sample with a mean of 0.012 percent. With the subsample of firms in collaborations, the effect tends to be larger because they are more likely to expand their networks. It implies that while the direct effect increases productivity by 0.097 percent, a firm can expect 0.012 (0.022 for firms in collaborations) additional efficiency gain through potential collaborations on average.

However, since both productivity and collaboration are time-persistent, their long-term effect could be more substantial. Figure 7 displays the impulse functions of a 10 percent R&D increase shock. The solid line is the direct effect of R&D in the canonical model. In the first period of shock, a firm directly benefits from investing more in R&D. The effect persists over time because productivity is time-persistent, but it decays as time goes on. However, if we consider endogenous network formation, a firm might enjoy additional efficiency gain in the following periods by expanding the collaboration network,

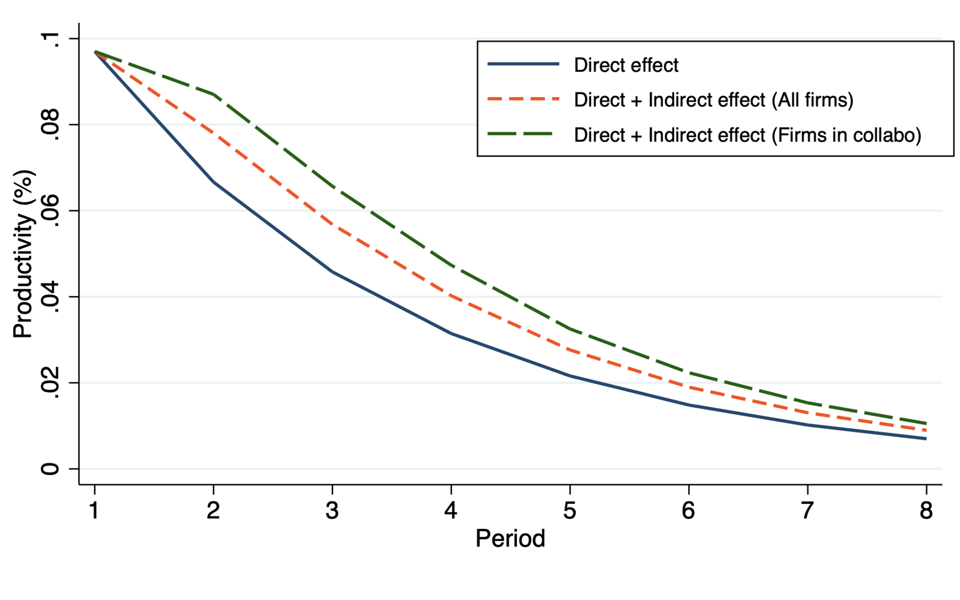



Figure 7: Impulse function of a 10% R&D increase shock

and therefore, R&D shock decays slowly. The short- and long-dashed lines represent the averaged total R&D effect in the endogenous network formation model averaged for all firms and firms in collaborations. Thus a long-term effect of a one-time 10 percent R&D increase shock is 0.31 percent in the canonical model with the direct effect but obtains 0.36 percent on average (0.40 if averaged for firms in collaborations) in the generalized model considering both the direct and indirect effects (i.e. For a median firm, by investing $147 million more in R&D, it would expect $165 million of output gain from the direct effect and $28 million of additional output gain from the indirect effect via expanding network.[18]). It indicates that the average effect of R&D is 16 percent underestimated in the canonical model. In the generalized model accounting for the endogenous network formation, the result suggests that R&D investment is more effective when firms collaborate in the economy, and a firm has more incentive to invest in R&D investment by considering the network effect.

## 5.4 Implications on R&D subsidy

Based on the estimation results, this subsection delves into the implications of R&D subsidy. The U.S. government has been providing R&D subsidies to firms to support R&D activities and promote productivity in the economy. Specifically, since the inception of the Economic Recovery Tax Act of 1981, firms elevating their R&D expenditures beyond

[18]It assumes the constant revenue level across periods.

a base level have been entitled to an R&D tax credit, aiming to incentivize further investment into R&D (Hall 1993). The effectiveness of an R&D tax credit in amplifying R&D investment has been explored in numerous studies. While its effectiveness varies across studies and data, in this subsection, I adopt a price elasticity of unity (i.e. a dollar-for-dollar increase in R&D spending from R&D tax credit) following Hall and Van Reenen (2000) and Bloom, Schankerman, and Van Reenen (2013). Then I analyze marginal private and social returns to R&D spending to give implications for R&D subsidy policies.

In the context of welfare, disentangling the R&D's effect into the *direct* and *indirect* effect through network formation is especially crucial to policymakers. The presence of collaborations amongst firms in the economy could cause market failure, given that firms do not account for external spillover effects through collaborations when determining R&D levels. This market failure is intensified if collaborations are endogenously determined, as fostering collaboration networks signifies an additional external benefit by potentially generating more future spillover effects in the economy. Thus, firms that actively collaborate or exhibit a higher propensity to collaborate may generate more social welfare than previously analyzed. This indicates that a targeted subsidy might offer greater benefits than a uniform subsidy.

First, one-time marginal private and social returns to R&D under the canonical model with exogenous network formation are given as follows:

$$\begin{aligned} \mathrm{MPR}^0_{it} &= \frac{Y_{it}}{R_{it}}\beta_2 \\ \mathrm{MSR}^0_{it} &= \mathrm{MPR}^0_{it} + \Sigma_{j\neq i} g_{ijt}\frac{Y_{jt}}{R_{it}}\beta_3 \end{aligned} \tag{24}$$

where $\beta_2$ and $\beta_3$ are coefficients for private R&D investment and spillovers from collaborating firms, respectively, and $Y_{it}$ and $R_{it}$ are sales and R&D expenditures in dollars. MPR considers the direct effect of firm $i$'s R&D on output. MSR includes MPR and the effect on the outputs of other firms diffusing through spillovers.

With the generalized model where we consider the *indirect* effect of R&D through endogenous network formation together, one-time expected marginal private and social returns are revised as follows:

$$\begin{aligned} \mathrm{MPR}^1_{it} &= \frac{Y_{it}}{R_{it}}\beta_2 + \frac{Y_{it+1}}{R_{it}}\cdot\frac{\partial S_{it+1}}{\partial r_{it}}\beta_3 \\ \mathrm{MSR}^1_{it} &= \mathrm{MPR}^1_{it} + \Sigma_{j\neq i} g_{ijt}\frac{Y_{jt}}{R_{it}}\beta_3 + \Sigma_{j\neq i}\frac{Y_{jt+1}}{R_{it}}\cdot\frac{\partial S_{jt+1}}{\partial r_{it}}\beta_3 \end{aligned} \tag{25}$$

In addition to the components given in equation (24), we observe additional gains through

Table 5: Private and social returns to R&D

| Group of firms | Canonical model (1) MPR (%) | (2) MSR (%) | (3) Wedge Percent Points | General framework (4) MPR (%) | (5) MSR (%) | (6) Wedge Percent Points | (7) Med Emp (1,000) | (8) Med R&D ($100m) |
|---|---|---|---|---|---|---|---|---|
| 1. All firms in collabo | 25.9 | 65.0 | 39.1 | 32.8 | 94.7 | 61.9 | 10.5 | 54.2 |
| 2. #Collabo ≤ 1 | 41.2 | 60.4 | 19.2 | 43.7 | 91.4 | 47.7 | 3.5 | 14.6 |
| 3. 1<#Collabo ≤ 3 | 31.7 | 77.5 | 45.8 | 37.3 | 114.4 | 77.1 | 5.7 | 30.0 |
| 4. 3<#Collabo ≤ 9 | 27.4 | 87.0 | 59.6 | 36.8 | 132.1 | 95.3 | 14.0 | 82.7 |
| 5. 9<#Collabo | 16.7 | 56.6 | 39.9 | 25.0 | 74.9 | 49.9 | 37.8 | 382.6 |

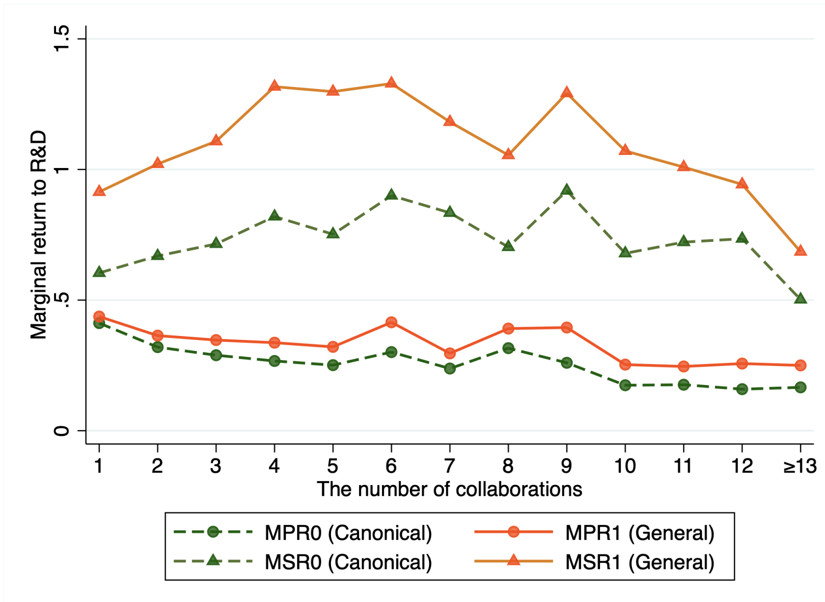
*Notes*: This table presents median values of the rates of return to R&D of each group of firms in collaborations according to equation (24) and (25) for two scenarios: (1) canonical model with exogenous network formation and (2) general framework with endogenous network formation. Column (3) and (6) contains the absolute difference between MPR and MSR under the canonical model and general framework, respectively. Columns (7) and (8) present the medians of employees in thousand and R&D investments in $100 million. Each group of firms with a different number of collaborations is based on the 1st, 2nd, and 3rd quartiles of the number of collaborations (1, 3, and 9 collaborations, respectively).

expanding the collaboration network and increasing spillovers in the next period in both MPR and MSR.

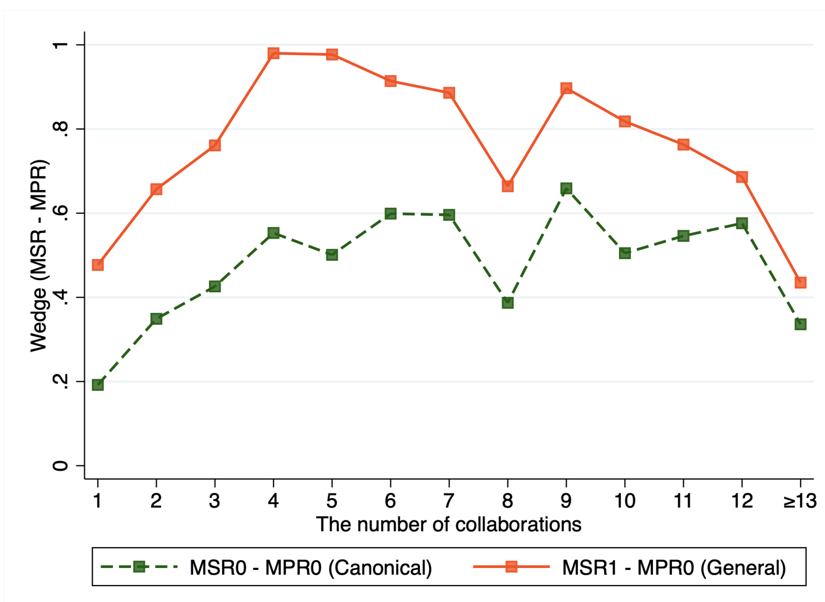
In deriving MPR and MSR, I cannot directly obtain the marginal effect from the derivation because the collaboration status $g$ is not differentiable. Instead, I obtain the private and social return from a 10 percent increase in R&D as done in section 5.3 and divide them by the amount of the increase in R&D.

Table 5 displays private and social returns to R&D under the canonical and general frameworks. Columns (1), (2), (4), and (5) represent median values of private and social returns to R&D in percentage for each firm group in collaboration. Columns (3) and (6) show the absolute difference between private and social returns. In the canonical model, MPR and MSR are 25.9% and 65.1%, respectively, when pooling all firms in collaborations together (i.e., a $1 increase in a firm's R&D expenditure increases its own output by $0.259 and total output by $0.651). The wedge percentage point between MPR and MSR is 39.2, which mirrors previous results in the literature. When considering endogenous network formation in the general framework, the median MPR and MSR rise to 32.8% and 94.7%, respectively, indicating a wider 61.9 percentage point wedge.

In rows 2-5 of Table 5, I categorize firms by the number of collaborations based on their 1st, 2nd, and 3rd quartiles. Within each model, MPR decreases as the number of collaborators increases because the overall firm size and R&D investments are larger in firms with many collaborators. However, MSR tends to amplify even as marginal values typically decline with escalating R&D investment because firms generate more




(a) Median MSRs and MPRs




(b) Wedges between the median MSRs and MPRs

Figure 8: Median MSRs and MPRs and their wedges (MSR - MPR) across firms with different number of collaborations

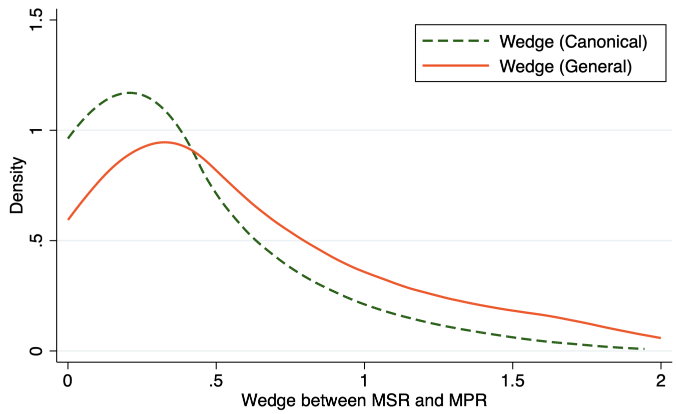
social benefits through more collaborators. However, this tendency recedes for firms with over 9 collaborations, as described in row 5, because large firms tend to collaborate with smaller firms to avoid potential competition, whereby the decline in marginal values supersedes the advantages of additional collaborations.

Compared to the canonical model with the exogenous network, once we allow the endogenous formation of R&D collaborations, MPR and MSR increase due to the additional gains from expanding the network in all groups. However, this additional gain is more evident in MSR; thus, the wedges between the median MPRs and MSRs, which need to be subsidized, are higher in the general model than in the canonical model–increased by 22.8 percent points. This increase is particularly pronounced for firms with fewer collaborators, as seen in rows 2 and 3. The wedges for those firm groups have more than or almost doubled. Although the wedge increase between the two models remains positive, it is less discernible for firms with over nine collaborators. This is because larger firms with already many collaborators have less incentive to expand their collaboration network or tend to collaborate with smaller firms to avoid potential competition by increasing their R&D, resulting in minimal additional social gains. On the other hand, by increasing R&D, relatively smaller firms with few collaborators are more prone to expand their R&D network with larger firms, which enhances social gains.

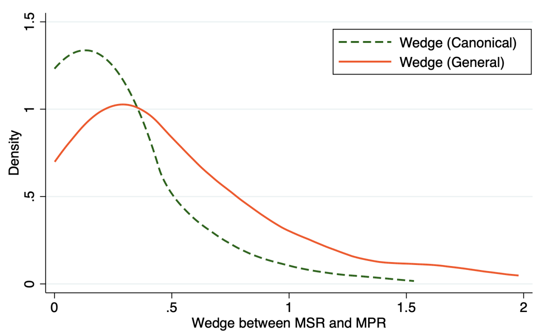
Figure 8 shows the depicted features prominently. Figure 8a displays the median MPRs and MSRs across the groups of firms with each number of collaborations. The lines with circles represent MPRs, those with triangles indicate MSRs, and dashed and solid lines denote canonical and general models, respectively. In the canonical model, the peak median MSR is in the 9-collaboration group; contrastingly, the general model suggests maximal MSRs in firms with 4-6 collaborations. The same pattern is observed in Figure 8b of the wedges between the median MSRs and MPRs. In addition, the increments of the median wedges for firms with 1-3 collaborators are also noticeable, resulting in larger wedges than those above 12 collaborators. To see the overall wedges across firms, Figure 9 displays the kernel density distributions of the wedges.[19] The distributions are right-shifted and flatter in the general model, indicating higher and more varied wedges. This pattern is more pronounced in firms with fewer collaborators, reaffirming previous median-value results.

This result suggests the following. There are two types of R&D subsidy: the uniform subsidy, such as R&D tax credit, and the targeted subsidy, such as the Advanced Technology Program (ATP) in the U.S. Especially the ATP puts more weight on collaborating firms. The results in this section imply that R&D investments from firms in collaborations

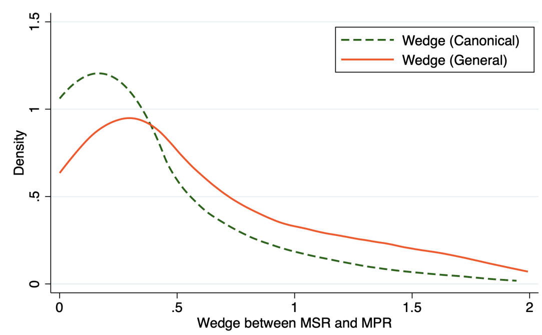
[19] To avoid simulation errors, I used the sample between the first and third quartiles.



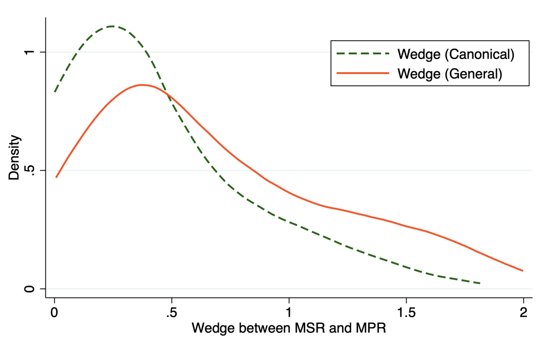

(a) All firms in collaborations



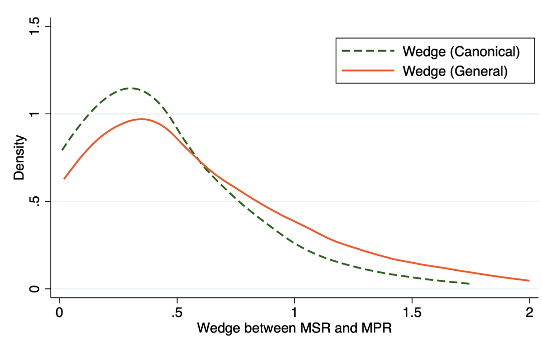

(b) #Collaborations = 1




(c) 1< #Collaborations ≤ 3




(d) 3 < #Collaborations ≤ 9




(e) 9 < #Collaborations

Figure 9: The distributions of the wedges (MSR - MPR) of each group of firms with different numbers of collaborators

inadvertently generate social benefits through networks and possible network expansion, which is not accounted for by firms' own decisions and leads to market failure. It supports the importance of the targeted subsidy program for collaborating firms, as in Konig, Liu, and Zenou (2019). In the canonical model, the targeted subsidy ranking favors large firms as they are better connected in the R&D network. However, in addition to that, this paper suggests that smaller firms with fewer R&D networks should also be considered subsidized since they are more prone to contribute to expanding the R&D networks with larger firms and, therefore, have high marginal social effects.

# 6 Concluding Remarks

A significant portion of the empirical literature on innovation and productivity emphasizes measuring private returns to R&D investment, typically associated with effects through innovation or cost reduction. Concurrently, a more recent subset of literature has concentrated on estimating the effects of spillovers from firm interactions. This paper bridges these two literature facets by incorporating the decision on R&D collaboration into a firm's dynamic model.

My research examines the R&D investment's additional role in collaboration network formation and discusses the direct and indirect effects of R&D accounting for endogenous network formation. It posits that firms, through strategic R&D collaboration, anticipate additional collaborations and knowledge spillovers from their R&D investments. To disentangle the direct and indirect effects, my model captures two key aspects. First, productivity is improved by both private R&D investment and spillovers from collaborations. Second, a firm makes dynamic decisions on collaborators based on the expected gain from collaborations. These features open an additional channel for R&D to improve productivity through networking.

Utilizing U.S. firms' R&D collaboration data from 1981-2001, this research quantifies R&D effects. The estimation results provide that a 10 percent increase in R&D directly improves productivity by 0.31 percent and indirectly by 0.05 through better collaborations in the long run. It suggests that if we ignore endogenous network formation, the effect of private R&D is underestimated by 16% on average. These results underline that firms have more incentive to invest in R&D because they might enjoy more collaborations and knowledge spillovers in the future. Policymakers are advised to subsidize R&D investments for collaborative firms due to the social benefits produced through current and potential networks. Also, I especially focus on smaller firms with fewer collaborations, which can generate substantial social advantages by expanding R&D networks with larger entities.

The idea suggested in this paper can be applied to other empirical settings. For example, the co-working network among researchers or students in schools might be endogenously determined and affected by the invested effort.

## References

Ackerberg, D., K. Caves, and G. Frazer. 2015. "Identification Properties of Recent Production Function Estimators." *Econometrica* 83 (6): 2411–2451.

Aguirregabiria, V., and C. Ho. 2012. "A Dynamic Oligopoly Game of the US Airline Industry: Estimation and Policy Experiments." *Journal of Econometrics* 168 (1): 156–173.

Arque-Castells, P., and D. F. Spulber. 2022. "Measuring the Private and Social Returns to R&D: Unintended Spillovers versus Technology Markets." *Journal of Political Economy* 130 (7): 1860–1918.

Auerbach, Eric. 2022. "Identification and Estimation of a Partially Linear Regression Model Using Network Data." *Econometrica* 90 (1): 347–365.

Badillo, E. R., and R. Moreno. 2016. "Are Collaborative Agreements in Innovation Activities Persistent at the Firm Level? Empirical Evidence for the Spanish Case." *Review of Industrial Organizaion* 49:71–2916.

Bjornerstedt, J., and J. Stennek. 2007. "Bilateral oligopoly – The efficiency of intermediate goods markets." *International Journal of Industrial Organization* 25 (5): 884–907.

Bloch, F., and M. O. Jackson. 2007. "The Formation of Networks with Transfers among Players." *Journal of Economic Theory* 133 (1): 83–110.

Bloom, N., M. Schankerman, and J. Van Reenen. 2013. "Identifying Technology Spillovers and Product Market Rivalry." *Econometrica* 81 (4): 1347–1393.

Cohen, W. M., and D. A. Levinthal. 1989. "Innovation and Learning: The Two Faces of R&D." *The Economic Journal* 99:569–596.

Collard-Wexler, A., G. Gowrisankaran, and R.S. Lee. 2019. ""Nash-in-Nash" Bargaining: A Microfoundation for Applied Work." *Journal of Political Economy* 127 (1): 1–463.

Comola, M., and A. Dekel. 2022. "Estimating Network Externalities in Undirected Link Formation Games." *Working paper.*

De Loecker, J. 2013. "Detecting Learning by Exporting." *American Economic Journal: Microeconomics* 5 (3): 1–21.

de Paula, Á. 2020. "Econometric Models of Network Formation." *Annual Review of Economics* 12:775–799.

Doraszelski, U., and J. Jaumandreu. 2013. "R&D and Productivity: Estimating Endogenous Productivity." *The Review of Economic Studies* 80:1338–1383.

Hagedoorn, J. 2002. "Inter-firm R&D Partnerships: an Overview of Major Trends and Patterns Since 1960." *Research Policy* 31 (4): 477–492.

Hall, B. H. 1993. "R&D Tax Policy during the 1980s: Success or Failure?" *Tax Policy and the Economy* 7:1–35.

Hall, B. H., and J. Van Reenen. 2000. "How Effective Are Fiscal Incentives for R&D? A Review of the Evidence." *Research Policy* 29:449–469.

Hall, B.H., J. Mairesse, and P. Mohnen. 2010. "Measuring the Returns to R&D." *Handbook of the Economics of Innovation* Vol. 2, Chapter 22. North-Holland, Amsterdam: Elsevier.

Hanaki, N., R. Nakajima, and Y. Ogura. 2010. "The Dynamics of R&D Networks in the IT Industry." *Research Policy* 39:386–399.

Hendel, I., and I. Nevo. 2006. "Measuring the Implications of Sales and Consumer Inventory Behavior." *Econometrica* 74 (6): 1637–1673.

Hernán, R., P. L. Marín, and G. Siotis. 2003. "An empirical evaluation of the determinants of Research Joint Venture Formation." *The Journal of Industrial Economics* 51 (1): 75–89.

Hsieh, C., M. D. König, and X. Liu. 2022. "Endogenous Technology Spillovers in Dynamic R&D Networks." *Working Paper.*

Ifrach, B., and G. Y. Weintraub. 2017. "A Framework for Dynamic Oligopoly in Concentrated Industries." *The Review of Economic Studies* 84 (3): 1106–1150.

Iyoha, E. 2020. "Estimating Productivity in the Presence of Spillovers: Firm-level Evidence from the US Production Network." *working paper.*

Johnsson, Ida, and Hyungsik Roger Moon. 2021. "Estimation of Peer Effects in Endogenous Social Networks: Control Function Approach." *Review of Economics & Statistics* 103 (2): 328–345.

Konig, M. D., X. Liu, and Y. Zenou. 2019. "R&D Networks: Theory, Empirics, and Policy Implications." *Review of Economics and Statistics* 101 (3): 476–491.

Leung, M. P. 2015. “Two-step Estimation of Network-formation Models with Incomplete Information.” *Journal of Econometrics* 188:182–195.

Lychagin, S., J. Pinkse, M. E. Slade, and J. Van Reenen. 2016. “Spillovers in Space: Does Geography Matter?” *Journal of Industrial Economics* 64 (2): 295–335.

Malikov, E., and S. Zhao. 2021. “On the Estimation of Cross-Firm Productivity Spillovers with an Application to FDI.” *The Review of Economic Studies,* 1–46.

Olley, G.S., and A. Pakes. 1996. “The Dynamics of Productivity in the Telecommunications Equipment Industry.” *Econometrica* 64 (6): 1263–1297.

Peters, B., M. J. Roberts, V. A. Vuong, and H. Fryges. 2017. “Estimating Dynamic R&D Choice: an Analysis of Costs and Long-run Benefits.” *The RAND Journal of Economics* 48 (2): 409–437.

Schilling, M. 2009. “Understanding the Alliance Data.” *Strategic Management Journal* 30 (3): 233–260.

Wooldridge, J. M. 2015. “Control Function Methods in Applied Econometrics.” *Journal of Human Resources* 50 (2): 420–445.

Zacchia, P. 2020. “Knowledge Spillovers Through Networks of Scientists.” *Review of Economic Studies* 87 (4): 1989–2018.

# A Data

## A.1 Number of collaborations a firm has

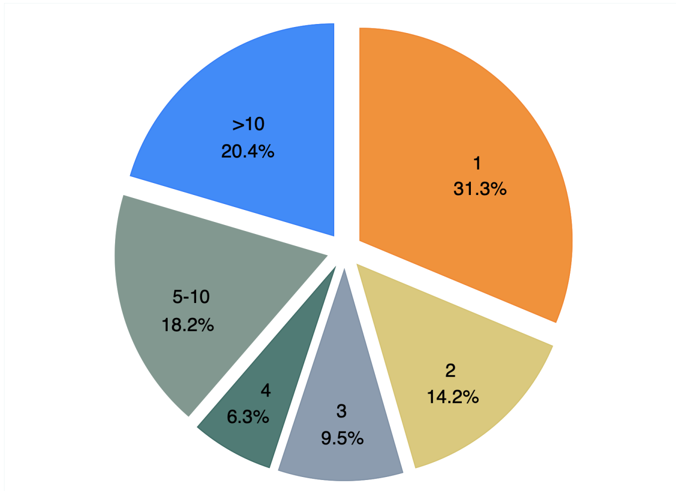



Figure 10: Shares of the number of collaborations a firm has

## A.2 Summary of the industries collaborating firms operate

To view the collaboration patterns, Table 6 displays the top five industries where collaborating firms are most commonly found. The second column lists the percentage of total collaborations in which each industry is involved, while the last column shows the shares of within-industry collaborations (i.e., firms in a collaboration operating in the same sector) among all collaborations. Most firms in collaborations operate in 'Computer and electronic products', 'Chemical products', 'Machinery', 'Motor vehicles, bodies and trailers, and parts', and 'Computer systems design and related services', covering up to 72% of total collaborations. But we observe more across-industry firm pairs than within-industry firm pairs. About 67% of collaborations in the sample are across-industry pairs, and most within-industry collaborations are found especially in the sector of 'Computer and electronic products' and 'Chemical products'.

## A.3 Preliminary reduced form evidence

Before introducing the structural model, I provide a preliminary evidence of the effect of R&D investments on the number of collaborations to motivate the necessity of the theoretical framework of a firm's dynamic model incorporating network formation. It

Table 6: Collaboration patterns across industries

| | Total share | Within-industry |
|---|---|---|
| Computer and electronic products | 31% | 17% |
| Chemical products | 23% | 11% |
| Machinery | 7% | 1% |
| Motor vehicles, bodies and trailers, and parts | 6% | 1% |
| Computer systems design and related services | 5% | <1% |

The second column is the percentage of collaborations in which either one of two firms in a collaboration is in each industry and the third column is the percentage of total collaborations in which both firms in a collaboration operates in the same industry.

has been theoretically and empirically supported that R&D investments or R&D intensity have a positive effect on forming more R&D collaborations in numerous papers using various data sets. Those studies mostly focus on simple linear or binary choice regressions using R&D investments as one of the key independent variables. However, the results from those studies might be diluted with unobserved heterogeneity in the tendency of firms with high or low R&D seeking collaborators and simultaneity problem that future collaborations might already have been known or determined. In this section, I provide more robust regression evidence exploiting an instrument for R&D investments.

To see their correlation more clearly, I ran the regression using the following equation:

$$\#\text{Collaborations}_{it} = \beta_0 + \beta_r \ln \text{R\&D}_{it-1} + \mathbf{X}_{it-1}\beta + \eta_t + \eta_m + \eta_i + \epsilon_{it} \tag{26}$$

where $\mathbf{X}_{it}$ includes firm characteristics such as labor, capital, and productivity[20] levels in logs, and $\eta_t, \eta_m, \eta_i$ are year-, industry-, and firm-fixed effects, respectively. The coefficient of the lagged R&D investments, $\beta_r$, is expected to capture the effect of R&D to foster the formation of R&D collaborations. However, it might have endogeneity probme as suggested. To address the endogeneity problem, I use R&D user cost induced by tax credit to instrument R&D investment levels as described in section 4.2 In this estimation, state tax credit and federal tax credit components of R&D user cost are used to instrument for R&D expenditures. The rationale for this identification strategy is that lagged R&D tax credit directly affects R&D investments but only indirectly affects the formation of collaborations.

Table 7 shows the regression results. Columns (1) and (2) are OLS and IVE results of the estimation of equation (26), respectively. In both results, lagged R&D investments have positive effects on the number of collaborations, but the effect becomes larger if I

[20]I used productivity level obtained from the structural estimation. See section 4.1 for the estimation strategy.

Table 7: Regression results on the number of collaborations

| Dependent var: # Collaborations$_{it}$ | | | |
|---|---|---|---|
| | (1) | (2) | (3) |
| | OLS | IVE | IVE |
| ln R&D$_{it-1}$ | 0.732$^{**}$ | 4.251$^{**}$ | 5.427$^{**}$ |
| | (0.182) | 1.183 | (2.486) |
| ln R&D$_{it}$ | | | −0.789 |
| | | | (2.222) |
| Productivity$_{it-1}$ | 14.674$^{**}$ | 9.696$^{**}$ | 9.072$^{**}$ |
| | (2.296) | (0.234) | (1.739) |
| ln Capital$_{it-1}$ | 2.114$^{**}$ | 0.901$^{**}$ | 0.716 |
| | (0.339) | (0.423) | (0.445) |
| ln Employees$_{it-1}$ | −1.763$^{**}$ | −3.003$^{**}$ | −3.113$^{**}$ |
| | (0.379) | (0.438) | (0.449) |
| First stage estimation | | | |
| (Federal) R&D user cost$_{it-1}$ | | −0.351$^{**}$ | −0.355$^{**}$ |
| | | (0.074) | (0.076) |
| (State) R&D user cost$_{it-1}$ | | −1.460$^{**}$ | −0.672 |
| | | (0.294) | (0.539) |
| (Federal) R&D user cost$_{it}$ | | | 0.011 |
| | | | (0.076) |
| (State) R&D user cost$_{it}$ | | | −0.923$^{*}$ |
| | | | (0.530) |
| Observations | | 12,695 | |

All specifications are controlled for year-, industry-, and firm-fixed effects. *Significant at 0.05% level, **Significant at 0.01% level.

use the instruments that have high explanatory power in the first stage. It implies that either firms with lower R&D tend to seek collaborators or future known collaborations reduce R&D investments. Without this endogeneity, the IVE result in column (2) suggests that lagged R&D increases the number of future collaborations–A 10 percent increase in R&D contributes to 0.4 more collaborators on average. In addition, in column (3), I add the same-period R&D as an additional variable to support the timing of forming R&D collaborations. The number of R&D collaborations depends on the lagged R&D, not the same-period R&D, as the effect of the same-period R&D is insignificant. It implies that the decision for the network formation relies on the lagged R&D investments, supporting that the formation takes time to be effective.

These results suggest that R&D investments boost future collaborations. It motivates us to study the additional effect of R&D through the formation of a collaboration network. In addition, since the decision for R&D collaborations is a firm's strategic game considering its dynamic benefit, it requires the micro-foundation to find the more precise effect of R&D on forming collaborations. The next session provides the theoretical framework for estimating the effect of R&D, including the strategic game of collaboration network formation.

# B R&D investment decision

In this framework, the firm uses R&D investment to buy improvement in expected future productivity. Notably, other firms' R&D investments also matter through spillovers. Firms therefore strategically choose their own R&D levels, based on the given R&D collaboration network determined in the previous period.

After the short-term decisions are made, firm-specific productivity levels, $\omega_t = (\omega_{1t}, ..., \omega_{Nt})$, are revealed. A firm has a value function at this stage, $V^r$, as follows and faces an R&D investment decision problem:

$$V_i^r(\mathbf{s}_{it}^r, \xi_{it}) = \pi(\omega_{it}, k_{it}) + \max_{r_{it}}\{E_t[\rho V_i^r(\mathbf{s}_{it+1}^r, \xi_{it+1})|\mathbf{s}_{it}^r, r_{it}] - C_r(r_{it}, \xi_{it})\} \tag{27}$$

where $\mathbf{s}_{it}^r \equiv (\omega_t, \mathbf{g}_t, \mathbf{x}_t, k_{it})'$ is a vector of state variables, $\mathbf{g}_t = [g_{ijt}]_{i,j}$ is a network matrix, $\mathbf{x}_t = [x_{ijt}]_{i,j}$ is a networking cost matrix, which will be explained in the following subsection, and $\rho$ is a discount factor. The last term $C_r(r_{it}, \xi_{it})$ represents the cost of implementing R&D where $C_r(\cdot)$ is a cost function and $u_{it}$ captures the randomness in cost variation across firms. The cost shock $u_{it}$ is private and independent across firms and time. It is therefore an incomplete game of firms. I assume that there exists a Markov

Table 8: Production function estimation

| Dependent var: $i_{it}$ | | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|---|
| $\omega_{it}$ | | 6.787 | 9.4470 | 22.3100 | 6.9592 | 10.9673 | 43.1362 |
| | | (0.0082) | (0.1569) | (2.4783) | (0.0155) | (0.2549) | (3.664) |
| $\omega_{it}^2$ | | | −0.1480 | −1.5516 | | −0.2290 | −4.0419 |
| | | | (0.0087) | (0.2700) | | (0.0145) | (0.4082) |
| $\omega_{it}^3$ | | | | 0.0510 | | | 0.1415 |
| | | | | (0.0098) | | | (0.0151) |
| $l_{it}$ | | 0.4044 | 0.4032 | 0.4033 | 0.5224 | 0.5177 | 0.5166 |
| | | (0.0017) | (0.0017) | (0.0017) | (0.0044) | (0.0043) | (0.0043) |
| $k_{it}$ | | 0.6021 | 0.6044 | 0.6043 | 0.4518 | 0.4593 | 0.4611 |
| | | (0.0015) | (0.0015) | (0.0015) | (0.0037) | (0.0037) | (0.0037) |
| $r_{it}$ | | −0.1063 | −0.1053 | −0.1052 | −0.1435 | −0.1397 | −0.1384 |
| | | (0.0009) | (0.0009) | (0.0009) | (0.0024) | (0.0024) | (0.0024) |
| $S_{it}$ | R&D spillover | 0.0270 | 0.0261 | 0.0262 | −0.0442 | −0.0399 | −0.0405 |
| | | (0.0001) | (0.0001) | (0.0001) | (0.0003) | (0.0004) | (0.0004) |
| | $\omega$ spillover | | | | −0.0270 | −0.0271 | −0.0267 |
| | | | | | (0.0005) | (0.0005) | (0.0005) |

All specifications are controlled for year- and industry-fixed effects, and within-group fixed effect regression is used. *Significant at 0.05% level, **Significant at 0.01% level.

Perfect Equilibrium that determines the optimal R&D levels. In the equilibrium, a firm's optimal strategy only depends on $(\mathbf{s}_{it}^r, \xi_{it})$ and takes other firms' strategies as given.

# C Strict monotonicity assumption in ACF

We would check the *ex-post* validity of strict monotonicity assumption in ACF in this section. To apply ACF method, we need the demand function for capital investment, $\mathscr{I}(k_{it}, l_{it}, \omega_{it}, r_{it}, S_{it})$, to be strictly monotonic in productivity. I first estimate the firm-specific productivity using GMM of the equation (18) under the assumption. Then I would check if the obtained productivity actually satisfies the strict monotonicity. I run the following reduced form estimation:

$$i_{it} = b_0 + b_1\omega_{it} + b_2\omega_{it}^2 + b_3\omega_{it}^3 + b\mathbf{X}_{it} + \xi_{it} \tag{28}$$

where $X_{it}$ includes $l_{it}$, $k_{it}$, $r_{it}$, and $S_{it}$. I also added year and industry fixed effects. In addition, to account for firm-fixed effects, within-group fixed effect estimation is used for regression.

Table 8 displays the regression results. Columns (1)-(3) are the results with only

Table 9: Estimation of R&D adjustment

| Dependent var: $\ln \text{R\&D}_{it}$ | (1) | (2) |
|---|---|---|
| $\ln \#\text{Collaborations}_{it}$ | 0.014** | 0.017 |
| | (0.005) | (0.010) |
| $\ln \text{R\&D}_{it-1}$ | 0.972** | 0.793** |
| | (0.003) | (0.015) |
| $\text{Productivity}_{it}$ | 0.231** | 0.063** |
| | (0.023) | (0.031) |
| $\ln \text{Capital}_{it}$ | −0.045** | −0.080** |
| | (0.004) | (0.011) |
| $\ln \text{Labor}_{it}$ | 0.025** | 0.119** |
| | (0.006) | (0.015) |
| $\ln \text{Sale}_{it}$ | 0.034** | 0.122** |
| | (0.006) | (0.014) |
| $\text{R\&D user cost}_{it}$ | −0.308** | −0.263** |
| | (0.032) | (0.037) |
| Firm-fixed effect | N | Y |
| Observations | 12,713 | |

All specifications are controlled for year- and industry-fixed effects. *Significant at 0.05% level, **Significant at 0.01% level.

R&D spillovers, and columns (4)-(6) are the ones with productivity spillovers as well. Considering that productivity $\omega$ spans from 7.9 to 10.5 in this sample of data, the first derivatives of capital investment $i_{it}$ with respect to productivity $\omega_{it}$ are always positive in all specifications. It implies productivity monotonically increases capital investment conditional on other variables. This result *ex-post* supports the strict monotonicity in productivity.

# D The adjustment of R&D after forming a link

In this section of the Appendix, I discuss how a firm's decision on R&D investment is adjusted upon forming collaboration. I simply approximate $E[r_{it+1}|n_{it+1}, \mathbf{X}_{it+1}]$ where $\mathbf{X}_{it}$ includes firm characteristics such as previous R&D investment, productivity, capital, and R&D user cost induced by R&D tax credit. I ran a linear regression with those variables and year-, industry-, and firm-fixed effects.

Table 9 shows the regression results. The effect of the number of collaborations is significant in column (1) without firm-fixed effects, but it becomes insignificant in column

Table 10: Estimation of the number of collaborations

| Dependent var: | $\ln \#\text{Collaborations}_{it}$ |
|---|---|
| $\ln R\&D_{it}$ | $-0.027$ |
| | (0.251) |
| $\ln R\&D_{it-1}$ | $0.661^{**}$ |
| | (0.303) |

All specifications are controlled for year-, industry-, and firm-fixed effects. *Significant at 0.05% level, **Significant at 0.01% level.

(2) with firm-fixed effects. As described in section 5.2, two conflicting incentives cancel out the effects. However, there could be a simultaneity problem if the same-period R&D investments affect the formation of collaboration. To check its potential effect, Table 10 shows the result of IV regression of the logged number of collaborations on current and previous R&D investments and other firm variables, such as lagged capital, labor, productivity, and sales, using R&D user cost as instruments. The coefficients for R&D investments are the effect of R&D on forming collaborations, and the estimated results suggest that the number of collaborations is explained by the previous R&D, not the current period R&D. Therefore, there is no evidence of the simultaneity problem in R&D investment in Table 9.