Of Rivalry and Sinergy:

The Role of Patents and Products Similarity on R&D Investment

Dario Marino

University of Chicago
Department of Economics

March. 2025

Introduction

- This paper focuses on the interaction of product and technology similarities between firms.
- My claim is that technology similarity between two firms helps R&D investments when product similarity is low, because firms can absorb part of the improvements of other firms and they can invest without helping competitors.
- In contrast product similarity between firms reduces R&D investment when technology is similar, because the improvement is also absorbed by competitors.
- Therefore, it would be necessary to grant patent protection only when product similarity between the assignee and the firms in its technology network is high.

Introduction

- Synergy between firms is important. In (Akcigit, Sina, 2023) we saw that the effect on US business dynamism of lower knowledge diffusion is estimated to be between 70 and 50 percent depending on the economic dimension.
- I will start by outlining a Cournot model to show how firms competing on quantity choose an optimal level of R&D.
- The optimal decision for each firm depends on patent protection and good similarity, because each firm uses every other firms' technological improvements weighted by their technology similarity.
- We'll then simulate this model and then move to the empirical part.

Introduction

- We have different empirical aims. First we need a confirmation of a positive effect of tech similarity and a negative effect of product similarity on R&D.
- Then we should also see what dominates on R&D decision when the two similarity dimensions interact. A positive value would mean that in their interaction the positive effect of sinergy between technologies dominates over the negative effect of product rivalry.
- We would also need to "slice" our network to see if at similar level of technology similarity R&d goes down when product similarity is higher.
- Knowing from (Aghion et al. 2005) that the effect of competition on R&D is non-linear, I could extend the model to observe similar relationships in our case.

Model

- In this model firms compete on quantity while simultaneously investing in technology to reduce marginal costs.
- These improvements can be transmitted between firms depending on the level of patent protection and the similarity between technologies.
- Let $i \in \mathcal{I} = \{1, \ldots, N\}$ index firms, each producing a differentiated product. Denote by q_i the quantity produced and by p_i the corresponding price. The inverse demand function is given by:

$$p_i = A - q_i - \sum_{j \neq i} \delta_{ij} \, q_j,$$

where $A > 0$ is a market size parameter and $\delta_{ij} \geq 0$ measures substitutability between products i and j.

Model

- Each firm i is endowed with a technology set $\mathcal{T}_i \subseteq \mathcal{T}$. For each technology $t \in \mathcal{T}_i$, the firm invests z_{it}. The total technology investment is defined as

$$z_i = \sum_{t \in \mathcal{T}_i} z_{it}.$$

Technological overlap between firms i and j is captured by ω_{ij}, with $\omega_{ii} = \frac{1}{\phi}$. Where ϕ is a measure of patent protection which reduces the sinergy between firms but not for the original firm with respect to its own technologies.

Model

- Firms experience a cost reduction from technology spillovers. The effective marginal cost for firm i is

$$MC_i = c_i - \phi \sum_{j \in \mathcal{I}} \omega_{ij} z_j ,$$

where c_i is the baseline marginal cost and ϕ represents the patent breadth (externality parameter).

- The profit function is

$$\pi_i = p_i q_i - MC_i q_i - \frac{\kappa}{2} z_i^2 ,$$

with $\kappa > 0$ capturing the convexity of technology investment costs.

Model

- Firms choose (q_i, z_i) to maximize their profits under Cournot competition. The regulator chooses ϕ to maximize overall welfare,

$$W = CS + \sum_{i \in \mathcal{I}} \pi_i,$$

where CS denotes consumer surplus.

Model

- We simulate the model with N=1000 firms with technology and product similarity normally distributed around 0.5. Because we have $N = 1000$ firms, we do not solve the system analytically. Instead, we use an *iterative best-response* method:
    1. **Quantity best-response:** Holding z_j (and other q_j) fixed, firm i's first-order condition w.r.t. q_i leads to:

$$\frac{\partial \pi_i}{\partial q_i} = 0 \quad \Longrightarrow \quad q_i = \frac{1}{2}\left[\alpha - c + \phi \sum_j \omega_{ij} z_j - \sum_{j \neq i} \delta_{ij} q_j\right].$$

    2. **R&D best-response:** Holding q_i (and other z_j) fixed, the first-order condition w.r.t. z_i gives:

$$z_i = \max\left\{0, \frac{\phi\,\omega_{ii}}{\kappa} q_i\right\}.$$

If $\omega_{ii} = 1$, that simplifies to $z_i = \frac{\phi}{\kappa} q_i$.

Model

- We update all z_i first (given the current q_i), then update all q_i (given the newly updated z_i), repeating until convergence or until a fixed number of iterations.
- After we converge to (q^*, z^*), we approximate *consumer surplus* by

$$CS \approx \sum_{i=1}^{N} \frac{1}{2} q_i^* \left(p_i^*\right) \quad \text{where} \quad p_i^* = \alpha - q_i^* - \sum_{j \neq i} \delta_{ij} q_j^* .$$

We define *total welfare* as

$$W(\phi) = \sum_{i=1}^{N} \pi_i + CS.$$

We sweep over ϕ from 0 to 1 in steps of 0.01, compute the equilibrium, then record consumer surplus and total welfare to see how they vary.

Model

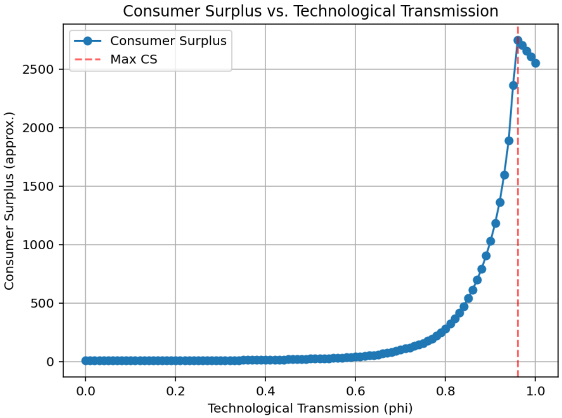


Figure 1: Consumer Surplus at different levels of Technological Transmission

Model

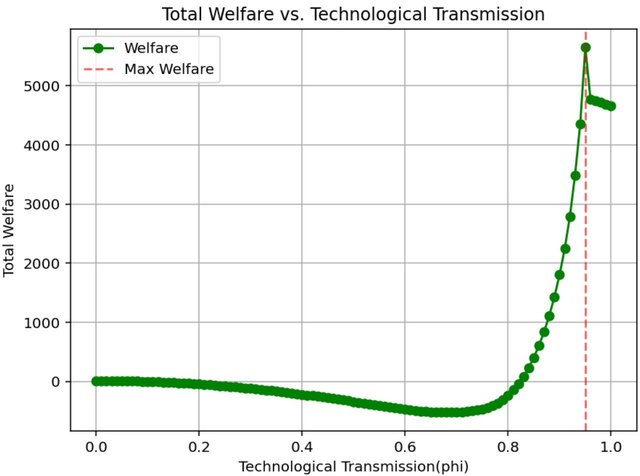


Figure 2: Total Welfare at different levels of Technological Transmission

Model

- A higher ϕ means that technologies are more easily transmitted between firms, meaning patent protection going down when ϕ increases.
- The model in this case is showing us that sinergy is winning over rivalry, namely that going on a more extreme framework of technology sharing is optimal for these type of markets. Our empirical section will confirm these results.
- However we also want to see what happens for markets where the parameters are not normally distributed around 0.5. For this reason I am looping this model for all possible values from 0 to 1 with 0.01 steps of both patent and product similarity.
- For each of these combinations we plot the optimal patent protection, so we can have a good idea of how to maximize welfare with different levels of product and patent similarity that we know can vary from sector to sector.

Model

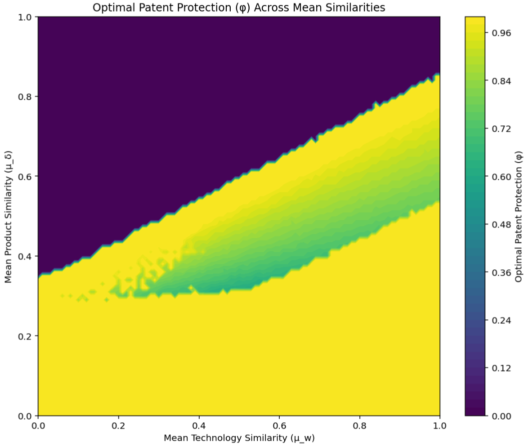


Figure 3: Optimal Technological Transmission with different Technology and Product Similarity

1. Introduction
2. Model
3. **Data**
4. Results

Product Similarity

- To quantify product similarity I have used a dataset from (Hoberg and Philips, 2010) based on product descriptions from 10-K statements filed yearly with the Securities and Exchange Commission.
- The authors measure the pairwise cosine similarity distance of all firms' 10-K statements using Word2Vec. The dataset has been recently updated to cover from 1989 to 2023. **Link**
- Companies with more than $10 million in assets and a class of equity securities that is held by more than 2000 owners must file annual and other periodic reports, regardless of whether the securities are publicly or privately traded. The authors have yearly 50,673 firms fillings pairwise cosine similarity.

Product Similarity

- Using SIC or NAICS for my research purpose would have several limitations. Neither adjusts significantly over time as product markets evolve, and neither can easily accommodate innovations that create entirely new product markets. In the late 1990s, hundreds of new technology and web-based firms were grouped into a large and nondescript SIC-based business services industry.
- More generally, fixed classifications like SIC and NAICS have at least four shortcomings: they only rarely re-classify firms that move into different industries, they do not allow for the industries themselves to evolve over time, and they impose transitivity even though two firms that are rivals to a third firm may not compete against each other. Lastly, they do not provide continuous measures of similarity both within and across industries.

Product Similarity

- The business description section of the document appears as Item 1 or Item 1A in most 10-K. These business descriptions are legally required to be accurate, as Item 101 of Regulation S-K legally requires that firms describe the significant products they offer to the market, and these descriptions must also be updated and representative of the current fiscal year of the 10-K. This recency requirement gives the measure the ability to change over time.
- A web crawling approach is used to extract only the business section from each 10-K filled in the SEC Edgar site. An example of Nvidia 10K 2025 annual 10-K report can be found at this link. You can also see a snapshot of the business description on the next page. **Link**

Product Similarity

Our Markets

We specialize in markets where our computing platforms can provide tremendous acceleration for applications. These platforms incorporate processors, interconnects, software, algorithms, systems, and services to deliver unique value. Our platforms address four large markets where our expertise is critical: Data Center, Gaming, Professional Visualization, and Automotive.

Data Center

The NVIDIA Data Center platform is focused on accelerating the most compute-intensive workloads, such as AI, data analytics, graphics, and scientific computing, delivering significantly better performance and power efficiency relative to conventional CPU-only approaches. It is deployed in cloud, hyperscale, on-premises and edge data centers. The platform consists of compute and networking offerings typically delivered to customers as systems, subsystems, or modules, along with software and services.

Our compute offerings include supercomputing platforms and servers, bringing together our energy efficient GPUs, CPUs, interconnects, and fully optimized AI and HPC software stacks. In addition, they include NVIDIA AI Enterprise software; our DGX Cloud service; and a growing body of acceleration libraries, APIs, SDKs, and domain-specific application frameworks.

Our networking offerings include end-to-end platforms for InfiniBand and Ethernet, consisting of network adapters, cables, DPUs, switch chips and systems, as well as a full software stack. This has enabled us to architect data center-scale computing platforms that can interconnect thousands of compute nodes with high-performance networking. While historically the server was the unit of computing, as AI and HPC workloads have become extremely large spanning thousands of compute nodes, the data center has become the new unit of computing, with networking as an integral part.

Our customers include the world's leading public cloud and consumer internet companies, thousands of enterprises and startups, and public sector entities. We work with industry leaders to help build or transform their applications and data center infrastructure. Our direct customers include original equipment manufacturers, or OEMs, original device manufacturers, or ODMs, system integrators and distributors which we partner with to help bring our products to market. We also have partnerships in automotive, healthcare, financial services, manufacturing, retail, and technology among others, to accelerate the adoption of AI.

At the foundation of the NVIDIA accelerated computing platform are our GPUs, which excel at parallel workloads such as the training and inferencing of neural networks. They are available in the NVIDIA accelerated computing platform and in

5

Table of Contents

industry standard servers from every major cloud provider and server maker. Beyond GPUs, our data center platform expanded to include DPUs in fiscal year 2022 and CPUs in fiscal year 2024. We can optimize across the entire computing, networking and storage stack to deliver data center-scale computing solutions.

While our approach starts with powerful chips, what makes it a full-stack computing platform is our large body of software, including the CUDA parallel programming model, the CUDA-X collection of acceleration libraries, APIs, SDKs, and domain-specific application frameworks.

Figure 4: NVIDIA 10-K (Annual report) 2025-02-26

Technology Similarity

- For technology similarity I use the measure developed in (Arts et al. 2023). To obtain firm technology portfolios, the authors rely on the DISCERN database that dynamically matches U.S. public firms to U.S. patents for the period 1980-2015 (Arora et al. 2021).
- The sample covers 1,345,945 US patents granted between 1980 and 2015 that are ultimately assigned to any public US firm. The technology portfolio of firm i in year t, they collected all the patents owned by firm i with a filing year between year t-5 and t-1, following (Ahuja and Katila 2001, Rothaermel and Deeds 2004, Hirshleifer et al. 2018).

Technology Similarity

- The authors train a Doc2Vec model on the entire corpus of US patents. Then they create a document-embedding vector for every patent, and average the patent-level vectors for all patents in a firms technology portfolio in a given year to map a firms spatial position in technology space in a given year.
- Next, they use these firms' technology portfolios to calculate tech similarity for every pair of firms and each year by means of cosine similarities, resulting in a total of 98,279,118 cosine similarities.
- They didn't use a BERT model because at the time it was limited to 512 tokens. Not anymore (**ModernBERT**) so it could be interesting to compute these values again.

Data

- For financial data I have used Compustat. For each firm I have: ID, fiscal year, total assets, R&D expenses, SIC code.
- I have used total assets to compute the HHI index at 4 digit SIC level.
- Each row is a relation between a firm and another firm, in each row we have both the product and technology similarity between the two firms. So I computed the HHI index for the first firm appearing in the row (HHI starting) and the HHI index for the second firm appearing in the row.
- The dataset can be considered a weighted network between firms where we have financial data for both firms and the weight of their relationship at the product and technological level.

Data

gvkey1	gvkey2	year	conm1	conm2	tech_similarity	product_similarity
105346	110039	2011	AMBIENT CORP	AMKOR TECH	0.41959804	0.0218
105346	110685	2011	AMBIENT CORP	L3 TECHNOLOGIES	0.53819847	0.0001

at_starting	emp_starting	xrd_starting	sich_starting	at_receiving	xrd_receiving	sich_receiving
21.874	0.088	11.665	4899	2773.047	50.386	3674
21.874	0.088	11.665	4899	15497	421	3812

Table 1: Firms Similarity Network Dataset

Results

$$R\&D_{i,t} = \beta_0 \cdot Tech_{j,t} + \beta_1 \cdot Prod_{j,t} + \beta_2 \cdot (Tech_{j,t} \cdot Prod_{j,t}) + \Gamma_{i,t} + \zeta_{SIC} + \Delta_t + \epsilon_{i,t}$$

- In the next page we are going to look the results of this regression with time and sectoral fixed effects.
- Our dependent variable is R&D investment.
- As we saw before every row represents the link between two firms with technology and similarity values.

Results

	Dependent variable: R&D Expenditure			
	No FE	Time FE	Time+SIC3 FE	Time+SIC4 FE
Tech-Prod Interaction	194.691***	191.368***	259.016***	249.906***
	(19.521)	(19.507)	(17.986)	(17.873)
Tech Similarity	69.963***	72.141***	37.478***	36.105***
	(1.602)	(1.617)	(1.563)	(1.577)
Product Similarity	3.424	4.604	−15.314*	−23.680***
	(8.934)	(8.928)	(8.310)	(8.272)
Assets (Starting)	0.072***	0.072***	0.073***	0.074***
	(0.0001)	(0.0001)	(0.00005)	(0.00005)
Employees (Starting)	−3.144***	−3.149***	−1.757***	−1.679***
	(0.036)	(0.036)	(0.037)	(0.038)
Assets (Receiving)	−0.0005***	−0.0005***	−0.0002***	−0.0002***
	(0.0001)	(0.0001)	(0.00005)	(0.00005)
Employees (Receiving)	−0.181***	−0.186***	−0.141***	−0.068*
	(0.036)	(0.036)	(0.037)	(0.038)
Concentration (Starting)	−2782.713***	−2786.832***	−2216.176***	−6196.441***
	(36.357)	(36.294)	(54.327)	(100.914)
Concentration (Receiving)	−607.816***	−611.935***	−173.955***	−428.554***
	(36.357)	(36.294)	(54.327)	(100.914)
Constant	−5.391***	−6.182	51.694***	77.849***
	(0.826)	(6.830)	(11.564)	(11.460)
Time FE	*No*	*Yes*	*Yes*	*Yes*
SIC3 Sector FE	*No*	*No*	*Yes*	*No*
SIC4 Sector FE	*No*	*No*	*No*	*Yes*
Observations	515, 472	515, 472	515, 472	515, 472
R^2	0.836	0.836	0.863	0.866

Note: Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Results

- Technology similarity has a positive coefficient, confirming that more technology similarity per se improves incentives for synergy.
- The interaction between technology and product similarity is positive, confirming the model result that showed that the aggregate effect of product and technology similarity is generally positive, namely sinergy dominates over rivalry.
- The other control variables confirm several literature findings, such as the fact that assets concentration decreases the incentives to invest in R&D

Results

- We still need to see if product similarity interacts negatively with technology similarity.
- The main mechanism in my model would make us expect that product similarity has similar effects on R&D throughout low levels of patent similarity.
- It should become significantly negative when technology similarity increases, because by increasing R&D expenditure when there is both high technology and product similarity the firm would help a competitor by investing more in R&D.
- I ran an interaction model to see if this was the case, we start by dividing the sample into different technology similarity categories.

Results

$$\text{TechCat}_i = \begin{cases} \text{tech\_0} & \text{if } 0.0 \leq \text{TechSimilarity}_i < 0.1 \\ \text{tech\_1} & \text{if } 0.1 \leq \text{TechSimilarity}_i < 0.2 \\ \vdots & \\ \text{tech\_9} & \text{if } 0.9 \leq \text{TechSimilarity}_i \leq 1.0 \end{cases} \tag{1}$$

Results

We then specify a regression model that includes interaction terms between product similarity and each technology category, using the lowest category (tech_0) as the reference group:

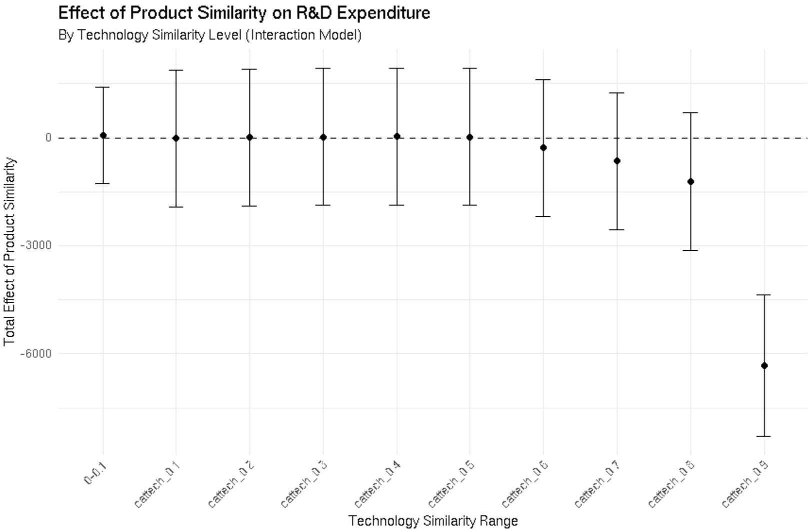
$$
\begin{aligned}
\text{R\&D}_i = \alpha + \beta_1 \text{ProductSimilarity}_i \\
+ \sum_{j=1}^{9} \gamma_j \text{TechCat}_{ij} \\
+ \sum_{j=1}^{9} \delta_j (\text{ProductSimilarity}_i \times \text{TechCat}_{ij}) \\
+ \boldsymbol{X}_i' \boldsymbol{\theta} + \lambda_t + \eta_s + \varepsilon_i
\end{aligned}
$$

Results

Where:

- $\text{R\&D}_i$ is the R&D expenditure for observation i
- α is the constant term
- β_1 is the coefficient for product similarity in the reference technology category (0.0-0.1)
- $\beta_1 + \delta_j$ represents the effect of product similarity when technology similarity is in range j
- γ_j are the coefficients for each technology category dummy variable
- δ_j are the coefficients for the interaction terms
- $\boldsymbol{X}_i$ is a vector of control variables
- λ_t represents time fixed effects
- η_s represents industry fixed effects (4-digit SIC)

Results


Effect of Product Similarity on R&D Expenditure
By Technology Similarity Level (Interaction Model)
Total Effect of Product Similarity
0
-3000
-6000
0-0.1
cattech_0.1
cattech_0.2
cattech_0.3
cattech_0.4
cattech_0.5
cattech_0.6
cattech_0.7
cattech_0.8
cattech_0.9
Technology Similarity Range


Thanks
FOR YOUR
ATTENTION