DIGITAL ACCESS TO SCHOLARSHIP AT HARVARD
DASH.HARVARD.EDU

HARVARD LIBRARY

Essays in Organizational Economics and Innovation

Citation
Sharoni, Brit. 2024. Essays in Organizational Economics and Innovation. Doctoral dissertation, Harvard University Graduate School of Arts and Sciences.

Link
https://nrs.harvard.edu/URN-3:HUL.INSTREPOS:37378740

Terms of use
This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Other Posted Material (LAA), as set forth at
https://harvardwiki.atlassian.net/wiki/external/NGY5NDE4ZjgzNTc5NDQzMGIzZWZhMGFlOWI2M2EwYTg

Accessibility
https://accessibility.huit.harvard.edu/digital-accessibility-policy

Share Your Story
The Harvard community has made this article openly available.
Please share how this access benefits you. Submit a story

HARVARD UNIVERSITY
Graduate School of Arts and Sciences

DISSERTATION ACCEPTANCE CERTIFICATE

The undersigned, appointed by the

Department of Economics

have examined a dissertation entitled

"Essays in Organizational Economics and Innovation"

presented by Brit Sharoni

candidate for the degree of Doctor of Philosophy and hereby certify that it is worthy of acceptance.

Signature ____________________

Typed name: Prof. Josh Lerner

Signature ____________________

Typed name: Prof. Zoe Cullen

Signature ____________________

Typed name: Prof. Claudia Goldin

Signature ____________________

Typed name: Prof. Oliver Hart

Date: April 23, 2024

Essays in Organizational Economics and Innovation

A dissertation presented

by

Brit Sharoni

to

The Department of Economics

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

in the subject of

Economics

Harvard University

Cambridge, Massachusetts

April 2024

© 2024 Brit Sharoni

All rights reserved.

Dissertation Advisor:
Professor Josh Lerner

Author:
Brit Sharoni

Essays in Organizational Economics and Innovation

Abstract

In this dissertation, I explore themes in organizational economics and innovation. The common thread running through all chapters of the dissertation is the recognition that production and innovation are not isolated processes; rather, they are intricately linked to the social networks individuals cultivate throughout their lives, whether they consist of colleagues or individuals who happen to influence them.

In the first chapter, I develop a tractable model of endogenous network formation among co-workers. In the model, workers aim to complete their tasks but may encounter failures, and when they do, they may reach out to one of their connections for help. I show that the extent to which workers invest in social networks is sensitive to the probability of failing to perform a task. For very low or very high failure rates, social networks offer limited value to workers, and therefore investment in connections is low. Intermediate failure rates, on the other hand, promote collaboration and, therefore, link formation. Furthermore, I uncover that excessive investment in social networks may occur because workers fail to recognize that their actions can impede other workers' access to assistance.

In the second chapter, I study knowledge spillovers between inventors who used to collaborate. Specifically, when inventors move to new locations, they carry knowledge and expertise, which may be a loss to their previous collaborators. But they might also become a bridge between otherwise disconnected innovation hubs, facilitating information flows and idea diffusion. I, therefore, study the effect of an inventor's relocation on their previous collaborators' productivity, both theoretically and empirically. The model helps to guide the

empirical analysis and to interpret the results. While empirically, I build a novel dataset and find sizeable positive effects on the productivity of inventors whose collaborators have relocated. These effects pertain not only to quantity, as gauged by the number of patents, but also to quality, as measured by the number of citations. I show that the core mechanism driving both effects is greater access to novel information networks and information.

And in the third chapter, my co-author and I develop a spatial growth model with a focus on mobility of innovators across space. The model highlights how inventor mobility, local knowledge externalities, and aggregate knowledge externalities jointly pin down the aggregate and distributional effects of innovation policy with a focus on regional and skill type heterogeneity. We combine the model with a quasi-natural experiment during Word War II, and enrich the data with information on local population size, skill composition, and innovator mobility. We use the reduced form estimates to discipline the model. In future work we will use the model to analyze the counterfactual impact of temporary and persistent innovation subsidies across US labor markets.

Contents

List of Tables

List of Figures

Acknowledgments

I extend my heartfelt gratitude to my dissertation advisors Zoë Cullen, Claudia Goldin, Oliver Hart, and Josh Lerner for their invaluable guidance and support. Zoë, for granting me a front-row seat to observe research at the highest levels. Claudia, for being the first to open the door for me and reassuring that transitioning between fields, though daunting, is achievable. Oliver, for standing by me in every step of the way, listening to my countless research ideas, and providing support throughout the entire process. And Josh, for the encouragement, patience, guidance and invaluable advice.

I am also deeply appreciative of my exceptional friends and colleagues at Harvard Michael Blank, Dev Patel, Jenna Anders, Benny Goldman, Adriano Fernandes and Pedro Degiovanni. I am immensely grateful for your support.

To my parents, my greatest supporters, who have helped me pursue and fulfill my dreams. You can take full credit for instilling in me the qualities necessary to embark on and complete this journey. I am grateful for the countless calls and visits.

To my brothers and sisters-in-law, who have always been there for me. Your encouragement have been invaluable to me, both academically and personally. Your belief in me has been a constant source of motivation, and I am deeply appreciative of the role you have played in my life during this challenging yet rewarding endeavor.

And to my spouse, I owe this achievement entirely to you. Your constant support, whether through guidance, offering a listening ear, or extending a helping hand, has been a source of strength and comfort. I cannot envision navigating this journey without you.

To my family

Introduction

In the first chapter of my dissertation, I look at how networks are endogenously formed among co-workers. Individuals have various strengths and weaknesses, and collaboration among colleagues can enhance a firm's overall productivity. I develop a tractable model of endogenous network formation among co-workers within the same firm. In the model, workers aim to complete their tasks but may encounter failures, and when they do, they may reach out to one of their connections for help. Therefore, workers face the tradeoff between the potential rewards of task completion and the costs tied to forming connections and requesting assistance.

I show that the extent to which workers invest in social networks is sensitive to the probability of failing to perform a task. For very low or very high failure rates, social networks offer limited value to workers, and therefore investment is low. In contrast, intermediate failure rates promote collaboration and link formation.

Furthermore, I uncover that excessive investment in social networks may occur because workers fail to recognize that their actions can impede other workers' access to assistance. They only see their own benefits, which they associate with having many weak links. I also explore approaches that can mitigate the over- or under-investment that often emerges, for example, whether remote work worsens inefficiency.

In the second chapter of my dissertation, I consider the effect of inventors' mobility on innovation. Specifically, I study, theoretically and empirically, how inventors' patenting productivity evolves as a result of their previous collaborators' relocation. The effect seems ex-ante, ambiguous. On the one hand, inventors leave with their skills and expertise when they relocate. So, in a world where this implies their co-inventors lose access to valuable information, the effect is expected to be negative. But those who relocate might also bridge between otherwise disconnected inventors in different innovation clusters, facilitating information flows and idea diffusion. That is, one view is that the network is reduced by one, whereas the other view is that the network is greatly expanded.

I formalize the argument using tools from network theory. In a simple patent production model, I highlight the dual role played by relocating inventors, acting as collaborators whose contribution is valuable and might be lost and as intermediaries facilitating information flow between inventors on the network, even when they move to different locations. The theory suggests a link between the collaboration's characteristics, like the frequency of joint patents or the identity of the new connections, and the effect size on the co-inventors' productivity after the move, suggesting predictions I can test empirically.

To estimate the effect, I combine a novel dataset and a matching design. A critical feature of the data should allow me to follow inventors' locations over time. The USPTO dataset, which provides information about inventors and their innovations, is tied to their patenting activity. Consequently, inventors are only observed in years when they apply for patents. Since most inventors do not patent annually, this dataset offers a limited view of their location history. To overcome this limitation, I utilize a second dataset comprising the universe of online professional profiles. This dataset includes details such as demographics, education, work locations, and employers on a monthly basis. I then use this information to construct a control group for the group of inventors experiencing a relocation of their co-inventors through an exact matching on the mover characteristics.

Using a difference-in-difference research design, I find sizeable positive effects on the quantity and quantity of patents, in contrast to the impact on inventors experiencing the unexpected death of a co-inventor or a collaborator, more broadly. Lastly, I show that a key mechanism driving the effects is greater access to a new information network and information sharing, confirming the model predictions.

In the third chapter of my dissertation, co-authored with Florian Trouvain, we investigate the impact of public R&D funding on domestic high-skilled migration in the United States. Building upon existing research investigating the impact of the expansion of R&D funding during and after World War II on patenting and innovation activity, we delve deeper into understanding the underlying mechanisms driving this effect, particularly focusing on the role of high-skilled migration. Specifically, we compare high-skilled migration patterns into counties that received this funding to those that did not. Rather than attributing this increase solely to the effect additional funding has current workers' productive, we explore the hypothesis that an influx of high-skilled workers into these counties contributes significantly to the reported productivity surge.

Our approach begins by constructing a two regions spatial Jones-Romer endogenous growth model, incorporating knowledge spillovers. We show that when high-skilled workers exhibit mobility, characterized by a non-zero labor elasticity, migration into funded locations is expected. Empirically, using census decennial data, we confirm that the proportion of high-skilled workers in counties that received funding is indeed higher post-funding. In particular, using a difference-in-difference research design, we compare the share of high-skilled workers in counties that receive some funding during the period between 1941 and 1948, to the share of high-skilled workers in counties that receive no such funding during that period. Our findings validate the model's prediction, indicating that the share high-skilled increases after the funding is distributed in counties that received it. While one might initially be concerned that this effect reflects the displacement of low-skilled workers, our analysis reveals that population growth in these areas also increases in these areas. This suggests that at least a portion of the effect can be attributed to the influx of new high-skilled individuals rather than merely an increase in the productivity of existing workers.

To establish the exogeneity of this shock, we ensure that its characteristics are not correlated with unobservable factors that may influence high-skilled migration. We verify this by testing for parallel trends in our analysis. Additionally, this shock coincides with the establishment of the Office of Scientific Research and Development (OSRD), which was primarily oriented towards addressing the immediate requirements for securing victory in the war. Its objectives were not focused on long-term strategies for distributing high-skilled workers across the United States, further supporting the exogeneity assumption.

Chapter 1

With a Little Help from My Friends: Productivity and Socialization in the Workplace

1.1 Introduction

An extensive literature in economics studies the effects of social networks in organizations while focusing on the hiring stage (Calvó-Armengol, 2004; Calvo-Armengol and Jackson, 2004; Calvó-Armengol and Jackson, 2007; Galeotti and Merlino, 2014; Merlino, 2014; Pallais and Sands, 2016; Merlino, 2019; Bolte *et al.*, 2020). Yet, the recent findings on the increasing rewards for social skills in the labor market (Deming, 2017) suggests that social networks within organizations hold significance not only during the hiring phase but also in subsequent stages.

Deming (2017) connects the increasing rewards associated with social skills to the concept of task trading within the workplace, which ultimately results in gains from specialization. This implies that individuals who possess strong social skills are better positioned to effectively trade tasks with others, thereby enhancing their productivity and efficiency.

On a related note, Jarosch *et al.* (2021a) develop a model illustrating how agents acquire knowledge and skills through interactions with their co-workers. This highlights the importance of social networks in facilitating learning and skill development within organizations.

However, despite these insights, there remains a gap in understanding the mechanisms underlying the formation of connections between workers and the incentives driving the emergence of more or less collaborative network structures in the workplace. In other words, while we recognize the significance of social networks in enhancing productivity and fostering learning, we have yet to fully grasp how these networks evolve and the factors influencing their structure within organizational settings. Addressing this gap could provide valuable insights into optimizing workplace dynamics and promoting effective collaboration among employees.

This paper delves into the role of social networks as a conduit for mutual help among coworkers within an organization and investigates its implications for workers' socializing efforts. Specifically, I develop a simple model of endogenous network formation that facilitates the exchange of help among colleagues who are connected to each other on the network. Each worker makes an ex-ante social investment, which then translates into a network of connections between individuals.

In the model, workers need to complete either easy or complex task, with task assignments being independent and identically distributed across agents. Workers assigned easy tasks can independently accomplish them and do so promptly. Conversely, those assigned complex tasks must rely on help from their network to complete their task. This help can only be provided by workers who are either unassigned or have already completed their own tasks.

The key trade-off facing a worker revolves around the social cost incurred by seeking help, on the one hand, and the consequent loss incurred if a task remains unfinished, on the other. Help carries a cost, albeit yielding a reward upon task completion. Concurrently, engaging in social interactions also imposes a cost. Consequently, workers must weigh the potential benefits, represented by the reward, against the associated losses, encompassing both the cost of help and the cost of socializing, when determining their degree of social engagement.

A reduction in the reward makes socializing less appealing, whereas reduced assistance costs or lower socializing expenses prompt workers to increase their social investment, leading to an anticipated increase in social investment, and as a result social connections.

Moreover, I also explore the mechanisms through which the organization can influence workers' social investment, either by incentivizing task completion or by regulating the proportion of tasks assigned to each worker. These considerations offer insights into how firms can effectively manage and incentivize social interactions among their workforce.

The model demonstrates that social investment is typically low when tasks are either extremely simple or highly complex, while peaking at intermediate levels of complexity. This phenomenon can be rationalized by the fact that help is unnecessary for easy tasks, as individuals can independently complete them, thus social connections hold little value. Similarly, in instances of highly complex tasks, the absence of available assistance renders social connections less valuable once more.

This finding holds implications in the ongoing discourse surrounding the advantages and disadvantages of remote work arrangements. Specifically, the model suggests that both easy and complex tasks are often faced independently, without the need or ability to ask collaborators for help, making them conducive to remote work setups. This observation aligns with recent research by (Bloom *et al.*, 2015), who demonstrate that the proportion of workers opting for remote work follows a U-shaped pattern in relation to wages. In other words, individuals are more likely to work from home if their wages are either very high or very low.

Finally, I delve into an alternative policy approach that firms can adopt to address the inefficient investment in social networks. Specifically, I investigate the possibility of assigning a portion of workers as designated helpers. These designated helpers are not tasked with specific assignments but are instead available to assist colleagues facing complex tasks when needed.

When the likelihood of being assigned a complex task is low, the firm opts not to designate any helpers. In such scenarios, workers are less likely to require assistance, resulting in lower social investment overall, and a limited inefficiency. Consequently, the cost the firm pays by losing out on tasks that designated helpers could have addressed is too high.

Conversely, when the likelihood of complex tasks is high, the firm opts to allocate a non-zero fraction of workers as designated helpers. This strategic allocation serves to mitigate inefficiencies and ensures that assistance is readily available when needed, thereby enhancing overall productivity.

Furthermore, my analysis reveals that when firms opt for an optimal reward structure for task completion, it fosters an efficient level of social investment, assuming workers face no financial constraints. Under an incentive-compatible contracting arrangement, firms can delegate tasks to workers, who stand to gain the full reward upon successful completion while receiving no compensation if the task remains unfinished. This setup ensures alignment between workers' incentives and the firm's objectives, leading to optimal levels of social engagement within the organization.

The rest of the paper is organized as follows. In section 1.2 we place our paper in the current literature. In section 1.3 we present the model and its results. Section 1.4 discusses some applications and Section 1.5 concludes.

1.2 Literature Review

Recent improvement in communication technology have opened new possibilities for remote work, a trend accelerated by the outbreak of the novel coronavirus in March 2020. A paper by Bloom *et al.* (2015), as well as recent work by Emanuel and Harrington (2020), find that workers that transition to working from home after spending some time in the office are on average more productive. Bloom *et al.* (2015) also find that they are less likely to be promoted. Whether workers are productive from home or not at least partially hinges on the importance of collaboration. If, for instance, workers rely on each other for help, and socializing is harder form home, productivity might fall as workers miss out on support from their colleagues.

Another strand of literature tries to tie differences in labor market outcomes to differences in one's workplace networks. Cullen and Perez-Truglia (2019) shows that socialization between men is highly correlated with future promotions, while there is no effect on women. Lindenlaub and Prummer (2021) explains the gender wage gap by suggesting that women networks are both more clustered and of a lower degree which, on the one hand, make them more exposed to peer pressure, and on the other hand less exposed to new information. Therefore, women thrive in occupations where uncertainty does not play a big role, but trust does.

Lastly, Battiston *et al.* (2021a) emphasises the importance of face to face interaction for collaboration in firms. They show that workers are more likely to provide informations to their team-mates if these are located in proximity. Our paper incorporates this finding by modelling socializing as more costly when performed remotely.

The paper also relates to Jacobs and Watts (2021) that measures professional networks in many firms across the globe.

1.3 The Model

There are n risk neutral workers. Time evolves in discrete periods. Each period begins with all workers handling a task and earning an exogenous fixed base salary f.

Tasks take one of two realizations. It can be complex, with an exogenous probability $b \in (0, 1)$, or easy, with the compliment probability $1 - b$. At the beginning of the period agents learn their task realization. In particular, any given worker fails to complete her assigned task with with probability b, independently across agents. Workers who complete their tasks become available, while the rest remain occupied.

Occupied workers, who did not complete their task,look for an unoccupied coworker with whom they can share the burden of the task with. Here is where the network plays a role. It prescribes who can help whom, as help can only provided between two connected coworkers, and only once at a time. In that sense, social connections provide another type of information, informing workers about their available resources.

An undirected graph g summarizes the links between all workers, where $g_{ij} = 1$ indicates that workers i and j know each other, and $g_{ij} = 0$ means they do not.

A worker that fails to complete her task approaches all of her linked coworkers in the goal of finding help. An unoccupied worker who is approached by an occupied worker randomly picks one of his coworkers to whom he provides the help. An occupied worker who does not find help through her network fails to complete her task, and must report to her manager.

In order to insure themselves against the risk of failing to complete their tasks, workers invest in social connections, which translates into link probabilities and their random network.

The network formation game follows Cabrales *et al.* (2011). Each worker i chooses a costly network investment $s_i \geq 0$ from the set of pure strategies $S_i = \mathbb{R}_+$, simultaneously. The marginal cost of a unit of investment is equal to $c > 0$. Given a profile $\mathbf{s} = (s_1, ..., s_n)$ of pure strategies, the probability a link between workers i and j is formed is given by

$$\mathbb{P}\left(g_{ij} = 1 \middle| \mathbf{s}\right) = \begin{cases} \min\left\{\frac{s_i s_j}{\sum_{k \in N} s_k}, 1\right\} & \text{if } \sum_{k \in N} s_k > 0 \\ 0 & \text{otherwise} \end{cases}$$

where $\sum_{k \in N} s_k$ is the total social investment induced by the profile $\mathbf{s} = (s_1, ..., s_n)$ of pure strategies.

If all workers exert the same level of effort s, the induced random graph is binomial. In this case the probability a link is formed between two arbitrary workers is given by $\min\left\{\frac{s}{n}, 1\right\}$ and the average degree is $\min\left\{\frac{s}{n}, 1\right\} \cdot (n-1)$.

Workers are paid a reward w whenever they complete their task successfully. Following Sandvik *et al.* (2020), workers find it costly to seek for help. They incur a cost $k > 0$ when receiving help through one of their linked coworkers. This captures the compensation workers get for the time invested in helping a connect colleague. It may be thought of as the cost of taking him for lunch.

We assume that $k < w$ in order to induce workers to interact with each other. If that were not the case, workers would have preferred to avoid interaction, and complete only the tasks that they can complete by themselves, as help would have been very costly.

If we consider the case of a large office setting, i.e. when $n \to \infty$, and a symmetric investment, s, the probability a link between an arbitrary pair of agents is formed, is given by

$$p(s) = 1 - \exp\left(-\frac{1-b}{b}\left(1 - \exp\left(-sb\right)\right)\right)$$

Proposition 1.3.1. Suppose that $s_i = s$ for all $in \in N$. Then, the matching rate between an arbitrary pair of agents is given by

$$p(s) = 1 - \exp\left(-\frac{1-b}{b}\left(1 - \exp\left(-sb\right)\right)\right)$$

$p(s)$ has the following properties:

1. It increases in the social investment, s.

2. It decreases in the non-completion rate, b.

Proof. See Appendix A.2. □

1.3.1 Workers' Problem

Let $\mathbf{s}$ be a strategy profile. We consider the system in a steady state and look at a typical period. Worker i's expected utility is given by

$$\mathbb{E}\left[U_i\left(s_i, s_{-i}\right)\right] = f + w \cdot \left(1 - b\left(1 - p_i\left(\mathbf{s}\right)\right)\right) - kb \cdot p_i\left(\mathbf{s}\right) - cs_i$$

where $p_i(s)$ is the probability that worker i receives help through at least one coworker in their network and is equal to

$$p(s) = 1 - \exp\left(-\frac{1-b}{b}\frac{s_i}{s}\left(1 - \exp\left(-sb\right)\right)\right)$$

whenever all other workers $j \neq i$ exert effort that is equal to s.

With probability b worker i fails to complete her task. With probability $p_i(s)$ she finds help through her linked coworkers and with probability $1 - p_i(s)$ she does not. The cost k is associated with the first case. In the second case, the worker does not complete the task, and therefore is not rewarded w for completing the task.

Definition 1.3.1 (Equilibrium)**.** A pure strategy profile $\mathbf{s}$ is an equilibrium if for all $i \in N$

$$\mathbb{E}\left[U_i\left(s_i, s_{-i}\right)\right] \geq \mathbb{E}\left[U_i\left(s_i', s_{-i}\right)\right] \quad \forall s_i' \in S_i$$

We focus on a symmetric equilibria with positive social investment. Which implies that all workers choose the same level of social investment. The symmetric social investment $s_i = 0$ for all $i \in N$ is always a symmetric equilibrium, but we restrict ourselves to those in which the social investment is positive. Thus, Agent i's chooses a social investment $s_i \in S_i$ that maximizes her expected utility when all other agents use the same strategy s

$$\max_{s_i \in S_i} \mathbb{E}\left[U_i\left(s_i, s\right)\right] = f + w \cdot \left(1 - b\left(1 - p_i\left(\mathbf{s}\right)\right)\right) - k_{\text{help}} b \cdot p_i\left(\mathbf{s}\right) - c s_i$$

where $s_j = s$ for all $j \neq i$.

If an interior solution exists, then $s_i = s^*$ solves the first order condition which is given by

$$\left[w - k_{\text{help}}\right] b \frac{\partial p_i}{\partial s_i}\left(s^*\right) = c$$

Proposition 1.3.2. An interior solution exists if and only if $c < \left(w - k_{\text{help}}\right)(1-b)\, b$. In this case, the unique solution s^* solves

$$\frac{\left(w - k_{\text{help}}\right)(1-b)}{s^*}\left[1 - \exp\left(-s^* b\right)\right] \exp\left(-\frac{1-b}{b}\left(1 - \exp\left(-s^* b\right)\right)\right) = c \tag{1.1}$$

Proof. See Appendix A.2. □

Firm's Problem

The firm, on the other hand, values completed tasks by v. It cannot observe the help network, and therefore cannot contract helping. That said, the cost k is exogenous and is not internalized by the firm. Thus, the firm's expected utility is given by

$$\mathbb{E}\left[U_F\left(s\right)\right]=\left(v-w\right)\left(1-b\left(1-p\left(s\right)\right)\right)-f$$

We begin by analyzing the first best solution. The social planner chooses the socializing effort , s, that maximizes the social surplus:

$$\max_{s}\ \left(v-k\right)\left(1-b\left(1-p\left(s\right)\right)\right)-cs$$

Proposition 1.3.3. An interior solution to the social planner's problem exists if and only if $c<\left(v-k\right)\left(1-b\right)b$. In this case, the unique solution s^*_{SP} solves

$$\left(v-k\right)\left(1-b\right)b\left[1-\exp\left(-s^*_{SP}b\right)\right]\exp\left(-\frac{1-b}{b}\left(1-\exp\left(-s^*_{SP}b\right)\right)\right)=c \tag{1.2}$$

Proof. See Appendix A.2. □

In general, for any given $w<v$, we are not able to rank s^* and s^*_{SP}. On the one hand, the social planner internalizes that the more crowded the network is, the lower workers' ability to find help is, as many of them now compete over the same unoccupied workers. This force pushes down the social investment. On the other hand, the social planner's reward for completing a task is $v>w$, which makes links more valuable. This force pushes social investment up. Therefore, we get over-investment when the first effect dominates the second, and under-investment otherwise.

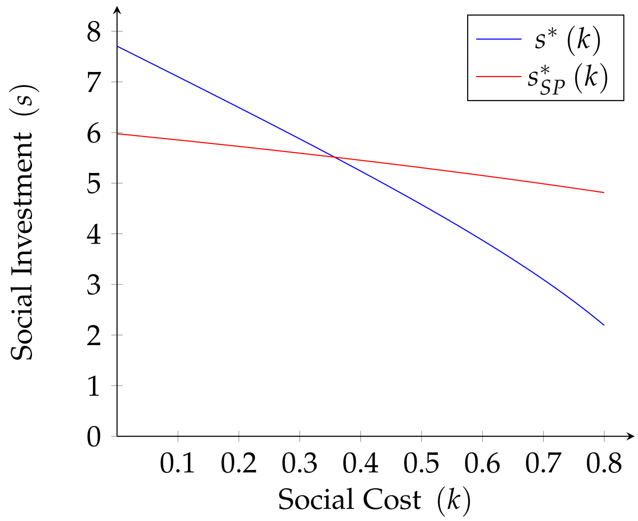


Figure 1.1: Over- and Under-Investment as a Function of the Social Cost

The point where these two graphs meet is the point where the two effects that determine the order in which s^* ans s^*_{SP} are ranked, exactly offset each other. In the range before they meet we get over social investment from the firm point of view, and then, under social investment.

The firm, on the other hand, cannot control workers' social investment, and therefore decides on the task completion reward. They take the workers' solution as given, and maximize its utility conditional on the incentive constraint and the individually rational constraint.[1] Thus, the firm solves:

$$\max_{k<w<v} \mathbb{E}\left[U_F\left(s^*\right)\right] = (v-w)\left(1-b\left(1-p\left(s^*\right)\right)\right) - f \tag{1.3}$$

$$\text{such that} \quad \text{(IR) } \mathbb{E}\left[u_i\left(s^*\right)\right] \geq 0$$

$$\text{(IC) } s^* \text{ is the solution to the workers' problem}$$

Proposition 1.3.4. A solution to the firm's problem, w^*, exists and is unique. The first best is achieved at w^* when workers do not face wealth constraints.[2]

[1]We normalize the workers outside option to zero.

[2]At the optimum $f < 0$, and therefore we must assume that the workers do not face wealth constraints.

Proof. See Appendix A.2. □

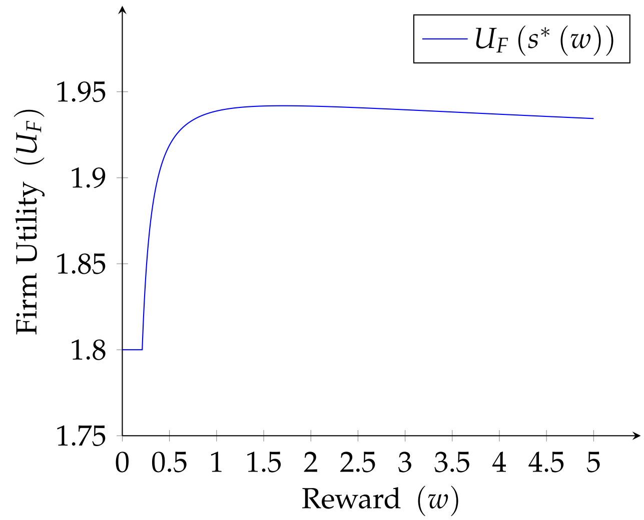


Figure 1.2: Firm's Utility as a Function of Reward w

Comparative Statics

In the Figures below, we present how s^* and s^*_{SP} change as a function of the parameters of the model.

Proposition 1.3.5. For any social cost $k \in (0, w)$, completion bonus $w < v$, and social investment cost c,[3] there exists a non-completion rate, $\bar{b} \in (0, 1)$, such that the social investment in equilibrium, s^*, increases on $(0, \bar{b})$ and decreases on $(\bar{b}, 1)$.

Proof. See Appendix A.2. □

[3] Such that a strictly positive unique solution to the workers' problem exists.

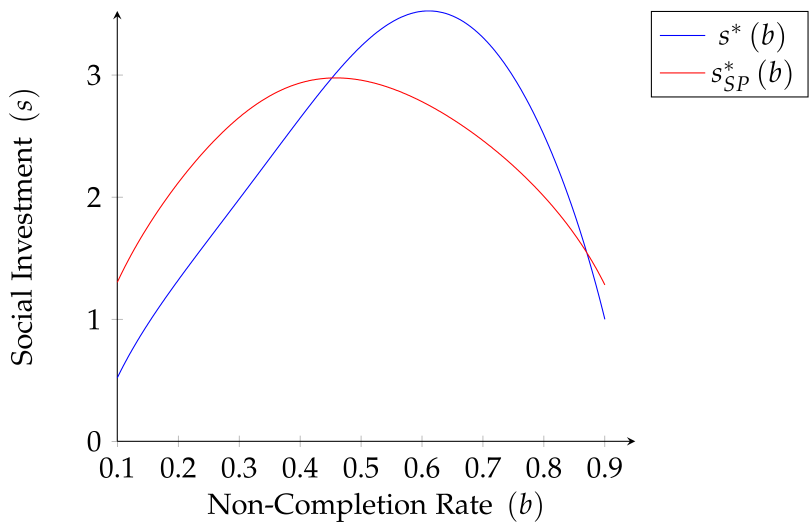


Figure 1.3: Social Investment Follows an Inverted U-Shape as a Function of the Non-Completion Rate

The increase in the non-completion rate is associated with two opposite effects. The first, the higher the non-completion rate is, the lower the fraction of workers that complete their task is. As a result, fewer workers are available to provide help and the social connections become less valuable. This effect pushes down the social investment. The second, when the non-completion rate increases, so does the likelihood workers will become overloaded. Thus, social connections become more valuable, which pushes the social investment up.

The latter effect dominates first, and this is why the social investment increases at the beginning. Then the former effect dominates and the social investment starts to fall.

1.4 Applications

1.4.1 Working from Home

One way in which workers can interact with each other is over coffee breaks, lunch breaks or happy hours. These type of events, which ease the formation of social links, usually take place in person, and sometimes happen randomly.

When workers work from home, these interactions may happen less often as 1. they have to be planned in advanced, and 2. workers may not live in proximity, which makes joint lunch breaks almost impossible. Therefore, the transition to work from home may be captured by an increase in the marginal cost of socializing, c.

In this subsection the firm is facing the decision of whether to have its workers work from home or from the office, given an exogenous completion reward w. The benefit of having workers work from home comes in the form of reduction in renting costs which are no longer needed. Let r denote the rental cost per square foot, and O the total office space needed in order to host all of the workers. Thus, the firm decides to keep its workers in the office if and only if

$$U_F\left(s^*\left(c\right)\right) - U_F\left(s^*\left(c'\right)\right) \geq rO$$

for c being the marginal cost of socializing in the office, and $c' > c$ the marginal cost of socializing when working from home.

This equation can be written as

$$\left(v - w\right) b \left[p\left(s^*\left(c\right)\right) - p\left(s^*\left(c'\right)\right)\right] \geq rO$$

Since the workers' solution, s^*, decreases in c, we get a threshold c^* such that the costs of renting an office equal the costs of working form home, which come in the form of lower completion rate.

When b is very high or very low c has little effect and working from home is preferred. In these cases socializing is very low, and help is rarely provided, and therefore the transition to working from home does not incorporate big losses. On the other hand, for medium values of b, working from home is favorable, especially for high marginal costs c'.

1.4.2 Designated Helpers

One way in which the firm may try to increase provision of help is through deciding on a designated set of workers who are not assigned a task at the beginning of the period, and whom their only task is to provide help. This decision has two opposite effects, on the one hand it means that some of the tasks the firm faces will never be completed as their are not assigned to any of the workers. On the other hand, more workers are available to provide help, which may imply that a higher fraction of the assigned tasks is completed.

In order to induce these workers to socialize with the rest we will assume that workers do not their role ex-ante, and they only realize that once tasks are assigned. Let $\delta \in (0, 1)$ be the fraction of workers who are not assigned a task. The remaining workers are assigned tasks, and face the same probability b of a hard task realization, independently across tasks. The probability a link between an arbitrary pair of workers is formed is therefore given by

$$p(s) = 1 - \exp\left(-\frac{\delta + (1-\delta)(1-b)}{(1-\delta)b}(1 - \exp(-(1-\delta)bs))\right)$$

Now, the fraction of idle workers is given by $\delta + (1-\delta)(1-b)$ and that of occupied workers by $(1-\delta)b$.

For a given value of w, the form solves:

$$\max_{\delta \in (0,1)} \mathbb{E}[U_F(s^*)] = (v-w)(1-\delta)(1-b(1-\delta)(1-p(s^*))) - f$$

$$\text{such that} \quad \text{(IR) } \mathbb{E}[u_i(s^*)] \geq 0$$

$$\text{(IC) } s^* \text{ is the solution to the workers' problem}$$

Proposition 1.4.1. There exists a unique solution to the firm's problem and it is given by $\delta^* = 0$ when $b \leq \frac{1}{2}$, and interior otherwise.

Proof. See Appendix A.2. □

The reason the firm prefers to assign tasks to all of its workers follows by the fact that an increase in δ is associated with two effects that lead to a decrease in the social investment, and therefore to a higher decrease in the fraction of tasks which are completed.

First, when δ increases workers do not need to seek for help with higher probability, which pushes down their social investment. Second, since more workers are available to provide help, links become less valuable which also pushes down the social investment.

1.5 Conclusion

This paper studies the role of social networks for firm and worker productivity and explicitly takes into account the cost and benefits of collaboration in the workplace. In each period workers are assigned a task that vary in its complexity. The realized complexity of the task determines if a task my be completed alone, or whether help is needed. Help, when needed, is provided through a network of social contacts. The endogenous network arises as a result of workers' costly social investment. The model pins down the level of the investment and highlights the trade-off between the costs incurred when receiving help and the losses realized when a task is not completed.

We show that, perhaps surprisingly, the increase in the level of complexity does not translate into an increase in social effort exerted by workers. This result can be understood as follows. On the one hand, complexity makes links more valuable as workers are more likely to seek help. Yet, higher complexity simultaneously reduces the fraction of workers that are available to provide help, ultimately pushing down social investment.

We also study how firms can achieve an efficient network structure. This is the case even when help is not observed, as long as there are no wealth constraints. The results of the model line up with recent practice of employers in light of a shift to work from home. Since networks do not play an important role for very complex or very easy tasks, these are precisely the tasks that could be done from home without substantial productivity losses.

Chapter 2

The Effect of Inventor Mobility on Network Productivity

2.1 Introduction

Inventors' mobility matters for innovation. When inventors relocate to innovation clusters they experience an increase to their patenting productivity (Moretti, 2021).[1] This effect can be attributed to several factors, including agglomeration externalities, knowledge spillovers, and the collaborative nature of innovation, where patents are often the result of teamwork (Jones, 2009).

While a large literature has focused on the effect of a reallocation of an innovator on their own patenting productivity, surprisingly little is know about the effects on the productivity of the movers' colleagues, who remain in the origin location. Given that most patenting activity involves team efforts, with more than 70 percent of patents filed with the United States Patents and Trademark Office in 2022 having at least two inventors listed, it seems reasonable that the reallocation of a colleague impacts the productivity of the inventors that

[1] See also Ellison and Glaeser (1997) and Bloom *et al.* (2021) who highlighted the existence of innovation clusters. And Ellison and Glaeser (1999), Ellison *et al.* (2010) and Greenstone *et al.* (2010) who document the advantageous outcomes they offer to both inventors and firms deciding to relocate there.

stayed in the original location.[2] And this effect might be as important as the productivity effects related to the mover's patenting activity. In this paper, I quantify the impact of an inventor's relocation on their former co-inventors' productivity, both theoretically and empirically.

Ex-ante, it is unclear whether a colleague's departure positively or negatively impacts the productivity of their former collaborators. On the one hand, when inventors relocate they take their skills and expertise with them, so in a world where the inventors who stay in the origin location are destitute since they lost access to all of their productive colleagues, the effect is expected to be negative.[3] But, on the other hand, in a world where their former collaborators expose them to new techniques and ideas, the situation becomes more complex. In this instance, it is not clear whether the inventors staying in the origin location are better off, even if they stop co-patenting with the mover entirely. Understanding which of these scenarios applies is crucial for innovation.

I develop this argument formally in a simple model of team production and patent creation. I build on recent advances in network theory to illustrate the distinct roles information sharing and collaboration play after a collaborator's relocation. In the model, the net effect of the move depends on the network characteristics, such as the intensity of the collaboration between inventors and the extent to which a colleague's move offers access to a new network and novel information. While the answer to this question is empirical in nature, the model helps to interpret the results and to guide the empirical test.

A significant obstacle in this work is the endogeneity of relocations, as they are determined by various observed and unobserved factors, with workers and firms deciding when the move occurs and its destination. One concern, for instance, is selection bias. If the inventors who stay in the origin location when their collaborator relocates are substantially

---

[2]Rising collaboration rates appear in the ascending pattern observed in both the average count of inventors listed on a patent and the proportion of patents generated through collaborative efforts, as demonstrated in Jones (2009). It is part of a longer trend that reflects a shift towards collaborative knowledge creation, facilitated by information sharing (Wuchty *et al.*, 2007).

[3]As was shown by Azoulay *et al.* (2010) and Jaravel *et al.* (2018) who use death shocks as a source of an exogenous variation.

different from the inventors who do not experience the relocation of their collaborators, the estimated results might reflect this difference instead of a true effect.

To overcome this challenge, I build on the methodology introduced in Jaravel *et al.* (2018). I create a control group of inventors who have never experienced a relocation of a collaborator but have collaborated with an inventor who is similar to a mover. The similarity is based on various observables using an exact matching procedure. I then identify all inventors who have collaborated with those in the matched control group. By adopting this approach, I can compare the treatment group (co-inventors of movers) to the control group in a difference-in-difference research design. This procedure helps to mitigate the endogeneity concern and enables a more robust estimation of the causal effect of co-inventor relocations on the stayers' productivity.

An additional hurdle is related to the structure of the patent data. Although the patent data allows for the tracking of inventors over time, observations are tied to patenting activity, meaning that information on inventors, such as location, is only observed when they apply for patents. Since most inventors do not patent every year, it becomes difficult to locate inventors with certainty during the time intervals between patents.

To surmount this challenge, I compile a novel dataset by merging information about inventors in the United States with their online professional profiles. This enables me to create a comprehensive longitudinal record of inventors based in the United States and, in turn, allows me to track their locations over time. Consequently, I can identify the timing of all moves and gain valuable insights into the characteristics of the move. These specifics, including details about the firms where the movers are employed and the geographical distances involved, enable a comprehensive exploration of the impacts of co-inventor relocations on innovation networks and productivity, as well as the underlying factors shaping these effects.

The procedure I employ results in a longitudinal data on approximately 300,000 inventors based in the United States between 1990 and 2022. I identify 49,902 inventors who relocated during this period. Following a co-inventor's move, I find that the inventors who remain in the original location increase their annual number of patent applications by an average of about 9 percent, when compared to the control group. This overall effect becomes evident over approximately three years and weakly increases over time. Notably, this increase in innovation quantity does not seem to compromise on quality, as there is an increase of almost 15 percent in the average number of adjusted forward citations, compared to the control group.[4]

I then explore the specific mechanism that drives the estimated positive effects. I first address a concern related to common team productivity shocks, which could bias the results upwards and overstate the positive effect of the move. If the timing of the relocation aligns with periods following a heightened patenting success, it is conceivable that the inventors staying in the origin location, who are also listed on these successful patents, might experience the benefits of the success in ways distinct from a relocation, yet still impacting their productivity, such as research grants or improved working conditions. I show that the size of the effect changes in ways that do not align with the common shock reasoning.

Next, I present evidence on information flows towards the inventors who stayed in the origin location. I show that, after the move, the inventors who remain in the origin location start citing patents which are produced in the mover's destination more frequently than before the move. Additionally, when compared to the control group, I show evidence that the inventors who remain in the origin location expand their network into the mover's destination location, after the move. The findings imply that some information flows from the mover's destination to their former collaborators after the move, either by an exposure to patents produced in that location or directly through collaboration with inventors who are located there. I support this reasoning by showing further evidence on the likelihood of patenting in the same technology class after the move.

---

[4] Hall *et al.* (2001) and Lerner and Seru (2021) emphasize the importance of weighted number of patents to account for the potential bias generated by patenting trends over time and field. Hall *et al.* (2005) and Lerner *et al.* (2011) show that the adjusted number of citations not only sheds light on an inventor's patenting activity and the influence of the patent, but is also shown to carry economic value.

Finally, to establish that this mechanism predominantly hinges on the access to new information networks, I examine the effect by contrasting two scenarios. In the first scenario, I analyze cases where the inventor who stayed had previously collaborated with inventors in the destination location *before* the move, and compare that to a scenario in which such collaborations did not exist. This comparison allows me to assess the difference between cases where the inventor gains access a new cohort of inventors potentially possessing novel information and cases where the information exchange may have already taken place. I find a positive and statistically significant impact when the network is new, while observing a statistically insignificant effect when access happened prior to the move. Specifically, in comparison to the control group, the left-behind inventors experienced a 10 percent increase in their annual number of patents and a 21 percent increase in their annual number of adjusted citations.

Collectively, these findings suggest that the main driving force behind the observed effects is access to a new network and, consequently, to new information resulting from co-inventor relocations.

This paper is connected to various strands of the literature. Firstly, it relates to studies that examine the effects of migration on productivity in the origin location (Kerr *et al.*, 2016, 2017; Kerr, 2008; Waldinger, 2012) and the horse race between "brain drain" and "brain gain."[5] I, not only focus specifically on domestic migration within the United States and the effects thereof, but also take into account some characteristics of the move and the mover, which explain some of the differential effects that I estimate.

Additionally, it relates to research papers such as Azoulay *et al.* (2010) and Jaravel *et al.* (2018) that study adjacent topics, and in particular the significance of team-specific capital and network structure and find a negative effect on the productivity and labor

[5] In a more recent paper, Prato (2022) examines the impact of migrants from Europe on innovation activity in the United States, as well as the reciprocal effect. Similarly, Bernstein *et al.* (2022) study the influence of high-skill migration on innovation within the United States. Moser *et al.* (2014) use the emigration of Jewish Germans from Germany to the United States to address this topic in the setting of Chemical innovation. Other papers highlight the possibility that reallocation of skilled labor into the innovation sector can have negative effects on technology adoption in other parts of the economy, resulting in a form of brain drain (see for instance recent work by Trouvain (2022)).

market outcome of inventors and scholars experiencing the unexpected death of one of their colleagues. In a sharp contrast, I estimate positive productivity effects, which also evolve over time, following a colleague's relocation. These results are quite surprising and hinge on the idea that in contrast to a death shock, which implies a complete discontinuation of any relationship, a colleague's relocation introduces the potential for spillovers, either directly trough continued collaboration or even in the form of network expansion, both of which are impossible when the colleague passes away. Figure 2.1 illustrates that both these avenues are open in the case of a relocation, and both can be used to share information, resulting in the positive spillovers we observe. Specifically, the links between inventors who stay in the original location after their co-inventor's move and inventors in the destination locations expand, leading to increased collaboration from 4 percent to almost 16 percent. And since experiencing a colleague's relocation is an order of magnitude more common than a colleague's death, it has important implications for innovation.

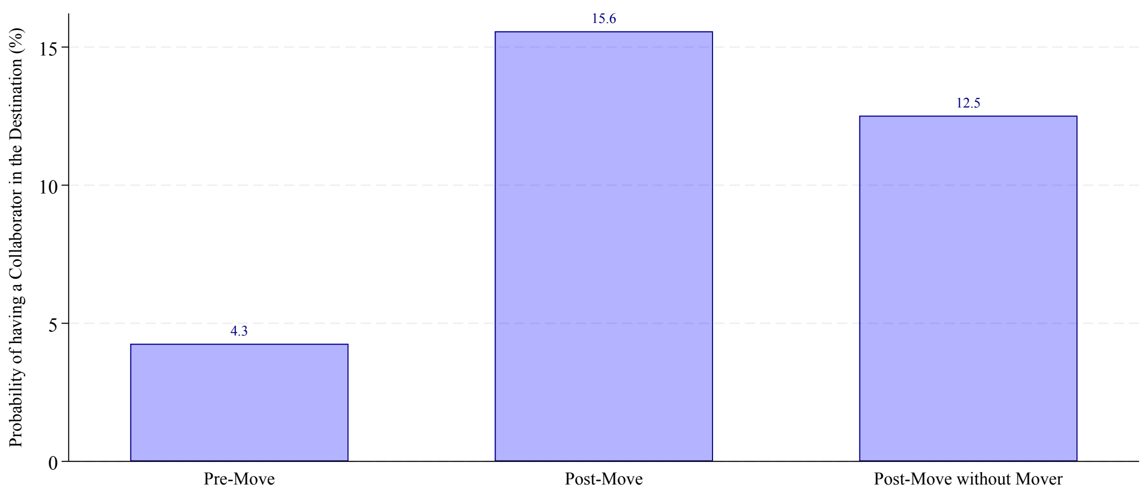


Figure 2.1: Collaboration Patterns Post-Move

Notes: This plot depicts a comparison of the likelihood of having a collaborator in the destination location of the mover before and after the relocation. The first bar on the left shows the probability of having a collaborator in the destination location prior to the move, while the subsequent two bars represent the probability after the move. The middle bar includes all inventors in the destination, and the right bar only counts inventors other than the mover in the destination. The probability is calculated by dividing the number of origin location inventors with collaborators in the destination by the total number of origin location inventors.

This research is also connected to Moretti (2021), who studies how the size of an innovation cluster affects the innovative output of inventors within that cluster. Leveraging the variations introduced in this paper, as well as the definition of economic regions, my study focuses on understanding the effects on inventors who stayed in the origin location when their collaborators relocate, rather than concentrating on the relocating individuals themselves. This approach enables me to delve deeper into the effects associated with relocations from a different perspective, and study the implications for the other group of inventors who might be affected by this relocation.

My results add to the results by Agrawal *et al.* (2008), by showing that collaboration with an inventor who moves to a new location, can serve as a channel though which the stayer can get access to a new network and, therefore, to promote knowledge spillovers between inventors who are not co-located. It also corroborates the findings of Zacchia (2020), which demonstrate that collaborative patent activities spanning different firms facilitate the transfer of knowledge between these firms. In contrast, my focus is on the individual productivity of inventors, and I use prior collaboration as a prerequisite for subsequent knowledge spillovers, rather than leveraging current collaboration for immediate spillovers. Moreover, I examine situations where an inventor relocates and analyze how this relocation affects the productivity of inventors who remain in their original location. This analysis includes cases where the relocation leads to the termination of the mover and the left behind collaborative relationships, and I do not assume a static framework. Moreover, some of the effects I capture follow from within firm relocations and spillovers.

My research is also closely related to Zacchia (2018). However, there are notable differences between our studies. First, I do not restrict the relocations to be undertaken by superstar collaborators, thus examining the question in a broader context. Second, my dataset offers more comprehensive information about inventors, including their locations even even in years in which they do not patent. This allows me to create a comparable control group through exact matching and use a different identification strategy that leverages recent advancements in the field. Additionally, I provide insights into the mechanisms behind the observed results, a facet not investigated in the referenced paper.

Lastly, this study is connected to other related literature that study team-based innovation and the significance of collaborative efforts in idea generation (Wuchty *et al.*, 2007; Crescenzi *et al.*, 2016; Jaffe *et al.*, 2015; Jones, 2009). It also has relevance to the theoretical literature concerning team-based or network-based human capital (Mailath and Postlewaite, 1990; Chillemi and Gui, 1997) and theories encompassing the transfer of knowledge among inventors (Stein, 2008; Lucas Jr, 2009; Lucas Jr and Moll, 2014; Jarosch *et al.*, 2021b). Additionally, while I focus on collaboration between geographical locations within the United States, Kerr and Kerr (2018) study the characteristics of collaborative patents where the inventors listed on them are located both within and outside of the United States.

The reminder of the paper is organized as follows. In Section 2.2, I outline the model. Section 2.3 describes the data and sample construction. Section 2.4 reports the estimates. Section 2.5 covers the mechanisms and Section 2.6 concludes.

2.2 A Model of Collaboration

In this theoretical section, I provide a simple framework that can be used to form predictions and elucidate the empirical findings on the productivity effect of a relocation on the mover's former patent collaborators.

The model incorporates two key drives of inventors' productivity: collaboration, through team-production, and information sharing on a network. While both information acquisition and collaboration are integral components of the patent production function, the channels through which they affect one's productivity are distinct. Therefore, the model allows me to explore the scenarios in which the cost of losing a collaborator due to a relocation can be offset by the opportunity to access a new information network through the mover in their new location.

2.2.1 Basic Framework

Inventors' Network

A society of n inventors is connected via a directed and weighted network, which has an adjacency matrix $\mathbf{W} \in [0,1]^{n\times n}$. A general element $w_{ij} \in [0,1]$ represents the status and the intensity of the relationship between inventor i and inventor j. One can think about w_{ij} as the share of the patents both inventor i and inventor j are listed on out of the total number of patents inventor i is listed on. Similarly, w_{ii} is the fraction of solo-patents produced by inventor i by themselves. In other words, w_{ij} is a proxy for the fraction of time inventor i spends with inventor j producing patents. Specifically, an entry $w_{ij} = 0$ implies that inventor i and inventor j do not collaborate on patents, and therefore are not connected. And a higher w_{ij} corresponds to a stronger relationship. Note that although the matrix $\mathbf{W}$ is not symmetric by assumption, collaboration is a reciprocal relationship, and therefore $w_{ij} > 0$ is and only if $w_{ji} > 0$.[6]

Figure 2.2 presents an example of a network with three inventors. In this example, inventor 1 collaborates with inventor 2 and inventor 3, with whom they spend 0.5 and 0.15 of their time, respectively. Inventor 2 and inventor 3 do not collaborate with one another, and they spend 0.25 and 0.35 of their time patenting with inventor 1, respectively.

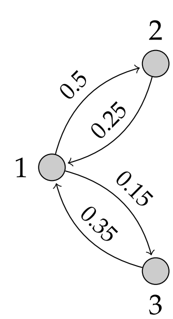


Figure 2.2: Network of Inventors

[6]The reciprocity is only with respect to zero. That is, either both the elements w_{ij} and w_{ji} are greater than zero, or both are equal to zero. However, when positive, I impose no assumption on whether they are equal or not.

Inventors also have a fixed ability level $\alpha_i \in \mathbb{R}_+$ and an initial knowledge level $k_i \in \mathbb{R}_+$. These concepts play a central role in patent production and information sharing. Specifically, inventors exchange their knowledge through collaboration and apply their skills when creating patents.

Information Acquisition

Inventors acquire information through their network of inventors, not just from their immediate connections, but also from inventors located farther away. In particular, inventors can acquire information from inventors they collaborate with directly, as well as their second degree connections.[7]

To formalize this concept, it is helpful to introduce the definition for second degree connections.[8]

Definition 2.2.1 (Second-degree Connection)**.** Inventor j is said to be inventor i's second degree connection if

1. Inventor i and inventor j are not directly connected ($w_{ij} = w_{ji} = 0$), and
2. There exists an inventor m such that inventor m is directly connected to both inventor i and inventor j ($w_{mj}, w_{jm} > 0$, $w_{im}, w_{mi} > 0$).

The intensity of the (indirect) relationship through inventor m between inventor i and inventor j in this case is given by the multiplications of the intensities of the relationships between inventor i and inventor m and inventor j and inventor m ($w_{im} \cdot w_{mj}$ and $w_{jm} \cdot w_{mi}$, respectively).

Building upon Definition 2.2.1, I can now generalize notion of collaboration intensity to second-degree connections. Specifically, let $\bar{w}ji$ denote the intensity of the (direct or

[7]The assumption that information can be acquired only through first- or second-degree inventors simplifies the notations and enhances the comparability with the empirical tests and the estimation conducted in the following sections. It, by no means, restricts the validity of the propositions. In Appendix B.2, I present an extension where I do not impose this assumption.

[8]This is usually referred to as a distance of two between inventor i and j.

indirect) relationship between inventor i and inventor j from inventor j's perspective. In the case where inventors i and j are directly connected, this weight equals the intensity of their direct relationship, denoted as wji. However, if inventors i and j are each other's second-degree connections, these weights are determined by the sum of all intensities across all the inventors who connect between them. Specifically, for all i and j it is given by,

$$\bar{w}_{ji} = \begin{cases} w_{ji} & \text{if } w_{ji} > 0 \\ \sum_{m=1}^{n} w_{jm} \cdot w_{mi} & \text{if } w_{ji} = 0 \end{cases}$$

In Figure 2.2, inventor 1 and inventor 2 are directly connected, while inventor 3 is inventor 2's second degree inventor, with inventor 1 connecting between them. The intensity of the indirect relationship in this case is equal to $\bar{w}_{23} = 0.35 \cdot 0.5 = 0.175$ and of the direct relationship is $\bar{w}_{21} = 0.5$.

The total information held by inventor i is the result of a combination between their initial knowledge and the information they acquire through first- or second-degree interactions with other inventors. Formally,

$$I_i = k_i + \sum_{j=1}^{n} \bar{w}_{ji} k_j \quad \forall i \tag{2.1}$$

where weighting the collaborators' knowledge by $\bar{w}_{ji}$ is meant to capture the idea that the information inventor i can learn from inventor j is proportional to a measure of the dependence between them. Intuitively, if inventor j spends only half of their time producing patents with inventor i, it is unlikely that inventor i can learn everything that inventor j knows in that time. Moreover, it implies that if inventor i and inventor j are only connected through inventor m, it should be unlikely for inventor i and inventor j to gather the same amount of information from inventor m, as inventor j has a direct access to them. Additionally, it imposes the implicit assumption that information cannot be gathered from both a direct and an indirect connection, and when inventors have both a direct and an indirect connections to an inventor on the network, the information transfers only through the direct connection. The idea is that a direct relationship provides superior access to information compared to an indirect one. Consequently, inventors are more inclined to act on the direct link rather than the indirect one.

Patent Production Function

Inventors produce patents in teams. Each inventor's total output relies on their own individual output, which is determined by their ability and the total information they hold, as well as the output contributed by their direct connections. Inventor i's output is, therefore, given by

$$
\begin{aligned}
y_i &= \alpha_i + I_i && \forall i \\
Y_i &= y_i + \sum_{j \neq i} w_{ji} y_j && \forall i
\end{aligned}
\tag{2.2}
$$

where y_i is inventor i's individual contribution, and it can be though of as the ability to produce solo-patents. Note the contribution of inventor i's collaborators is weighted by the time inventor j spends collaborating with inventor i, denoted as w_{ji}. It captures the idea that inventor j contributes more to the total output of inventors they spend more working on patents with.

This production function reflects the substitution between different sources that drive patent production. It emphasizes the tradeoff between the information inventors access through the network and the direct benefit the inventor gains from co-patenting, which comes in the form of the output contributed by their direct collaborators. This tradeoff will be the main focus of the next subsection.

2.2.2 Two Period Model

In general, patents are produced in various geographical locations. As a consequence, inventors may move around. In this part, I study the effect of a relocation on the total output of the mover's former collaborators in the eyes of the model. Specifically, the magnitude and the direction of this effect will be contingent on the specifics of the connections between these inventors, and on how their network changes in response to the relocation.

To start with, consider two time periods and two geographical locations. Let $t = 1$ represent the time before any relocation occurs, and $t = 2$ reflects the time after the relocation has taken place. Assume that the weight matrix $\mathbf{W}$ is fixed across time periods. I impose this assumption for compatibility with the empirical section where the intensity of the relationship between the inventors is measured only prior to the relocation and not after, especially since there are limitations in measuring based on what is observable.

Next, since the collaboration network can undergo changes between the two time periods, I introduce new notations which capture the state of the network in each one of these periods. Denote by $\mathbf{G}(t) \in \{0,1\}^{n\times n}$ the undirected adjacency matrix at time t. The ij-th entry represents the collaboration status between inventors i and j at time t, with the entry equal to one if they co-patent, and zero otherwise. This relationship is reciprocal. Additionally, let $\mathbf{S}(t) \in \{0,1\}^{n\times n}$ be the undirected information exchange network at time t. This is a symmetric matrix whose entries are equal to one whenever the inventors are engaged in information sharing. In the initial period, information sharing occurs exclusively when inventors co-patent.[9] However, after the relocation, in period $t = 2$, inventors who previously co-patented in $t = 1$ may cease their collaboration in period $t = 2$, and yet still engage in information sharing. Formally,

$$\begin{aligned} g_{ij}(1) = 1 \quad &\Longleftrightarrow \quad s_{ij}(1) = 1 \qquad\qquad \forall i,j \\ g_{ij}(2) = 1 \quad &\Longrightarrow \quad s_{ij}(2) = 1 \qquad\qquad \forall i,j \end{aligned}$$

Lastly, $I_i(t)$ is the level of information held by inventor i at time t, where the directed and second degree connections are now measured based on the entries of the network $\mathbf{S}(t)$.[10]

[9]This assumption follows by the empirical limitations. In the data I cannot observe any relationship between inventors that is not ties to co-patenting.

[10]In the first period, the elements of the matrices $\mathbf{S}(1)$ and $\mathbf{G}(1)$ are equivalent to the indicators $\mathbb{1}\left\{w_{ij} > 0\right\}$. Therefore, the direct and second degree connections on $\mathbf{W}$, $\mathbf{S}(1)$ and $\mathbf{G}(1)$ are the same. However, in the second period, since information sharing can take place even when the inventors do not collaborate, it does not necessarily hold.

Information Acquisition the Two Period Model

To accommodate some degree of continuity across the two periods, I impose the following assumption:

Assumption 2.2.1. The information inventors acquire in the first period cannot be forgotten and therefore, it is not subject to relearning.

This implies that inventor i begins period $t = 2$ with information level that is equal to $I_i(1)$, rather than k_i, as it is at the beginning of period $t = 1$. Intuitively, once techniques and ideas are acquired, they cannot be unlearned. Once learned, inventors can apply them again without relearning. In particular, equation (2.1) becomes

$$\begin{aligned} I_i(1) &= k_i + \sum_{j=1}^{n} \bar{w}_{ji} k_j && \forall i \\ I_i(2) &= I_i(1) + \sum_{j=1}^{n} \mathbb{1}\left\{\bar{w}_{ji}(2) - \bar{w}_{ji}(1) > 0\right\} \cdot \left[\bar{w}_{ji}(2) - \bar{w}_{ji}(1)\right] k_j && \forall i \end{aligned} \tag{2.3}$$

where $\bar{w}_{ji}(t)$ corresponds to the indirect weight based on the links in the matrix $\mathbf{S}(t)$, which are weighted by the matrix $\mathbf{W}$.[11] Note that the multiplication by the element $\mathbb{1}\left\{\bar{w}_{ji}(2) - \bar{w}_{ji}(1) > 0\right\}$ imposes the restriction that information can only be acquired in the second period, and cannot be lost.

Patent Production in the Two Period Model

Production in both periods follows the same reasoning as in equation (2.2), with output depending on both individual and scaled collaborative outputs. That is,

$$\begin{aligned} Y_i(t) &= y_i(t) + \sum_{j=1}^{n} g_{ji}(t)\, w_{ji} y_j && \forall i \\ y_i(t) &= \alpha_i(t) + I_i(t) && \forall i \end{aligned}$$

[11] Although the matrix $\mathbf{W}$ is fixed across the two time periods, the potential differences in the matrices $\mathbf{S}(2)$ and $\mathbf{S}(1)$ can lead to differences in $\bar{w}_{ji}(t)$ across the time periods, if new links are formed and/or old links are severed.

Define by $\Delta Y_i = Y_i(2) - Y_i(1)$ the change in inventor i's total output between the two periods.

Predictions

Let inventor i and inventor j be two inventors who are listed on at least one joint patent at time $t = 1$. And assume that inventor j relocates. In order to isolate the effect stemming directly from alternations in the patenting relationship with inventor j, assume that any change to inventor i's network involve adjustments related directly to inventor j.[12] To put it differently, my assumption is that the only connection of inventor i that could potentially change is the one with inventor j, and all other attributes, such as the abilities and information held by inventor i's other connections, remain constant across both time periods.

As a result of inventor j's relocation, the relationship between inventor i and inventor j can evolve in different ways, which will determine the sign of the effect.

Proposition 2.2.1. If, in period $t = 2$, inventor i and inventor j no longer collaborate or engage in information exchange with one another ($g_{ji}(2) = 0$ and $s_{ji}(2) = 0$), then the effect of the relocation on inventor i's output is negative ($\Delta Y_i = 0$). However, if information sharing takes place, the effect can be strictly positive.

Proof. See Appendix B.1.

This Proposition states that, ex-ante, the effect on inventor i's productivity is ambiguous. If inventor i and inventor j break their collaboration and cease any communication, then inventor i only bears the outputs costs to that are associated with the loss of a collaborator. This loss is proportional to inventor j's output in the first period.[13] On the other hand, the

---

[12]Empirically, I show that the number of inventors who use inventor j's old or new connections is very small, so I can abstract away from that in the model. However, I still address this possibility empirically.

[13]This alteration to inventor i's network resembles that of a death shock caused by an inventor's death. The result in this proposition is consistent with results in previous studies (see, for example, Azoulay *et al.* (2010) and Jaravel *et al.* (2018)) which utilize death shocks in innovative sectors to study the effect on the survivors and find negative effects.

effect that is driven by exposure to a new informational network operates in an opposite way to that of losing a collaborator. Therefore, as long as the information acquired through inventor j is valuable enough, it can counteract the negative effect due to the loss of inventor j's output and lead to a positive sign.[14]

This proposition underscores the significance of information sharing. It asserts that, in contrast to death shocks, the termination of collaboration between inventors does not necessarily imply they stop engaging in information sharing. Hence, the discontinuation of collaboration does not inevitably lead to a negative impact on the inventors who remain in the original location after their collaborator relocates. The sign of effect hinges on the tradeoff between the benefits from shared information and the potential cost of losing a collaborator.

The potential informational benefits depend on the identity of inventor j's new connections after the move, as outlined in the next proposition.

Proposition 2.2.2. Denote by $N_i(t)$ and $N_j(t)$ the set of inventor i's and inventor j's collaborators at time t, respectively. Holding the collaboration and information status between inventors i and j fixed, as well as the level of information inventor j gets access to after the relocation, if

1. The number of inventors j collaborates with after the move is fixed across two scenarios $\left(\left|\tilde{N}_j(2)\right| = \left|N_j(2)\right|\right)$, and

2. The number of new connections j makes that are not collaborators of inventor i prior to the move is larger under one of the scenarios $\left(\tilde{N}_j(2) \setminus N_i(1) \subseteq N_j(2) \setminus N_i(1)\right)$

then, the size of the effect under $N_j(2)$ is greater. In other words,

$$\Delta Y_i\left(N_j(2)\right) \geq \Delta Y_i\left(\tilde{N}_j(2)\right)$$

With strict inequality as long as $s_{ji}(2) = 1$.

[14] It is also important to note that if the inventors maintain their collaborative links, inventor i benefits from a weakly higher level of information and a weakly stronger collaborator. In that case, both effects operate in the same direction, leading to a positive effect.

Proof. See Appendix B.1.

This proposition states that, holding everything fixed but the connections between inventor j's new collaborators and inventor i's collaborators, a greater overlap between inventor i's collaborators and inventor j's new collaborators leads to a lower effect on inventor i's output. The intuition relies on Assumption 2.2.1, which posits that inventor i cannot relearn information they already possess. Consequently, when inventor j relocates, inventor i and their immediate connections remain connected, meaning that inventor i doesn't acquire new information through inventor j's new connections, provided that these connections are formed with inventor i's existing connections. However, in the scenario described in the proposition, as there is no impact on inventor j's individual output, when inventor j forms connections with inventors not directly linked to inventor i, inventor i gathers additional information. And this results in a higher output. This emphasizes that it is not just the act of sharing information that results in the potentially positive effect, but rather the ability to access new and previously unknown information.

Conversely, if a change in the output is associated with a common shock following, for example, a significant success in a patent on which both inventor j and inventor i are listed, then effect would be synchronous with the timing of the relocation but would not be triggered by any modifications in the network. This situation is equivalent to a shock affecting the innate abilities of both inventor i and inventor j, denoted as α_i and α_j, respectively. Consequently, it would not hinge on the identity of the inventors in the destination location, as Proposition 2.2.2 suggests.

Proposition 2.2.3. Let $\alpha_i(t), \alpha_j(t)$ be the innate abilities of inventors i and j at time t, respectively. And assume that $\alpha_i(2) > \alpha_i(1)$ and $\alpha_j(2) > \alpha_j(1)$. If the conditions in Proposition 2.2.2 hold, and the effect is driven solely by the changes in the inventors innate ability, then

$$\Delta Y_i\left(N_j(2)\right) - \Delta Y_i\left(\tilde{N}_j(2)\right) = 0$$

That is, the effect does not depend on the characteristics of the network inventor i gets access to.

Proof. See Appendix B.1.

The assumption that the effect is solely driven by a shock to inventors' abilities implies that the same amount of information is acquired in both scenarios. In that case, since the abilities are also equal across these cases, the effect should be the same.

Note that this proposition is not limited to that specific scenario, and a similar behavior should take place regardless of the which variation in the network I utilize.[15]

Another important feature of the collaboration network is the strength of the connections between the inventors. Given that both the amount of information acquired and the output produced through the network are scaled by these weights, it plays a crucial role in determining the magnitude of the effect.

Proposition 2.2.4. Let $\tilde{w}_{ji} > w_{ji} > 0$ be two weights. Assuming the status of collaboration and the status of information sharing remain constant in the second period under these two weights, the size of the effect under $\tilde{w}_{ji}$ is greater in absolute terms. In other words, $\left|\Delta Y_i\left(\tilde{w}_{ji}\right)\right| > \left|\Delta Y_i\left(w_{ji}\right)\right|$.

Proof. See Appendix B.1.

This proposition highlights that a more intense collaborative relationship between inventor j and inventor i leads to a more pronounced effect, measured in absolute terms. This follows since the intensity of the connection between the inventors scales the information gains and also the benefits or costs associated with continued collaboration or severance of the links. Therefore, if the difference between the information gained and the collaboration effect is negative, multiplying this difference by a larger number amplifies the negative effect. Similarly, if the overall effect is positive, scaling it by a larger number magnifies its magnitude.

[15] In the empirical section, I will use different variations such as the gender or distance of the move, which I do not model here.

2.3 Data and Descriptive Statistics

The dataset used for conducting the analysis is the product of a combination of two data sources, which together allow me to follow the innovative activity of inventors, as well as their locations within different economic areas in the United States over time. The first is the patent data from the US Patent and Trademark Office (USPTO). The second is provided by Revelio Labs, and it includes public employment information and other characteristics available on online professional profiles. The merge between these datasets opens up the opportunity to track inventors' employment history, including, but not limited to the firm they are employed at, the physical location, and the role they are taking within the structure of the firm.

2.3.1 Patent Data

The patent data covers all the U.S. patents granted between 1976 and 2021 and it was downloaded directly from PatentsView.[16] For each patent, this dataset includes information on the dates at which it was applied and granted, the individuals who were part of the team working on the patent, the firm to which the patent was assigned, the CPC class and subclasses, as well as backward and forward citations.[17,18] Furthermore, the city and the state the inventor resides in at the time of the application are also part of this dataset.[19]

---

[16]According to their website, PatentsView is "a collaboration between the USPTO, American Institutes for Research (AIR), University of Massachusetts Amherst, New York University, University of California, Berkeley, Twin Arch Technologies, and Periscopic which started in 2012." It allows a bulk data download of raw as well as disambiguated, or processed data on patents and applications applied to with the USPTO.

[17]CPC or Cooperative Patent Classification is a patent classification system. It was developed in a collaboration between the USPTO and the EPO (European Patent Office) in the goal of constructing a consistent classification system across these two entities. The CPC is in use since 2013, but older patents were given this classification retro-actively. It has five hierarchies, where each layer in the hierarchy refines the subject to which the invention relates to.

[18]Backward citations are usually referred to as the prior work a given patent cites as a relevant reference. While forward citations are the citations a given patent receives.

[19]This is the inventor's home address and not the location of the firm the patent is assigned too. The reason one would prefer to focus on the inventor's residence location is that the listed address for the assignee is usually the location of the headquarters, which might not be the the actual branch the inventor is employed at. Since I am interested in the physical location the inventor works at, it makes more sense to use their home

The raw data initially lacks consistent identifiers for patent inventors, making it challenging to accurately track and analyze inventor information. However, PatentsView provides a valuable dataset that has undergone a disambiguation process. This process involves assigning similarity scores to inventors using various algorithms, allowing for a reliable and standardized inventor identification. Additionally, PatentsView follows a similar procedure for assignees, ensuring that a disambiguated assignee information is consistently represented in the dataset. Using this dataset, I can create a panel of inventors over time, where each year includes information on all the granted patents the inventor applied for.

Since I follow relocations within the United States, I do not include any inventor whose home address was listed outside of the United States for any of the years 1976-2022.[20] There are about 1.2 million unique inventors in the final dataset.

2.3.2 Online Professional Profiles Data

This is an individual level database I received access to through a company named Revelio Labs. It is comprised of about 1.25B professional profiles and it provides their entire employment and education history as posted online by the end of 2022. The information in these profiles is made available and published by the individuals themselves. It includes, but is not limited to, the firm they are employed in, the role they take in the firm, and the institute at which they acquired education at that time. In general, the information is supplied in a panel structure such that for each firm-position combination, there exists a starting date and an ending date.

This is the universe of global professional profiles, and as such it covers all employees who have such a profile.[21]

---

address due to the likelihood of proximity between the two.

[20]For the purpose of this paper, and based on the the geographical units I utilize, I do not include the American territories of the United States.

[21]Thinking about the population as a whole. Since this data is collected from online professional profiles, it is possible that the data is being drawn from a non-representative sample of the population as a whole. However, as this paper focuses on high-skilled workers, with a very specific occupation – inventors – this issue should be less of a problem.

2.3.3 Data Construction

The ultimate goal is to create a panel dataset that encompasses yearly entries, providing crucial information on an inventor's location, employer, and patenting behavior. To achieve this, a merger between the patent data and Revelio Labs data is essential. This linking process is accomplished by utilizing the inventor's home address provided during the patent application. Focusing on inventors who applied for a patent for the first time after 1990, and considering online profiles became available around 2008, the linking procedure involves using the name of the inventor, the state they lived in, and the name of the assignee listed on the application.[22] The linking procedure is described in full details in Appendix B.3. This procedure results in approximately 300,000 successful matches, which represents about 30% of all US-based inventors who have exclusively lived within the US and have never obtained a patent before 1990.

Focusing on inventors applying for the first time after 1990 allows me to deal with the potential bias involved with the inventor's decision on which old employment information to include, as well whether to even open such a professional profile account. Since such profiles are usually more beneficial for individuals currently part of the labor force, it is less likely that a retiree or an individual close to retirement will have such an account. Building on Kaltenberg *et al.* (2023) who show an empirical evidence suggesting that most inventors apply for their first patent in their late 20s or early 30s, focusing on 1990 seems reasonable.[23]

Ultimately, the merged dataset effectively tracks inventors over time, providing annual records of their employers, job position, and the specific city and state of the establishment where they work. Additionally, the dataset includes valuable information on their patenting activity, such as the number of patents and citations received. The inventor's patenting behavior offers insights into both the quantity and quality of their innovative contributions.

[22]Although online profiles were introduced around 2008, it is up to the individual to decided on the employment history they feed in. An individual can decide to use of their employment history, dating back as far as possible, and they can also decide to focus on the more recent past.

[23]These inventors will be in their late 40s to 50s by the time online professional profiles became available. Therefore, it is likely that these individuals will still find it beneficial to have an online account.

The number of patents filed in a given year serves as a measure of quantity, while the number of citations received indicates the quality of the patents. However, it is essential to address potential biases and inaccuracies in measuring patent citations, as highlighted by Hall *et al.* (2001) and Lerner and Seru (2021). They pointed out that a simple citation count may be misleading due to changes in innovative activity over time and across CPC classes. To mitigate these issues, I adopt an adjusted measure for patent citations. This involves normalizing each patent's citation count by the average citation count for all other granted patents in the same year and CPC class. For the citation year, I refer to the application year of the cited patent.[24]

The inventor's location and firm are defined based on their employment history, ensuring accurate and reliable information in the dataset. During the calendar year, if an inventor changes employers or relocates, the employer and location information for each entry in the panel dataset are determined based on the maximum number of days spent in a specific year. This means that the employer entry will reflect the firm where the inventor was employed for the greatest number of days throughout the year, where ties are broken randomly. Similarly, the location entry will be based on the place where the inventor spent the majority of their time during that year, where, again, ties are broken randomly. This approach ensures that the panel dataset provides the most accurate and representative information regarding the inventor's employer and location for each year.

Appendix Table B.1 presents summary statistics related to patent activity and demographics for the final sample, comparing it with the full sample of inventors. In both sets of samples, the distribution of innovation activity is skewed towards higher values. This characteristic is evident in both the quantity and quality aspects of innovation. Moreover, a marginal distinction exists between the two samples, with inventors in the linked sample showcasing slightly higher average productivity than their counterparts in the full sample. This discrepancy could potentially be attributed to the fact that the linking tends to favor

[24] The reference year is the year the patents were applied for. That is, the average number of citation is calculated relative to the application date rather than the grant date.

inventors with more observations. This is a result of the linking process, which primarily considers inventors with higher accuracy rate.[25]

2.3.4 Movers, Left Behinds and Sample Construction

Following Moretti (2021), a move in this setting will be defined as a relocation between the U.S. Bureau of Economic Analysis' (BEA) "economic areas." These "economic areas" are determined by the patterns of labor commuting, effectively defining local labor markets. They consist one or multiple MSAs, which serve as hubs of economic activity, along with the adjoining counties that share economic interdependencies with them. Therefore, the distinction between "economic areas" is more pronounced, making it less likely for spillovers to naturally occur between these locations.

The United States comprises 179 economic areas that span the entire country, which can vary in size depending on their location. Notably, in larger regions like New York, Boston, or San Francisco, the economic areas tend to be larger than their corresponding MSAs as illustrated in Figure 2.3 (Johnson, 2004).

In each inventor-year observation, I determine the economic area by considering the location where the inventor states their professional work is conducted. Unlike the address provided in patent applications, which often pertains to the corporate headquarters, the city and state reported by the inventor on their professional profiles likely accurately represent the physical location of the office where they work. This method ensures that the economic area assigned to each inventor reflects their actual workplace location, providing a more precise and representative measure of the network of inventors they gain access to.

With the information regarding the inventors' locations over time, I identify a group of movers as those who have relocated between economic areas, and I designate the year of their move as the first year in their new location. By examining the employment history

[25]The scenario where the patent's assignee matches the employing firm of all inventors involved isn't universally applicable. Due to this, inventors associated with patents where the listed assignee differs from their employing firm won't be considered for matching. Conversely, if their patent history includes multiple instances, some of which under their employing firm, the likelihood of successfully linking them with their online profile increases.

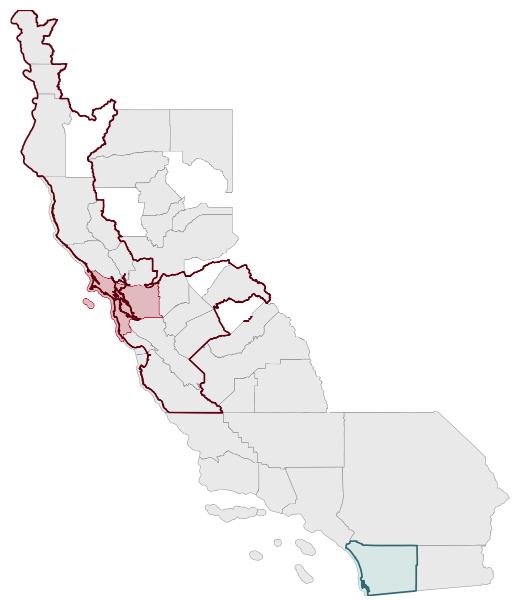


Figure 2.3: Economic Areas Examples: San Francisco and San Diego

Notes: This map illustrates two economic areas: San Francisco (indicated by a red outline) and San Diego (indicated by a green outline) within the state of California. Notably, in smaller cities like San Diego, the MSA (depicted in green) and the economic area (outlined in darker green) coincide. However, in larger cities such as San Francisco, the corresponding MSA (colored in red) is smaller than the corresponding economic area, delineated in bordeaux.

provided in professional profiles, I can identify an inventor as a mover even if they have not filed any patents in that specific year or at any point afterward, which is an advantage in my setting.[26] Overall, I identify 49,904 movers who moved at least once. If an inventor moved more than once, I only consider the first move.[27] Moreover, as the network of co-inventors is an important part of the analysis, I only consider moves after the first patent. If the inventor moved prior to the first patent, I ignore this move.

Table 2.1 reports the descriptive statistics of the moves. It shows that moves tend to

[26] If I had solely relied on patent data, this identification would not have been feasible. The patent data only includes observations on inventors when they apply for a patent, and as most inventors do not file for patents every year, this would have likely resulted in a delayed identification of the inventor's move year.

[27] As a robustness check, I repeat the analysis while including only inventors who relocated only once.

be associated with a more productive destination, as measured by different specification. Where cluster size in a given year is defined in Moretti (2021) as the number of inventors in an economic area X CPC class pair, excluding the mover, as a share of all inventors in that CPC class and year. Specifically, it shows that about 75 percent of the inventors are relocating to an areas that is associated with a higher productivity. Moreover, the mover tends to patent in the modal technology class in this location, as defined by the CPC class, with more than 50 percent satisfying this condition.

Table 2.1: Descriptive Statistics on Moves

Move Associated with a More Productive Location by	
Number of Patents in Location (%)	51.78
Weighted Number of Patents by Citations in Location (%)	51.85
Number of Patents in Location X Technology Class (%)	75.15
Weighted Number of Patents by Citations in Location X Technology	75.18
Number of Inventors in Location X Technology Class (%)	74.97
Cluster Size (%)	74.97
Patent in the Same Technology Class (%)	53.88
Patent in the Same Technology Class in the Origin Location (%)	57.09
Patent in the Same Technology Class in the Destination Location (%)	57.17
Moves within the Same Firm (%)	17.55
Continue Patenting after the Move (%)	52.84
Avg. Number of Patents at Time of Move (#)	0.72
Avg. Stock of Citations at Time of Move (#)	1.03
Avergae Distance of the Move (Miles)	980.59

Notes: This table includes information about the moves. It shows the percent of moves that are associated with a more productive destination, based on different measures of productivity, as well as, whether the mover tends to patent in the technology classes that are associated with this location.

In accordance with Jaravel *et al.* (2018), I establish a group of "placebo movers." These placebo movers consist of inventors who appear similar to those who actually relocated but, in reality, did not change their location and were not co-inventors of any of the movers. To achieve this, I implement an exact matching procedure, matching the movers based on specific criteria: the cumulative number of patent applications at the time of the move, the year of the first patent, the time of the move, and the CPC class of the last patent before the move. In case of ties, the matched inventors is picked randomly. Accounting for the total number of applications inventors applied for prior to the move allows me to control for some degree of productivity, while considering the year of the first patent serves as a measure of experience in patenting. Lastly, by utilizing matching based on the CPC class of the final patent before relocation enables me to mitigate potential biases that could arise from patenting activities associated with specific technological categories, or industry shocks. For instance, consider the IT revolution that unfolded in the early 2000s. If inventors specializing in IT-related domains were more inclined to relocate across economic areas, it might create a situation where the inventors left behind appear more productive due to the influence of this technological shift, rather than their connection to the mover.

Ultimately, I successfully find an exact match for 43,123 movers, accounting for roughly 87 percent of the total movers I identified in my dataset. In Table 2.2, I present the summary statistics of the real and placebo movers' characteristics at the time of their move. The real and placebo movers are perfectly balanced in terms of their first year of patenting, cumulative number of applications, year of the move, and CPC class, as per the matching procedure construction. According to the table, at the time of the move, real and placebo movers have applied for 2.63 patents on average, that were eventually granted. The average number of years between the move and the first patent is 5, while the average year of the move is 2012. Additionally, despite not being directly matched on these characteristics, the real and placebo movers also exhibit balance concerning the number of adjusted citations, averaging 2.75 for the real movers and 2.58 for the placebo mover, and gender distribution, with an average of 86 percent male among the real movers and also the placebo movers. Overall, the table demonstrates that both groups are also evenly balanced in various characteristics that encapsulate their innovation-related activities. This further strengthens the credibility of the matching process, a crucial factor in deriving accurate insights from the study's findings.

Next, I construct the co-inventor network for both real and placebo movers. This group comprises any inventor who has previously collaborated on a patent with either a real

Table 2.2: Summary Statistics on Real and Placebo Movers

	Real Movers				Placebo Movers			
	Mean	Median	Std. Dev	# Obs.	Mean	Median	Std. Dev	# Obs.
First Patent Year	2007	2008	9	43,123	2007	2008	9	43,123
Move Year	2012	2014	8	43,123	2012	2014	8	43,123
Patent Stock	2.63	1.00	3.12	43,123	2.63	1.00	3.12	43,123
Average Number of Patents Per Year	0.70	0.50	0.64	43,123	0.70	0.50	0.64	43,123
Adjusted Citations Stock	3.75	0.93	12.43	43,123	3.54	0.89	13.99	43,123
Average Adjusted Citations Per-Year	1.01	0.23	3.68	43,123	1.00	0.22	7.76	43,123
Male	0.86	1.00	0.34	38,991	0.85	1.00	0.36	39,471

Notes: This table presents summary statistics for both the real movers and the matched placebo group that I constructed. The variables are assessed at the time of the real or placebo move and encompass both the total count and the average values over the period before the move. What becomes evident from this table is that, even among variables that were not specifically matched, the real movers and the placebo group appear to be evenly balanced.

or placebo mover before their respective moves. These inventors are referred to as real and placebo "stayers," respectively. To ensure a reliable analysis, I exclude inventors who formerly co-patented with more than one real or placebo mover, leaving me with 23,553 real stayers and 15,401 placebo stayers. The summary statistics for these groups are presented in Table 2.3, which demonstrates the balance between these groups in terms of their patenting activity and characteristics.

On average, both real stayers and placebo stayers indicate their initial job on their online professional profile as 1997. Their first patent, on average, is applied for about 10 years later, and they experience the mover of their collaborates, on average, in 2015. Real stayers apply for an average of 6.54 patents before the move, with a cumulative total of 6.97 adjusted citations. In contrast, placebo stayers apply for an average of 7.10 patents prior to the move, accompanied by an average cumulative total of 7.59 adjusted citations. They are also balanced in term of gender, with 84 percent of real stayers and 85 percent of placebo stayers being male.

Remarkably, this balance is achieved despite not conducting the matching procedure at the stayers level. This implies that through an exact matching procedure, which was aimed to establish a foundation for identifying causal effects, I created a comparable control group which will help me to mitigate any biases that could arise from patenting activities and the timing of inventors' relocations. Through CPC class matching, I ensure that the placebo movers are patenting within the same technology category as the real movers. Consequently, any effect driven solely by the patenting category of the movers should be effectively neutralized, as the placebo stayers will be exposed to it in a manner similar to the real stayers.

Table 2.3: Summary Statistics Real and Placebo Left Behind Pre-Move

	Real Left Behind				Placebo Left Behind			
	Mean	Median	Std. Dev	# Obs.	Mean	Median	Std. Dev	# Obs.
First Year in Sample	1997	1998	10	23,553	1997	1998	10	15,401
First Patent Year	2007	2009	9	23,553	2007	2009	9	15,401
Move Year Mover	2015	2016	6	23,553	2015	2016	6	15,401
Patents Stock	6.54	3.00	11.66	23,553	7.10	3.00	13.41	15,401
Average Number of Patents Per Year	0.40	0.20	0.71	23,553	0.42	0.20	0.73	15,401
Adujsted Citations Stock	5.77	2.27	12.67	23,553	6.32	2.19	14.54	15,401
Average Adujsted Citations	0.36	0.13	0.79	23,553	0.38	0.13	0.88	15,401
Male	0.84	1.00	0.36	21,457	0.85	1.00	0.35	14,036

Notes: The information presented in this table offers summary statistics for both the real stayers and the placebo stayers group. An inventor is considered to be a "stayer" if they have collaborated with an inventor who eventually relocates. As such, the real (placebo) stayers are listed on a patent with real (placebo) movers before their respective moves. The variables are examined at the time of the real (placebo) move and cover both the cumulative count and the average values throughout the period prior to the move. Notably, it's important to observe that although I didn't specifically match characteristics of the stayers, the dataset seems to be well-balanced.

2.4 The Productivity of Left Behind Inventors

In this section, I outline the methodology utilized to estimate the average treatment effect of an inventor's relocation on the innovation activity of their co-inventors who remain in the original location and do not move themselves. This effect is identified through a difference-in-differences research design, where the control group consists of co-inventors who did not experience the relocation of their collaborators but share similar characteristics with the treatment group, as detailed in Section 2.3. By selecting the co-inventors in this manner, I address the potential concern that co-inventors who experience a move of their peers might be substantially different than those who did not experience a relocation of their collaborators. This approach helps mitigate any bias that could arise from productivity disparities between the two groups, enabling a more reliable analysis of the effects of the inventor's move on their co-inventors' productivity.

2.4.1 Dynamic Effects

Building on the identification strategy in Jaravel *et al.* (2018), I employ OLS regressions with a full set of leads and lags around the inventor's move. This approach enables me to study how the relocation influences the co-inventor's productivity over time. Furthermore, the methodology serves as a means to validate the research design by testing whether there is any observable effect of being left behind before the actual relocation event takes place, and by that addressing some of the concerns regarding a potential common shock. Specifically, using the constructed treatment and control group, I stack the moving events together such that the time index t represents the time relative to the move, which takes place at time $t = 0$, thereby addressing the concerns about estimation of pre-trends raised by Roth (2022). I then estimate the following OLS regression:

$$Y_{it} = \sum_{k=-9}^{9} \beta_k^{\text{Real}} \mathbb{1}_{\left\{L_{it}^{\text{Real}}=k\right\}} + \sum_{k=-9}^{9} \beta^{\text{All}} \mathbb{1}_{\left\{L_{it}^{\text{All}}=k\right\}} + \alpha_i + \alpha_t + \varepsilon_{it} \tag{2.4}$$

where L_{it}^{Real} are the leads and lags around the time of the relocation for the real stayers. Similarly, L_{it}^{All} are the leads and lags around the time of the relocation for both real and placebo stayers. $\left\{\beta_k^{\text{Real}}\right\}_{k=-9}^{9}$ and $\left\{\beta_k^{\text{All}}\right\}_{k=-9}^{9}$ are the predictive effects associated with the respective leads and lags, where k denotes the time relative to the move. I also include individual (α_i) and year (α_t) fixed effects.[28] To account for a possible serial correlation between inventors who are associated with the same mover, I cluster the standard errors at the mover level as in Jaravel *et al.* (2018).

If the move is as good as an exogenous variation, the coefficients $\left\{\beta_k^{\text{Real}}\right\}_{k=-9}^{9}$ identify

[28]These fixed effect also include the time elapsed since the first patent, i.e., experience in the patenting innovation.

the causal effect of experiencing a relocation of a co-inventor k years relative to the year of the move. If being experiencing the relocation of a co-inventor results in an increase in the inventor's productivity k years relative to the move, β_k^{Real} should be positive. Conversely, if it negatively affects the inventor's productivity, the coefficient should be negative, and if there is no effect, it would be zero. However, an identification concerns that needs to be addressed is selection bias. If the inventors who experienced the relocation of their co-inventor display, on average, higher productivity compared to the overall population of inventors, then the estimated effect might not reflect true causality. To mitigate this issue, I use a control group of inventors who share similarities with those experiencing a relocation of a co-inventor but did not experience that themselves. In the analysis below, I demonstrate that there is no evidence of statistically significant pre-trends, effectively addressing this concern. Specifically, the results indicate that the lag terms β_{-9}^{Real} to β_{-1}^{Real} are not statistically significant before the move occurs. This finding suggests that prior to the move, the inventor's productivity is not influenced by the future move, thus supporting the valid causal interpretation of the estimates.

In Figure 2.4, I report the point estimates derived from regression equation (2.4). No pre-trends are reported and the point estimates seem to be downward sloping. It also reveals that the enhancement in the productivity of inventors who experienced the relocation of their co-inventors, relative to the placebo group, exhibits a positive effect. This improvement becomes statistically significant approximately three years after the initial event when considering the annual number of patents and the weighted metric of annual number of patents, termed adjusted citations. Also, although the effect weakly increases over time, it appears to stagnate around six years after the move.

This outcome aligns logically with the notion that there is a time lag involved. It requires some duration for the relocating inventor to amass new knowledge and subsequently transmit it to their former collaborators in the origin location. Moreover, the delayed adjustment in citations can be attributed to the usual pattern where citations trail behind patents.[29]

2.4.2 Baseline Regression

As a way of summarizing the results, I utilize a second specification that directly compares the pre-period and post-period to estimate the average treatment effect of the move, instead of examining the effect over time. In this alternative approach, the dummy variables $PostMove_{it}^{\text{Real}}$ and $PostMove_{it}^{\text{All}}$ turn one after the time of the real or placebo move, respectively. This allows for a more concise assessment of the move's impact without focusing on the dynamic changes over time. This specification is given by:

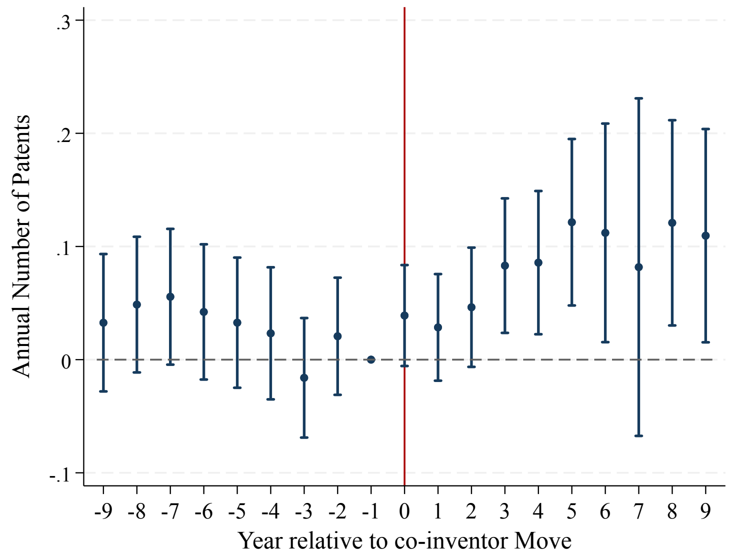
$$Y_{it} = \beta^{\text{Real}} PostMove_{it}^{\text{Real}} + \beta^{\text{All}} PostMove_{it}^{\text{All}} + \alpha_i + \alpha_t + \varepsilon_{it} \tag{2.5}$$

I also include individual and fixed effects, as before. And cluster the standard errors at the mover level.

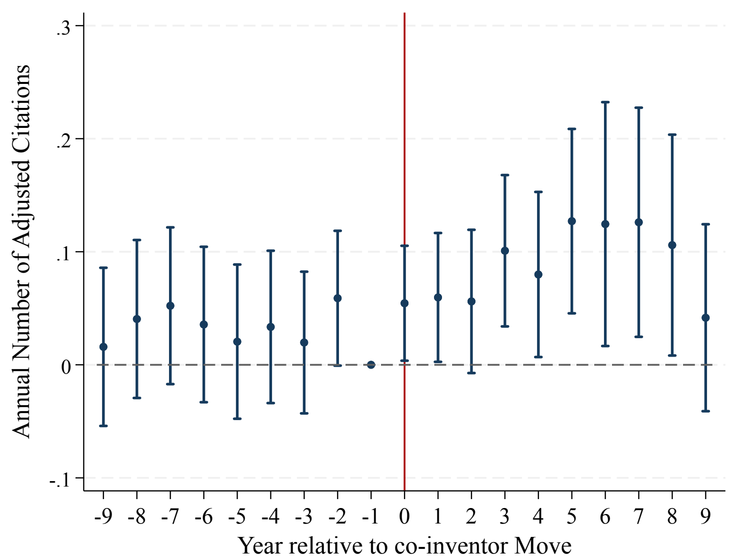
Table 2.4 presents the outcomes of the regression analyses. The results indicate a positive and statistically significant impact on both the annual number of patents and the annual number of citations. In column (1), the coefficient β^{Real} equals 0.045 and is statistical significance. This signifies that in comparison to the placebo stayers, those who truly experienced the relocation of their co-inventor tend to exhibit a higher productivity. When compared to the average annual patents post mean for the control group (0.5), this translates to approximately a 9 percent rise in the number of patents per year compared to the control group.

Similarly, as seen in column (2), the coefficient β^{Real} stands at 0.051 and is highly statistically significant. This suggests that, relative to the placebo group of inventors, real stayers generate a higher annual number of weighted-patents on average, as implied by the number of annual adjusted citations. In percentage terms, this signifies more than 15

[29] I use the date of patent application, not the date of patent granting. So while it's generally observed that patent application timing aligns closely with R&D (as indicated by studies such as Griliches (1998) and Griss (1993)), my findings do not conflict with this notion, because the mover should logically not engage in R&D in the new location before the move takes place. Consequently, both sets of results can indeed remain valid, and patent and citation counts can take time to respond.



(a) Annual Number of Patents



(b) Annual Number of Adjusted Citations

Figure 2.4: Dynamic Effects

Notes: This plot presents the effect of an inventor's move on their collaborators' productivity. The vertical lines represents a 95% confidence interval, while standard errors are clusters at the mover level. The dependent variables are the annual number of patents and the annual number of adjusted citations, which are defined in Section 2.3.

percent rise in the annual number of adjusted citations compared to the placebo stayers group.

Table 2.4: Baseline Regression Results

	(1) Annual Number of Patents	(2) Annual Number of Adjusted Citations
$PostMove^{\text{Real}}$	0.045**	0.051***
	(0.019)	(0.018)
Control Post Mean	0.5	0.34
Percentage Change	+8.95%	+14.92%
Observations	555815	555815
Individual FE	Yes	Yes
Year FE	Yes	Yes
Experience FE	Yes	Yes

Notes: The information presented in this table presents the results from regression equation (2.5). The unit of analysis in these regressions is inventor-year. The dependent variable in column (1) is the number of patents per year and in column (2) is the number of adjusted citations per year, as defined in Section 2.3. Standard errors are clustered at the mover level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

A second identification concern relates to a common shock, involving unobserved time-varying productivity shocks at the mover level that could lead to more productive stayers. If the mover's opportunity to relocate is influenced by their prior work and success, and the stayer is associated with any of these patents, there is a possibility that the stayer, despite not moving themselves, experiences increased productivity due to the success of these patents and not directly due to the move. This increased productivity may be linked to factors like greater ease in obtaining research grants, which allows them to access better equipment and additional resources. Consequently, the observed effect might be driven by the joint previous success rather than the mere connection to someone who moved. To address this concern, I conduct a series of heterogeneity tests to account for potential confounding factors and ensure the validity of the estimates in the next section.

2.4.3 Additional Results and Robustness Checks

Within vs. Across Firm Moves. Although inventors within the same company are typically encouraged to collaborate on patents, it is conceivable that when a relocation involves switching employers, these inventors may no longer have the opportunity to co-patent with each other. If this scenario holds true, it could create obstacles to collaboration and, potentially, hinder the exchange of information. To address this possibility, I have confirmed that my results are robust to within and across firm moves, and the results are presented in Appendix Table B.4.

Bad vs. Good Moves. Inventors relocate to various locations, and the characteristics of the location they move to might generate consequences for their co-inventors who remain in the origin location. Research indicates that relocating to a bigger innovation cluster leads to increase in the productivity of the inventor making the move (Moretti, 2021). Building upon this finding, I show that when the mover relocates to a bigger innovation cluster, defined by a higher concentration of inventors patenting in the same model CPC class, the effect of the productivity of those who stay is larger in a statistically significant way. This implies that the effect is contingent on the nature of the relocation and whether it provides opportunities for heightened patenting activity. The results are presented in Appendix Table B.6.

Additional Robustness Checks. In Appendix B.4, I report additional robustness checks showing that the results do not depend on the type of firm the inventors move to, and are not driven by movers who relocated more than once.

2.5 Mechanism: New Information as the Driving Force

In this section, I show that the long-lasting positive effect on the productivity of the mover's co-inventors is a result of the opportunities created by the mover for the stayers to access new information. First, I rule out that the effect is driven by a common shock, such as access to more resources. Second, I show that the effect is not solely driven by cases where the stayers substitute the mover for their associated mover's former collaborators. That is, it is not driven by having more opportunities to collaborate with the former collaborators of the mover. Third, I demonstrate that the effect is primarily attributed to the acquisition of new information, rather than information in general. I show that the stayers cite patents produced in the destination location more extensively, and evidence on network expansion into the destination location, after the move. I further report that when the mover switches to a different CPC class, the stayers are more likely to follow them. I proceed by highlighting the asymmetric nature of this effect, particularly regarding inventors who shared a history of extensive collaboration with the mover prior to the relocation. As anticipated, this subgroup experiences a more pronounced effect, owing to the potential likelihood of maintaining some form of relationship in these cases. I conclude by showing two facts: first, I show that the effect remains positive even when I exclude patents produced in collaboration with the mover, indicating that stayers acquire skills that contribute to their independent work. Second, I show that the effect is predominantly prevalent in cases where the mover relocates to a location where the stayer has no prior collaborators.

2.5.1 Ruling Out the Common Shock Reasoning

Heterogeneity by Sex Differences. The research design I employ, while addressing some potential bias, doesn't completely eliminate the common shock concern. This concern arises when the effect is influenced by individual-level changes over time, such as access to more resources. For instance, these could be resources that were previously occupied by the mover or those resulting from a successful patenting activity for which the mover was "rewarded" with a relocation opportunity. To address this concern I leverage patterns that common shocks can not produce. I do so by showing that the size of the effect depends on similarities between the mover and the stayers that are not thought to be correlated with success, but rather are similar in nature to homophily.[30]

[30]An extensive body of literature has emphasized the significance of homophily, a concept characterized by individuals forming relationships with others who share similar attributes. Numerous studies underscore the role of homophily in forging robust connections (see McPherson *et al.* (2001) for a survey of the literature

Table 2.5: Heterogeneity by Sex Differences

	Same Sex		Opposite Sex	
	(1) Annual Number of Patents	(2) Annual Number of Adjusted Citations	(3) Annual Number of Patents	(4) Annual Number of Adjusted Citations
$PostMove^{\text{Real}}$	0.059***	0.058***	-0.011	0.020
	(0.022)	(0.021)	(0.037)	(0.029)
Control Post Mean	0.501	0.349	0.496	0.304
Percentage Change	+11.72%	+16.7%	-2.28%	+6.42%
P-Value H_0: Diff. $= 0$			0.09	0.26
Observations	364681	364681	109289	109289
Individual FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Experience FE	Yes	Yes	Yes	Yes

Notes: The information presented in this table presents the results from regression equation (2.5) on two different samples. Columns (1) and (2) correspond to cases where the mover and the left behind are of the same sex, and columns (3) and (4) cover the cases where the mover and the left behind are of opposite sexes. The unit of analysis in these regressions is still inventor-year. The dependent variable in columns (1) and (3) is the number of patents per year and in columns (2) and (4) is the number of adjusted citations per year, as defined in Section 2.3. Standard errors are clustered at the mover level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

The findings presented in Table 2.5 report differential effects which are contingent on shared characteristics between the mover and the stayer. It is evident that when the mover and the stayers are of the same sex, the effect is positive and statistically significant. This effect implies that compared to the placebo stayers, the real stayers experience an average of about 12 percent increase in the annual number of patents and almost 17 percent increase in the annual number of adjusted citations. On the other hand, when considering the cases where the mover and the stayer have opposite sexes, the effect becomes negative and statistically insignificant. I can reject that the coefficients are the same at the 10 percent level.[31]

These results dispel concerns related to common shocks, as there is no basis to assume that a common shock would disproportionately impact inventors based on both their own

in sociology). See also Currarini *et al.* (2009), Currarini *et al.* (2010) and Bramoullé *et al.* (2012) who model the origins of homophily.

[31]These outcomes align with the findings in Cullen and Perez-Truglia (2023) which indicate that positive outcomes are observed when men interact with men, but no similar effect is found for women interacting with women.

sex and that of their respective movers. If the effects were exclusively driven by a common shock, the logical expectation would be for the results to remain consistent across all sex combinations between the mover and the stayer, as Proposition 2.2.3 suggests. However, the findings in Table 2.5 demonstrate that this is not the scenario.

I find qualitatively similar results when conditioning on the race similarities between the mover and the stayer. I report these results in Appendix Table B.7.

Heterogeneity by the Distance of the Move. Another evidence that the effect is not driven by a common shock comes in the form of the asymmetric effect of the distance between the mover and the stayer after the move.[32]

Specifically, I calculate the distance between the actual mover and the real stayer by utilizing the location coordinates of the city where the inventor's workplace is situated.[33] It is defined as the distance between the mover's destination and the location of the stayer. For the placebo mover, I impose the same distance between the origin and the destination of the mover they match to.

The outcomes presented in Table 2.6 outline the estimation findings obtained by splitting the dataset based on the distance the mover travels in their relocation. The threshold for determining the distance between the mover and the left behind is set at the 50th percentile of the distance distribution, which is equivalent to 714 miles.[34]

The table indicates that in this scenario, the significance of network novelty outweighs physical proximity. The effect is substantial and statistically significant, especially when the inventor relocates to a significantly more distant location. It reveals that in comparison

[32]It corresponds to the geographical distance between the mover's initial location (and thus the location of the stayer) and the mover's new destination. In simpler terms, it represents the spatial gap between the mover and the stayer subsequent to the relocation.

[33]The coordinates correspond to the nearest city listed by the inventor as their workplace location. These coordinates indicate a random location within that city, and are persistent across cities. These coordinated align with the coordinates provided by Google Maps.

[34]For a better understanding, consider the following comparisons: the distance between Boston and Detroit spans 613 miles, while Boston to Chicago covers 851 miles. Boston to New York is a distance of 191 miles, whereas Boston to San Francisco spans an extensive 2699 miles.

Table 2.6: The Heterogeneity by Distance from the Mover

	Short Distance (Bottom 50%)		Long Distance (Top 50%)	
	(1) Annual Number of Patents	(2) Annual Number of Adjusted Citations	(3) Annual Number of Patents	(4) Annual Number of Adjusted Citations
$PostMove^{\text{Real}}$	0.029	0.027	0.061^{***}	0.073^{***}
	(0.033)	(0.028)	(0.020)	(0.021)
Control Post Mean	0.519	0.348	0.481	0.333
Percentage Change	+5.5%	+7.89%	+12.68%	+21.98%
P-Value H_0: Diff. $= 0$	0.398	0.195		
Observations	279200	279200	276615	276615
Individual FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Experience FE	Yes	Yes	Yes	Yes

Notes: The information presented in this table presents the results from regression equation (2.5) applied to two distinct subsets of data. The first subset corresponds to left behind inventors who are located in a closer proximity to their respective mover, and the second corresponds to left behind inventors who are located at a greater distance form their respective mover. Close proximity is defined as distance that is shorter than 714 mile (1149 km). The distance between the placebo left behind and placebo mover is defined to be equal to the distance between the real inventor and the real left behind they are matched to. Columns (1) and (2) pertain to scenarios where the left behind and the mover are close to each other, and columns (3) and (4) delve into cases where they are located at a greater distance. The outcome variable in columns (1) and (3) is the number of patents per year, while columns (2) and (4) utilize the number of adjusted citations per year, as elaborated in Section 2.3. Standard errors are clustered at the mover level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

to the placebo stayers, real stayers experience an about 13 percent increase in their annual number of patents and a 22 percent increase in their annual number of adjusted citations post move.[35] This result might seem to stand in a contradiction to the extensive research that underscores the significance of face-to-face interaction and proximity for fostering productive and successful collaboration (Battiston *et al.*, 2021b; Emanuel *et al.*, 2023). However, following Saxenian (1994) who suggests that mere proximity is insufficient for information exchange among inventors and that collaboration is necessary, it is possible that the prerequisite of an existing prior connection, which is used in order to form the foundation for potential collaboration, can compensate for the absence of later physical proximity. Not to mention that a distance location makes it more likely for new information to flow.

[35] I find similar results for alternative thresholds such as the top and bottom 10th percentiles or the top and bottom 25th percentiles are chosen.

Given that distance should not be correlated with common shocks, these results provide additional evidence against the common shock explanation. If the entire effect were attributed to a common shock, the magnitude and presence of the effect should not vary based on the distance of the city to which the mover relocates.

2.5.2 Ruling Out Firm and Network Effects

Substitution Effect. Another explanation to the positive effect estimated may be that the stayers can substitute collaboration with the mover by collaborating with the mover's former collaborators. To test whether this is the force that drives the result I split the sample into two, based on whether the stayers collaborate with some of the mover's former collaborators prior to the move.[36] The results reported in Table 2.7 show that, although the effects when some replacement exists are larger in a statistically significant way, the effects are still positive and statistically significant when there is no evidence of replacement. Moreover, the effects seem to be relatively similar in size to the overall effect, and the group of inventors who experience some replacement is relatively smaller, suggesting that replacement is unlikely to be the driving force of the effect.

Network Effects. To investigate whether the diffuse of network effects is an important channel, I consider the group of real and placebo second degree connections. The results reported in Table 2.8 are obtained form specification (2.5) and show that the relocation has no significant effect on the productivity of their second degree connections. Hence, suggesting that network diffusion does not take place.

2.5.3 The Effect is Driven by the Access to New Information

In this subsection, I provide evidence that the improvement in innovation productivity resulting from the relocation of a co-inventor is primarily influenced by new information that the mover exposes the stayer to. I do that in three steps. In the first step, I provide

[36] These type of collaborations might include the mover as part of the team.

Table 2.7: Heterogeneity based on Replacement

	No Replacement		Some Replacement	
	(1) Annual Number of Patents	(2) Annual Number of Adjusted Citations	(3) Annual Number of Patents	(4) Annual Number of Adjusted Citations
$PostMove^{\text{Real}}$	0.042**	0.045**	0.169	0.315**
	(0.020)	(0.018)	(0.144)	(0.135)
Control Post Mean	0.483	0.329	1.159	0.801
Percentage Change	+8.79%	+13.69%	+14.6%	+39.37%
P-Value H_0: Diff. $= 0$	0.382	0.047		
Observations	544529	544529	11286	11286
Individual FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Experience FE	Yes	Yes	Yes	Yes

Notes: The information presented in this table presents the results from regression equation (2.5) applied to two distinct subsets of data. The first subset corresponds to stayer who started co-patenting with the mover's prior collaborators after the mover, and the second corresponds to the stayers who did not. The mover's former collaborators are defined as inventors their patented with before the relocation. A stayer experiences some replacement if, after the move, they collaborate with at least one of the movers' former collaborators, with whom they have never collaborated with prior to the move. Columns (1) and (2) pertain to scenarios where the stayer experiences no replacement, and columns (3) and (4) delve into cases where they do. The outcome variable in columns (1) and (3) is the number of patents per year, while columns (2) and (4) utilize the number of adjusted citations per year, as elaborated in Section 2.3. Standard errors are clustered at the mover level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

evidence that information sharing between the mover and the stayer takes place. I show that the stayer is more likely to cite patents produced by inventors in the destination location or to collaborate with inventors in the destination location after the move. I also show that when the mover changes the main CPC class they patent in, the stayer is also more likely to patent in the same patent class. And finally, I show that the increase in the productivity is also positive and statistically significant when I exclude the patents on which the stayer and the mover collaborate on, suggesting that the information acquired by the stayer is used even on patents the mover does not collaborate with them on.

As a second step I show that the effect is larger when the mover and the stayer have a more intense relationship prior to the move. These are the inventors who are more likely to engage in information sharing, and this result offers a support for that.

In the last step, I show that the effect is larger when the stayer does not have any prior

Table 2.8: Second Degree Inventors

	(1) Annual Number of Patents	(2) Annual Number of Adjusted Citations
$PostMove^{\text{Real}}$	0.018	0.019
	(0.026)	(0.023)
Control Post Mean	0.658	0.447
Percentage Change	-0.78%	4.3%
Observations	349693	349693
Individual FE	Yes	Yes
Year FE	Yes	Yes
Experience FE	Yes	Yes

Notes: The information presented in this table presents the results from regression equation (2.5) on the sample of second degree co-inventors. The dependent variable in column (1) is the number of patents per year and in column (2) is the number of adjusted citations per year, as defined in Section 2.3. Standard errors are clustered at the mover level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

collaborators in the location the mover relocates to. This result suggest that it is the access to new information that drives the results rather than information in general.

Share of Citations Made to Mover's Destination. Citations are considered a valuable measure to account for knowledge spillovers. Citations are references made to prior work that is relevant for the current patent scope and was filed before the citing application's filing date. An elevated proportion of citations to patents produced in the mover's destination indicates a significant exposure to information in that location.

To examine how the citation behavior of the stayers changes after the mover's relocation, I calculate the share of citations the stayers attributed to patents located in the mover's destination. A patent is said to be produced in a certain location if at least one inventor who is listed on the patent is located there at the time of the application.

The findings in Table 2.9 display the outcomes when considering the share of citations attributed to patents located in the mover's destination. For the placebo movers, I enforce the destination of the real mover they are matched to.

The results indicate that, regardless of whether the inventors patent in a given year

Table 2.9: The Effect of Relocation on Citations to the Destination

	(1) Annual Percentage of Citations in Destination	(2) Annual Percentage of Citations in Destination
	All	Only When Patenting
$PostMove^{\text{Real}}$	0.047^{**}	0.247^{***}
	(0.023)	(0.095)
Control Post Mean	0.436	2.065
Percentage Change	+10.71%	+11.95%
Observations	555815	142296
Individual FE	Yes	Yes
Year FE	Yes	Yes
Experience FE	Yes	Yes

Notes: The information presented in this table presents the results from regression equation (2.5). In column (1), I consider all the observations in the panel, whether or not the inventor patented in that year, while column (2) covers only the years when the inventor patents. The unit of analysis in these regressions is still inventor-year. The dependent is the share of citations given to patents with at lease one inventor in the mover's destination. Standard errors are clustered at the mover level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

or not, there is a notable increase of approximately 11 percent to 12 percent in the share of citations directed toward inventors situated in the mover's destination, relative to the placebo group.

Share of Collaborators in Mover's Destination. Another way to gauge the interaction between the stayer and the mover, after the relocation, is by examining the proportion or count of collaborators the stayer has in the mover's destination. If the collaboration patterns between the stayer and other inventors in the mover's destination undergo changes after the move, it implies the existence of collaboration spillovers. This suggests that the relocation of a former collaborator to that location opens up opportunities for network expansion, which can be highly valuable, both at the individual level and for the firm involved, as these spillovers can lead to new and beneficial collaborative opportunities, fostering innovation and knowledge exchange beyond the direct interactions between the mover and the stayer.

I calculate the annual share of collaborators located in the destination per year, and use it as the dependent variable in regression equation (2.5).[37] Here, I, again, assign the placebo mover the destination of the real mover they are matched to.

Table 2.10: The Effect of Relocation on Share of Collaborators to the Destination

	(1) Annual Percentage of Collaborators in Destination	(2) Annual Percentage of Collaborators in Destination
	All	Only When Patenting
$PostMove^{\text{Real}}$	1.176***	5.318***
	(0.096)	(0.219)
Control Post Mean	1.525	7.229
Percentage Change	+77.10%	+73.56%
Observations	555815	142296
Individual FE	Yes	Yes
Year FE	Yes	Yes
Experience FE	Yes	Yes

Notes: The information presented in this table presents the results from regression equation (2.5). In column (1), I consider all the observations in the panel, whether or not the inventor patented in that year, while column (2) covers only the years when the inventor patents. The unit of analysis in these regressions is still inventor-year. The dependent variable is the share of collaborators in the destination location. Standard errors are clustered at the mover level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

The results presented in Table 2.10 indicate that, in comparison to the placebo stayers group, the real stayers experience an increase in the annual share of collaborators they have in the destination location ranging from 74 percent to 77 percent.

New Technology Classes. Another indication of information flows comes in the form of the knowledge needed in order to patent in a certain CPC class. When the mover starts patenting in a new CPC class after the move, the stayers can either start patenting in this CPC class as well, or to keep patenting in the CPC class they are familiar with. I define the CPC class as the model CPC class that the inventors patent in during that year. I look at the effect on the likelihood of the stayers to patent in the same CPC class as the mover, conditional on the mover changing theirs.

[37] Similar results are observed if, alternatively, I calculate the share of collaborators in the destination per patent and then average the values across all patents in the same year.

Table 2.11: Patenting in a Different Technology Class

	(1) Patent Class Same as Mover Post Move
$PostMove^{\text{Real}}$	0.052***
	(0.008)
Control Post Mean	0.4
Percentage Change	+13.11%
Observations	262490
Individual FE	Yes
Year FE	Yes
Experience FE	Yes

Notes: The dependent variable in this table is an indicator of whether the stayer patents in the same CPC class as the mover. The sample is restricted to include only the cases where the mover changed their patenting class after the move. The results are obtained from regression equation (2.5).Standard errors are clustered at the mover level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

The result reported in Table 2.11 show that the real stayers are, on average, 13 percent more likely to change their technology class to be the same one as the mover after the move, compared to the control group. And this result is statistically significant.

This result suggest that there is information flow between the mover and the stayer, even when the information is not necessarily relevant to the CPC class that the stayers used to patent in.

Excluding Patents with the Mover. One might be worried that the estimated effect is solely driven by the increase in the productivity of the mover. If the mover and the stayer maintain their collaborative relationship, and the mover becomes more productive, this can lead to an increase in the productivity of the stayer through team production, as suggested by the model. To address this concern, I examine the productivity effects of the real stayers when I exclude the patents the stayer produces in collaboration with the mover. In that case, I estimate the effect on the annual number of patents and annual number of citations the stayer produces without the mover, which serves as a measure of productivity that is uncorrelated with the mover's productivity.

Table 2.12: Baseline Specification Excluding Patents Produced with the Mover

	(1) Annual Number of Patents	(2) Annual Number of Adjusted Citations
$PostMove^{Real}$	0.035^{**}	0.039^{***}
	(0.016)	(0.015)
Control Post Mean	0.369	0.259
Percentage Change	+9.47%	+15.17%
Observations	555815	555815
Individual FE	Yes	Yes
Year FE	Yes	Yes
Experience FE	Yes	Yes

Notes: The information presented in this table presents the results from regression equation (2.5), when I exclude all the patents the stayer co-patented with the mover. Column (1) reports the effect on the annual number of patents, while column (2) shows the effect on the annual number of adjuated citations. The unit of analysis in these regressions is still inventor-year. The dependent variable is the share of collaborators in the destination location. Standard errors are clustered at the mover level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

The results in Table 2.12 show that there is a positive and statistically significant effect on both the annual number of patents and the annual number of adjusted citations. This implies that even when considering only the patents are not produced in collaboration with the mover, the real stayers are more productive. One explanation to these results can be that the real stayers acquire information from the mover than they can later apply in their own work, even when the mover does not participate.

Heterogeneity by the Intensity of the Collaboration. To account for how the intensity of the collaboration between the stayer and the mover affect the results, I rum separate regression conditioning on this characteristic. Specifically, I calculate the number of patents that the mover and the staying inventor are jointly listed on. The frequency of collaboration between inventors can indicate their reliance on each other and therefore, frequent collaborators are more likely to maintain their relationship outside of the office setting, enabling them to sustain their connection even when not co-located. If that is indeed the case, the model predicts that effect should be stronger.

To capture the intensity of the connection, I create a dummy variable (*StrongLink*) which takes the value one for the top 50th percentile of the distribution of joint patents prior to the move, indicating a strong connections, and zero for the bottom 50the percentile, representing weak connections.[38] Table 2.13 presents the results for two separate regression analyses: one for staying inventors with strong links and another for those with weak links.

Table 2.13: Effect Size and The Strength of the Link

	Strong Link		Weak Link	
	(1) Annual Number of Patents	(2) Annual Number of Adjusted Citations	(3) Annual Number of Patents	(4) Annual Number of Adjusted Citations
$PostMove^{\text{Real}}$	0.314**	0.294**	0.027	0.030
	(0.154)	(0.121)	(0.023)	(0.021)
Control Post Mean	0.971	0.483	0.438	0.305
Percentage Change	+32.36%	+60.86%	+6.07%	+9.85%
P-Value H_0: Diff. $= 0$			0.06	0.03
Observations	18380	18380	393254	393254
Individual FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Experience FE	Yes	Yes	Yes	Yes

Notes: The information presented in this table presents the results from regression equation (2.5) on two different samples. Columns (1) and (2) correspond to cases where the mover and the left behind are connected through a strong link, and columns (3) and (4) cover the cases where the mover and the left behind are connected through a weak link. A strong links is defined as collaborating on more than one patent prior to the move. The unit of analysis in these regressions is still inventor-year. The dependent variable in columns (1) and (3) is the number of patents per year and in columns (2) and (4) is the number of adjusted citations per year, as defined in Section 2.3. Standard errors are clustered at the mover level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

The results reported in Table 2.13 indicate that the effect is large and statistically significant when the connections between the mover and the stayer are strong. Conversely, for weak links, the effect diminishes and is no longer statistical significant.

Access to a New Network. I presented compelling evidence supporting the notion that the observed increase in productivity is, in part, influenced by knowledge spillovers and network expansion. Now, I delve further into the analysis and demonstrate that these

[38] The reason I pick the top and bottom 50th percentile is that the 50th percentile is exactly one joint patent between the mover and the left behind. Since one joint patent is a requirement by the definition of being left behind, any other split will only limit the number of observations for the group with strong links.

spillovers have a significant impact primarily when the staying inventor is exposed to a new location. In other words, I show that the effect on the productivity of a staying inventor who already has collaborators in the mover's destination prior to the move is considerably smaller. This finding suggests that it is the access to the information in a new location that plays a crucial role in influencing productivity, rather than the relocation itself and the exposure to inventors in general.

In order to account for the collaboration patterns in the mover's location, I create a dummy variable (*NewNetwork*). This variable takes the value one for the top 10th percentile of the distribution of the share of collaborators in the mover's destination prior to the move, indicating that the inventor gains access to a new network through the mover. Conversely, it takes the value zero for the bottom 10th percentile, indicating that the inventor did not access a new network through the mover.

Table 2.14 presents the results for two separate regression analyses: one for inventors who gain access to a new network through the mover and another for those who did not, as defined by the *NewNetwork* dummy variable. The estimation indicates that the effect is primarily noticeable among inventors who lack collaborators in the destination location before the move. In particular, when compared to the placebo left behinds, the real left behind experience a 10 percent increase in the annual number of patents and a 21 percent increase in the annual number of adjusted citations following the move. Conversely, when an inventor has previously gained access to the network in the destination location through collaborations with inventors already situated there before the move, the effect is diminished and lacks statistical significance.

These results also contradict the common shock concern, as the common shock would have affected the staying inventor's productivity regardless of the collaboration patterns in the mover's location, as Proposition 2.2.3.

Table 2.14: The Effect of Relocation and the Access to a New Network

	New Network		Old Network	
	(1) Annual Number of Patents	(2) Annual Number of Adjusted Citations	(3) Annual Number of Patents	(4) Annual Number of Adjusted Citations
$PostMove^{\text{Real}}$	0.048^{*}	0.069^{***}	0.056	0.054
	(0.029)	(0.027)	(0.059)	(0.065)
Control Post Mean	0.472	0.325	0.604	0.397
Percentage Change	+10.25%	+21.32%	+9.25%	+13.54%
P-Value H_0: Diff. $= 0$			0.09	0.08
Observations	333731	333731	37373	37373
Individual FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Experience FE	Yes	Yes	Yes	Yes

Notes: The information presented in this table presents the results from regression equation (2.5). In column (1), I consider all the observations in the panel, whether or not the inventor patented in that year, while column (2) covers only the years when the inventor patents. The unit of analysis in these regressions is still inventor-year. The dependent variable is the share of collaborators in the destination location. Standard errors are clustered at the mover level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

2.6 Conclusion

In this paper I study the impact of an inventor's relocation on the productivity of their former collaborators. I leverage a novel dataset that combines patent and inventor information from the USPTO with inventors' online professional profiles, which enables me to identify the movers, track the exact timing of the move, and accurately assign the origin and destination location.

I construct a model that integrates team-based collaboration and information sharing within a network. This model allows me to explore each channel in isolation and offers insights into the underlying dynamics at play. Consequently, I am able to delve into the mechanisms potentially underpinning the outcomes and formulate hypotheses about these channels, which I subsequently investigate through empirical analysis.

I find a positive effect of the relocation of co-inventors and an increase in the productivity of the inventors they "leave behind." The productivity measures under consideration encompass the annual count of patents and the annual number of adjusted citations, which collectively reflect both the quantity and quality of the patents produced.

Heterogeneous treatment effects underscore the dominant role of information sharing across different locations. The findings not only point to elevated annual patent and citation numbers due to this information sharing, but also highlight that the inventor who remains in the original location expands their network of collaborators and information network into the mover's destination. Moreover, inventors remaining in the original location cite patents originating from the new destination location more frequently, which is consistent with information sharing.

The tradeoff between increasing productivity through enhancing agglomeration on the one hand and addressing spatial inequality on the other is at the core of debates around place-based policies. In this paper, I highlight that, under certain conditions, an inventor's relocation can lead to sizeable spillover effects in the origin locations on their former collaborators. For these beneficial effects to materialize, the degree of information exchange is essential, and the information shared depends on the inventor's history of collaboration with the mover, their informal links, and the extent to which an innovator's move leads to accessing new information networks as perceived by the collaborators staying in the origin location. In essence, relocations can lead to brain gain by promoting information sharing across potentially distanced geographical locations, which should be considered when debating the policy tradeoff between innovation and spatial inequality.

Chapter 3

Public R&D Funding and High-Skilled Workers Mobility[1]

3.1 Introduction

One the the most remarkable feature of the US economy is a steady increase of income per capita over the last 150 years of roughly 2% annually, see Jones (2016). A large literature following Solow (1956) has pointed to the role of technological change and productivity growth as the fundamental driver of long-run growth. Romer (1990) develops a theory that micro-founds long-run technological change, jumpstarting a vast literature on endogenous growth. Purposeful innovation and skilled labor – the key inputs into the innovative process – are at the center of this largely theoretical literature.

In this paper we develop a quantitative spatial growth model, and use a natural experiment in the form of an innovation subsidy, to make progress towards a realistic theory of innovation across space and time. The model highlights the role of the spatial mobility of skilled labor in shaping the economy's response to local innovation subsidies. Taking account of this labor reallocation is of first-order importance in understanding the local and aggregate effects of innovation policy. We use the model to estimate dynamic knowledge

[1]Co-authored with Florian Trouvain.

externality, where we make a crucial distinction between local and aggregate externalities. The former is operating on the level of a local labor market, while the latter is national (or even international). Dynamic knowledge externalities are the central parameter in this literature. They shape the extent to which market-based innovation is efficient, and raise the possibility of welfare enhancing policy interventions.

There is, unfortunately, no agreement as to how large these externalities are.[2] Part of the disagreement is due to the endogeneity of innovation: isolating a clean source of variation to causally estimate an elasticity is challenging. Moreover, extrapolating from local to aggregate spillovers is non-trivial.

We make progress by combining the natural experiment first exploited in Gross and Sampat (2023) with a spatial growth model. The experiment allows us to leverage a clean source of variation, and the theoretical model allows us to map the estimates into structural parameters. The model features local and aggregate externalities, and allows for imperfect mobility of inventors across space. This allows us to distinguish local knowledge externalities, aggregate knowledge externalities, and negative spillovers driven by a reallocation of skilled labor across space.

We proceed in two steps. First, we set up a benchmark spatial growth model that combines model elements from Romer (1990) and Jones (1995) with a simple theory of imperfectly mobile spatial labor supply using Eaton and Kortum (2002) extreme value formulation. The model makes precise the role of skilled labor in the innovative process, and highlights how spatial mobility shapes the response of the economy to an innovation subsidy, both on the local and on the aggregate level.

Second, we build on the works of Gross and Sampat (2023), which study the the long-lasting impact of the largest R&D shock in U.S. history on the country's innovation system.

[2]Romer (1990), Grossman and Helpman (1991), and Aghion and Howitt (1990) are so-called first generation growth models with strong knowledge spillovers, which sustain long-run growth even when resources devoted to innovation are fixed. Jones (1995) suggests a semi-endogenous version with a weaker dynamic spillover. Empirically, Bloom *et al.* (2020) offer evidence that ideas are getting harder to find, consistent with spillovers much weaker than what is assumed in the first generation growth models, and more recent Neo-Schumpeterian literature, see for instance **?**. Peters (2021) offers estimates that are higher than the ones implied by Bloom *et al.* (2020) but substantially below the parameterizations used in the first-generation endogenous growth models.

The authors find sizeable positive effects on the growth rate of patents not only during the World War II period, when R&D funding was allocated, also extending into the post-war era.

We enrich their impressive dataset by supplementing it with information on the skill composition in U.S. counties sourced from the Census. This additional data allows us to estimate the impact on the share of high-skilled workers in counties that received funding as part of this shock. Our rationale is that the surge in innovation activity observed in these counties may not solely stem from enhancing the productivity of existing workers but could also be attributed to an influx of high-skilled migrants, who contribute to the innovation efforts and patenting activity.

By employing a difference-in-difference research design, we find a 20 to 30 percent increase in the share of high-skilled workers in the affected counties. Since affected counties also experience a population growth, we conclude that our result does not reflect the displacement of low-skilled workers. This suggests that at least a portion of the effect can be attributed to the influx of new high-skilled individuals rather than merely an increase in the productivity of existing workers.

Literature Review. Theoretically, we build on the seminal work of Romer (1990), and other first generation endogenous growth models (see Aghion and Howitt (1990) and Grossman and Helpman (1991)). An important distinction drawn by Jones (1995) is the one between so-called strong and weak scale effects. Strong scale effects refer to dynamic knowledge externalities that are so strong that they can sustain perpetual productivity growth with a constant number of workers performing R&D, a proposition that Jones (1995) argues is unreasonable. Recent work, especially Bloom *et al.* (2020), confirm that ideas appear harder to find over time, using relatively aggregate data. We contribute to this debate by leveraging a natural experiment to provide new estimates on local knowledge externalities, and interpret them through the lens of a spatial growth model. We relate our findings to theirs, and a vast literature that has studied local knowledge externalities, see for instance Moretti (2019). A relevant paper in this context is Peters (2021), which develops a spatial growth models and disciplines it using arguable exogenous shifts in migrant inflows. Our framework is complementary to his, as we focus on a innovation subsidy while his context is more broadly focused on the effect of immigration. We also have a different model of local and aggregate knowledge externalities, which matters quantitatively in how to map cross-sectional estimates into policy relevant counterfactuals.

From an empirical standpoint, this paper builds upon the methodology utilized and datasets collected in Gross and Sampat (2023) and Gross and Sampat (2024). Expanding upon their findings that public funding enhances innovation activity and patenting, our focus lies on elucidating the mechanisms driving the notable upsurge in patenting and innovation behavior observed in counties that received government funding. While their analyses maintain a fixed labor supply and assumes no mobility, we investigate the impact of domestic migration of high-skilled workers on the reported increase in innovation activity.

Another strand of the literature examines the spillover effects stemming from publicly funded R&D. For instance, Myers and Lanahan (2022) highlight significant geographical spillovers catalyzed by small firms benefiting from public R&D funding, particularly within the same technological domain. Similarly, Azoulay *et al.* (2019), who concentrate on biomedical research and leverage public R&D windfalls, study the effects of public R&D funding on patenting and innovation, revealing substantial impacts of public funding on patenting activity. Additionally, Bloom *et al.* (2013) explore the trade-off between technological spillovers and product market rivalry, demonstrating the potency of technology spillovers in fostering positive effects on firms operating within similar domains. In order to distinguish between technology spillovers and product market rivalry, they leverage a tax scheme reform which can be thought of a R&D subsidy. They underscore the dominance of technology spillovers in driving innovation. Lastly, Van Reenen (2021) provide a comprehensive overview of human capital policies aimed at stimulating innovation particularly focusing on policies facilitating the entry of individuals from underrepresented groups into the innovation sector. Our contribution to this literature lies in examining how the migration of already high-skilled workers can bolster innovation in areas that have benefited from public funding.

Furthermore, a related paper by Moretti (2021) demonstrates that high-skilled migrants, such as inventors, exhibit heightened productivity upon relocating to innovation clusters. Our findings support these results, revealing that the increase in innovation activity within a cluster can indeed be attributed to the influx of high-skilled workers into that region.

3.2 Model of Innovation in Space

We follow Romer (1990) and Jones (1995) and set up an expanding variety idea-based growth model similar to Trouvain (2022). We abstract away from long-run growth for now, and focus on a steady state. Our equilibrium concept can readily be extended to one of a balanced growth path with population growth as driving force of long-run productivity growth.

Time is continuous. There are two locations indexed by $j \in \{1, 2\}$, a measure H of skilled labor endowed with one unit labor, and a measure L of unskilled labor endowed with one unit of labor. Regions produce ideas, and an undifferentiated final good is traded frictionlessly across space.

3.2.1 Innovation

Within each region a flow of new ideas $\dot{A}_j^{\mathrm{N}}$ is created using the following entry technology

$$\dot{A}_j^{\mathrm{N}} = \frac{A_{\mathrm{W}}^{\phi} A_j^{\zeta}}{f_{\mathrm{E},j}} H_j$$

where A_j^{ζ} is a local knowledge spillover, A_{W}^{ϕ} is a global knowledge spillover with $\sum_j A_j = A_{\mathrm{W}}$ representing all ideas in the economy, and $f_{\mathrm{E},j}$ is a fixed entry cost which consist of a contemporaneous congestion externality $H^{1-\lambda}$ and a fundamental location-specific productivity term γ_j, i.e., $f_{\mathrm{E},j} = \frac{H_j^{1-\lambda}}{\gamma_j}$. The sign and size of the parameters ζ and ϕ are unrestricted but we will assume that they are smaller than unity. The former governs the local knowledge externality, while the latter governs the aggregate knowledge externality. Together, this implies the following law of motion for net ideas in location j

$$\dot{A}_j = \gamma_j A_{\mathrm{W}}^{\phi} A_j^{\zeta} H_j^{\lambda} - \delta A_j.$$

Denote the value of an innovation as

$$V_{\mathrm{I}} = \int_t^{\infty} e^{-(r+\delta)s} \pi_s ds$$

where π are the flow profits that accrue to the innovator. The net present discounted value of an ideas follows from geometric discounting using the interest rate r and death shock δ. In a free entry equilibrium (with positive innovation), the value of an idea equals the entry cost

$$V_{\mathrm{I}} = \frac{f_{\mathrm{E},j} w_{H,j}}{A_{\mathrm{W}}^{\phi} A_j^{\zeta}}$$

where w_H is the high-skilled wage rate, and w_L is the low skilled wage rate. Note that V_{I} does not have a j subscript, which follows from the assumption that ideas are immediately adopted in every location so the benefit of innovation is unrelated to where the idea has been created. The cost, on the other hand, is location specific and has a j subscript.

3.2.2 Production

Final goods producers combine capital goods with unskilled labor to produce final output according to a constant returns to scale technology,

$$Y_j = \left(\left(\int x_{i,j}^{\frac{\sigma-1}{\sigma}} di \right)^{\frac{\sigma}{\sigma-1}} \right)^{\alpha} L_j^{1-\alpha}.$$

Final goods and capital goods are traded frictionlessly across regions so we can drop the j subscripts. By symmetry of capital goods the production functions collapses to

$$Y = X^{\alpha} \left(A_{\mathrm{W}} L \right)^{1-\alpha}$$

where X is the total quantity of capital goods, and A is the total number of capital goods, so that the amount of capital per specific variety is $x_i = \frac{X}{A} = \bar{x}$. First order conditions imply

$$\frac{\partial Y_j}{\partial X} = p_{X,j}$$
$$\frac{\partial Y_j}{\partial L} = w_{L,j}$$

and the final goods producer makes zero profits.

Capital goods can be produced from final goods one-for-one, i.e.,

$$x_i = \frac{y_i}{c}$$

but a markup is charged consistent with the model of monopolistic competition across differentiated capital goods so that

$$p_x = \frac{\sigma}{\sigma - 1} p_Y c.$$

We normalize the final goods price to unity so $p_x = \frac{\sigma}{\sigma-1} c$.

3.2.3 Labor Supply and household choices

Workers supply their labor inelastically, but they are mobile across space. For now, suppose this gives rise to reduced-form spatial labor supply equations of the form

$$H_j^S = B_H w_{H,j}^{\theta_H}$$
$$L_j^S = B_L w_{L,j}^{\theta_L}$$

where parameters θ_H and θ_L govern the elasticity of labor supply. In addition, suppose that an inter-temporal consumption smoothing problem implies an equilibrium real interest rate of $r = \rho$ where ρ is the household discount factor. Since final goods are freely traded, wages of unskilled workers are equalized across space and their location choices don't matter for anything that follows. We will relax this assumption later.

3.2.4 Market clearing and Free Entry

Goods market clearing requires

$$C + \int y_i di = Y.$$

Total spending on capital goods follows from the first order condition, which can be inverted to provide the equilibrium quantity of capital goods

$$X = A_{\mathrm{W}} L \left(\frac{\alpha}{c} \frac{\sigma - 1}{\sigma} \right)^{\frac{1}{1-\alpha}}.$$

This in turn can be used to derive aggregate output

$$\begin{aligned} Y &= X^{\alpha} \left(A_{\mathrm{W}} L \right)^{1-\alpha} \\ &= \left(\frac{\alpha}{c} \frac{\sigma - 1}{\sigma} \right)^{\frac{\alpha}{1-\alpha}} A_{\mathrm{W}} L. \end{aligned}$$

The innovators' rents from selling differentiated capital goods thus read

$$\begin{aligned} \pi &= \frac{\alpha Y - cX}{A_{\mathrm{W}}} \\ &= \frac{p_x X}{A_{\mathrm{W}}} \frac{1}{\sigma} \\ &= \frac{\alpha}{\sigma} \left(\frac{\alpha}{c} \frac{\sigma - 1}{\sigma} \right)^{\frac{\alpha}{1-\alpha}} L \end{aligned}$$

where we conveniently summarize several constants using the parameter $\Lambda = \frac{\alpha}{\sigma} \left(\frac{\alpha}{c} \frac{\sigma-1}{\sigma} \right)^{\frac{\alpha}{1-\alpha}}$ so the flow profits read ΛL.

Since L is fixed, the effective discount rate equals $\rho + \delta$, and the value of an innovation (but not the cost) is independent of where it was invented we have

$$\begin{aligned} V_{\mathrm{I}} &= \int_t^{\infty} e^{-(r+\delta)s} \pi_s ds \\ V_{\mathrm{I}} &= \frac{\Lambda L}{\rho + \delta}. \end{aligned}$$

Free entry then implies an arbitrage condition across locations of the following form

$$\frac{f_{E,1} w_{H,1}}{A_1^{\zeta}} = \frac{f_{E,2} w_{H,2}}{A_2^{\zeta}}$$

where the aggregate externality conveniently cancels. After accounting for the congestion externality and fundamental research productivity differences across locations hidden in the fixed cost, we get

$$\frac{H_1^{1-\lambda} w_{H,1}}{\gamma_1 A_1^{\zeta}} = \frac{H_2^{1-\lambda} w_{H,2}}{\gamma_2 A_2^{\zeta}}. \tag{3.1}$$

3.2.5 Steady State Solution

In the steady state, we have $\dot{A}_j = \dot{H}_j = 0$. Using the law of motion of idea production it follows that ideas produced in a location are proportional to the number of skilled workers in that location

$$A_j = \left(\frac{A_W^{\phi} \gamma_j}{\delta} H_j^{\lambda} \right)^{\frac{1}{1-\zeta}}. \tag{3.2}$$

Combining (3.1) and (3.2), i.e., the free entry condition and the equilibrium measure of ideas as a function of equilibrium skill endowment, allows us to derive an equilibrium labor demand equation

$$\left(\frac{H_1}{H_2} \right)^{1-\lambda} \frac{w_{H,1}}{w_{H,2}} = \frac{\gamma_1 \left(\frac{\gamma_1}{\delta} H_1^{\lambda} \right)^{\frac{\zeta}{1-\zeta}}}{\gamma_2 \left(\frac{\gamma_2}{\delta} H_2^{\lambda} \right)^{\frac{\zeta}{1-\zeta}}}$$

which simplifies to

$$\left(\frac{w_{H,1}}{w_{H,2}} \right) = \left(\frac{\gamma_1}{\gamma_2} \right)^{\frac{1}{1-\zeta}} \left(\frac{H_1}{H_2} \right)^{\frac{\lambda}{1-\zeta} - 1} \tag{3.3}$$

Two aspects of (3.3) are noteworthy: First, for any $\lambda + \zeta < 1$, labor demand is downward sloping in the relative wages. This is in line with the standard intuition. However, when local scale effects are strong and congestion effects weak, with $\lambda + \zeta > 1$ being the key condition, this relationship is reversed and higher wages coincide with higher labor demand, which is a distinct possibility in models with increasing returns. Second, fundamental research productivity is positively related and explodes when $\zeta \to 1$.

Next, we use the reduced-form constant elasticity labor supply function to solve for the endogenous relative high-skilled wage across regions

$$\left(\frac{H_1}{H_2}\right)^{\frac{1}{\theta_H}} = \left(\frac{\gamma_1}{\gamma_2}\right)^{\frac{1}{1-\zeta}} \left(\frac{H_1}{H_2}\right)^{\frac{\lambda}{1-\zeta}-1} \tag{3.4}$$

where the right hand side of equation (3.4) maps fundamental differences in relevant spatial productivity differences taking account of local externalities and congestion forces, while the left hand side combines labor supply elasticities, local externality, and congestion forces. Rearranging makes clear that this model features a solution with positive innovation in each location only if

$$\frac{\lambda}{1-\zeta} < 1 + \frac{1}{\theta}.$$

In that case, the coefficient $\frac{\lambda}{1-\zeta} - \left(1 + \frac{1}{\theta}\right)$ is positive, and equation (3.4) can be inverted to reveal the relative skilled labor supply across space

$$\left(\frac{H_1}{H_2}\right) = \left(\frac{\gamma_1}{\gamma_2}\right)^{\frac{1}{1-\zeta-\lambda+\frac{1-\zeta}{\theta_H}}}.$$

It will be useful to define the relative share of skilled labor in location j as χ_j

$$\chi_1 = \frac{1}{1+\left(\frac{\gamma_1}{\gamma_2}\right)^{\kappa}} H$$

where $\kappa := \frac{1}{1-\zeta-\lambda+\frac{1-\zeta}{\theta_H}}$.

The labor supply elasticity is a key parameter. If it is very low, an equilibrium will always be guaranteed. Even if there are strong increasing returns due to ζ being large, workers are unwilling to move. Of course, in such a scenario the wag gaps across regions would be rather large. To see this, assume that an interior solution obtains and derive the wage ratio

$$\left(\frac{w_{H,1}}{w_{H,2}}\right) = \left(\frac{\gamma_1}{\gamma_2}\right)^{\frac{1}{\theta_H}\frac{\frac{1}{1-\zeta}}{1+\frac{1}{\theta_H}-\frac{\lambda}{1-\zeta}}}$$

where a small θ_H leads to massive wage differences even if there are only small differences in fundamentals.

If there are no congestion forces $\lambda = 1$, and there are no local scale effects $\phi = 0$ and $\zeta = 0$,then the allocation of skilled labor across space collapses to

$$\left(\frac{H_1}{H_2}\right) = \left(\frac{\gamma_1}{\gamma_2}\right)^{\theta_H}$$

with relative wages read

$$\frac{w_{H,1}}{w_{H,2}} = \left(\frac{\gamma_1}{\gamma_2}\right).$$

This special case constitutes an efficient benchmark.

Aggregate productivity can be derived as follows. Note that the ratio of ideas across locations can be expressed as

$$\frac{A_1}{A_2} = \left(\frac{\gamma_1}{\gamma_2}\left(\frac{\chi_1}{\chi_2}\right)^{\lambda}\right)^{\frac{1}{1-\zeta}}.$$

Next, note

$$A_1^{1-\zeta-\phi} = \left(\left(1+\frac{A_2}{A_1}\right)^{\phi}\gamma_1\left(\chi_1 H\right)^{\lambda}\right)$$

which follows from the steady state link between ideas and skilled labor in each location. Combining these two expressions delivers the total level of ideas in location 1, which then directly be used to compute the aggregate stock of ideas

$$A_1 = \left(\left(\gamma_1\left(\chi_1\right)^{\lambda}+\gamma_2\left(\chi_2\right)^{\lambda}\right)^{\phi}\gamma_1\left(\chi_1\right)^{\lambda(1-\phi)}\right)^{\frac{1}{1-\zeta-\phi}} H^{\frac{\lambda}{1-\zeta-\phi}}$$

so

$$A_{\mathrm{W}} = \left\{\left(\gamma_2\left(\chi_2\right)^{\lambda}\right)^{\frac{1-\phi}{1-\zeta-\phi}}+\left(\gamma_1\left(\chi_1\right)^{\lambda}\right)^{\frac{1-\phi}{1-\zeta-\phi}}\right\}\left(\gamma_1\left(\chi_1\right)^{\lambda}+\gamma_2\left(\chi_2\right)^{\lambda}\right)^{\frac{\phi}{1-\zeta-\phi}} H^{\frac{\lambda}{1-\zeta-\phi}}.$$

If one were to introduce population growth, the long-run growth rate would equal $g_{A_{\mathrm{W}}} = \frac{\lambda}{1-\zeta-\phi}g_H$ similar to Jones (1995).

3.2.6 Effect of a Subsidy

We differentiate between persistent and temporary subsidies. First, suppose that a policy maker subsidies innovation in region 1 by reducing the fixed cost of entry by some factor

$\tau_s < 1$, which means that the entry cost fall. This is identical to an increase in local research productivity γ from the innovator's point of view. We summarize the effects on skilled labor allocation and relative skilled wages using the elasticity of each variable with respect to the subsidy (or fundamental productivity)

$$\frac{\partial \log \frac{w_{H,1}}{w_{H,2}}}{\partial \log \tau_{s,1}} = \frac{1}{\theta_H} \frac{\frac{1}{1-\zeta}}{1 + \frac{1}{\theta_H} - \frac{\lambda}{1-\zeta}}$$
$$\frac{\partial \log \frac{H_1}{H_2}}{\partial \log \tau_{s,1}} = \frac{1}{1 - \zeta - \lambda + \frac{1-\zeta}{\theta_H}}.$$

We summarize to key takeaways. The stronger the local knowledge spillover ζ, the larger are the effects of a subsidy. A higher spatial labor supply elasticity amplifies the growth in the amount of skilled labor, and mutes the response of wages.

A temporary subsidy has no long-run effects in our model, unless in a special knife-edge case when

$$1 \approx \frac{1}{1 - \zeta - \lambda + \frac{1-\zeta}{\theta_H}}.$$

You can verify that this is not the same condition that we spelled out before in order for a spatial equilibrium with multiple locations to exist. To see that these two conditions are not the same, consider the case with perfectly immobile labor but strong scale effects, say $\zeta = 1$ and $\lambda = 0$. In that case, a temporary subsidy would lead to more ideas produced during the time interval of the subsidy. As the subsidy expires, fundamental research productivity reverts, but the strong knowledge spillover ensures that the temporary effect persists through time.

3.3 Data

The data used in this paper comes from two sources. The first is the OSRD funding collected and digitized by Gross and Sampat (2024) and the second is US Census data at the county level. Together, these datasets allow me to follow workers migration between different counties in the US, following the monetary shock associated with research funding during

World War II.

3.3.1 OSRD Funding

Innovation during World War II era is mostly associated with the development of the nuclear bomb and the Manhattan Project. However, beyond these notable endeavors, the period spanning from 1941 to 1948 witnessed a significant surge in government-funded research and development (R&D) across various domains. Figure 3.1 shows that the government funding as a share of GDP increased during that period.[3]

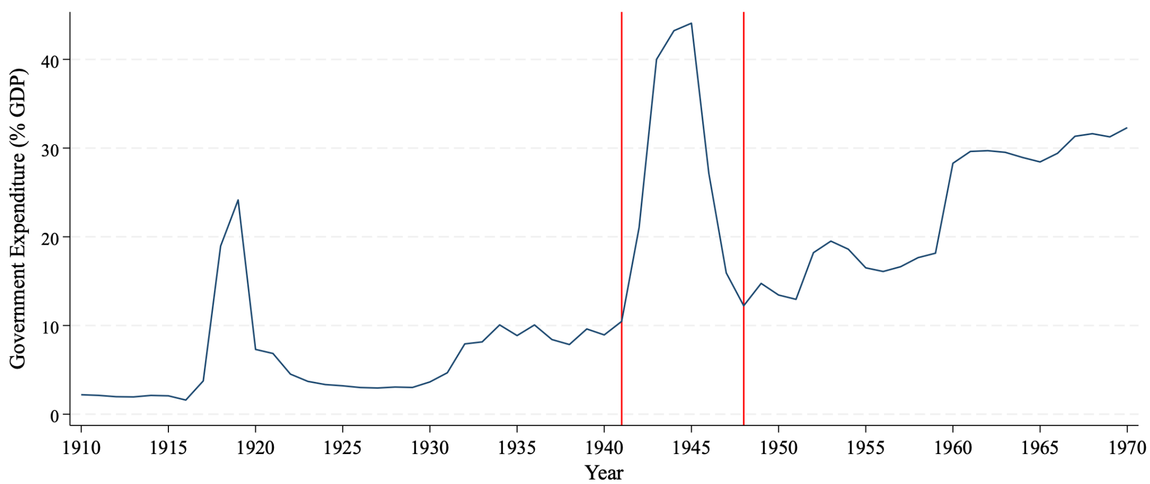


Figure 3.1: Government Funding as Percent of GDP

Notes: This figure shows the US government funding during the OSRD organization increased as a share of GDP. The data is taken from the IMF.

This increase in funding was driven by the recognition that both the US military and its allies would require substantial innovation to secure victory in the war and to facilitate post-war recovery efforts. Consequently, R&D initiatives extended beyond weaponry and protective gear to encompass support for civilians and returning soldiers.

To ensure adequate allocation of resources to R&D, the National Research Defense

[3]Even though this figure includes all government funding, and not just R&D funding, the high surge suggests that this event was unusual, especially when compared to World War I.

Committee (NRDC) was established. The NRDC not only oversaw the distribution of funding to research institutions and universities but also fostered collaboration between these entities and the scientists engaged in their work. In response to the wartime demand, the NRDC expanded in 1941, giving rise to the Office of Scientific Research and Development (OSRD).

The OSRD operated by disbursing funding through various contractors, ranging from individuals to universities and research institutes. Gross and Sampat (2024) collected a comprehensive collection of OSRD contracts, a total of 2,254 contracts. These records include details such as the recipients' identities, locations, funding amounts, and the specific divisions from which the funds were allocated. Additionally, the documentation encompasses the inventions and resulting patents arising from these contracts.

3.3.2 US Census data

The US decennial census data spanning from 1920 to 1990 includes information on the location, occupation, education, wage and other demographics such as age and gender on a selected share of the US population. By utilizing this dataset, we can track the evolving proportion of high-skilled workers over time and across counties.

Specifically, high-skilled workers are defined as those with a minimum of four years of college education, and their share is computed out of the working age population (i.e., between 25 and 60) at the county in each one of these years. Due to data constraints, we in order to extend our analysis beyond 1940 we must interpolate education to previous years, as educational attainment was not surveyed prior to that year.[4] We rely on Forstall (1996), which provides comprehensive population data at the county-level, when accounting for the total population in a specific county.

A challenge we encounter is the dynamic nature of county boundaries over time. Since our analysis operates at the aggregation level of county-year, we must address this issue. To

[4] We employ various skill measures, including occupation-based criteria, and consistently observe similar outcomes.

do so, we adopt the methodology outlined in Gross and Sampat (2023), which uses county boundaries in their 1920 delineations. This approach ensures continuity in our geographical units, facilitating the comparison of skill compositions and population, more generally, across counties over time.

We link this information to the information about the counties that received the OSRD funding shock to study how the shock affects the composition of high-skilled workers and migration into that location more generally.

3.4 Empirical Analysis

This section presents the methodology employed to estimate the average treatment effect of the World War II R&D government funding shock on high-skilled migration. We utilize a difference-in-differences research design, comparing counties that received government funding (treatment group) with those that did not receive any such funding (control group).

3.4.1 Identification

One concern we encounter is the potential endogeneity of funding allocation, which raises the possibility that funds were directed towards counties already poised for higher growth in high-skilled populations. An example can be big institutions with room for expansion.

Our identification assumption rests on the premise that OSRD funding allocation was not influenced by unobserved factors related to high-skilled migration. In other words, the counties receiving funding were not inherently more conducive to attracting high-skilled workers compared to neighboring counties during World War II.

As explained by Gross and Sampat (2023), the decision-making process for OSRD funding allocation was primarily driven by short-term military needs, without consideration for long-term outcomes. The primary objective was victory in the war, without forethought regarding post-war consequences. Consequently, it is likely that allocation decisions were not correlated with the specific outcomes under examination in this project.

3.4.2 Baseline Effects

Building on the methodology outlined in Gross and Sampat (2023), we run an OLS regressions incorporating leads and lags around the period of OSRD funding from 1941 to 1948. This approach allows us to investigate the impact of this funding shock on the high-skilled workforce share and population growth within these counties. Additionally, to ensure the validity of our research design, we conduct tests to assess the presence of parallel trends before the occurrence of the shock.

Specifically, we compare the share of high-skilled workers and population growth over time in counties that receive some government funding relative to counties that did not enjoy any such funding. The OLS regression we estimate is

$$\text{FracHS}_{ct} = \sum_{t=1940}^{1970} \beta_t \cdot \text{Treated}_c \cdot \text{Year}_t + \alpha_c + \delta_t + \varepsilon_{ct} \tag{3.5}$$

where c is county and t is year. The years in the sample are 1930 to 1970 in jumps of 10 years. The variable treatment is equal to 1 when the county received some OSRD funding and 0 otherwise. Year_t are year dummies.

Figure 3.2 shows the results from estimating equation (3.5). The figure ensures that there are no pre-trends, supporting our research design. Moreover, we can see that the fraction of high-skilled workers increase after the OSRD shock, and is on an upward trend until 1970. The overall shock represents between 20 to 30 percent increase in the share of high-skilled workers.

To gain deeper insights into the underlying dynamics of this surge and discern whether it stemmed from the displacement of low-skilled workers or an actual augmentation in the county's population, we re-estimate equation (3.5) using population growth as the outcome variable.

Figure 3.3 depicts the impact of the shock on population growth. We note some pre-trends suggesting that funded counties were predisposed to experiencing population growth, a trend which persists and intensifies following the OSRD funding shock. This indicates that the rise in the high-skilled workforce share subsequent to the shock was not solely

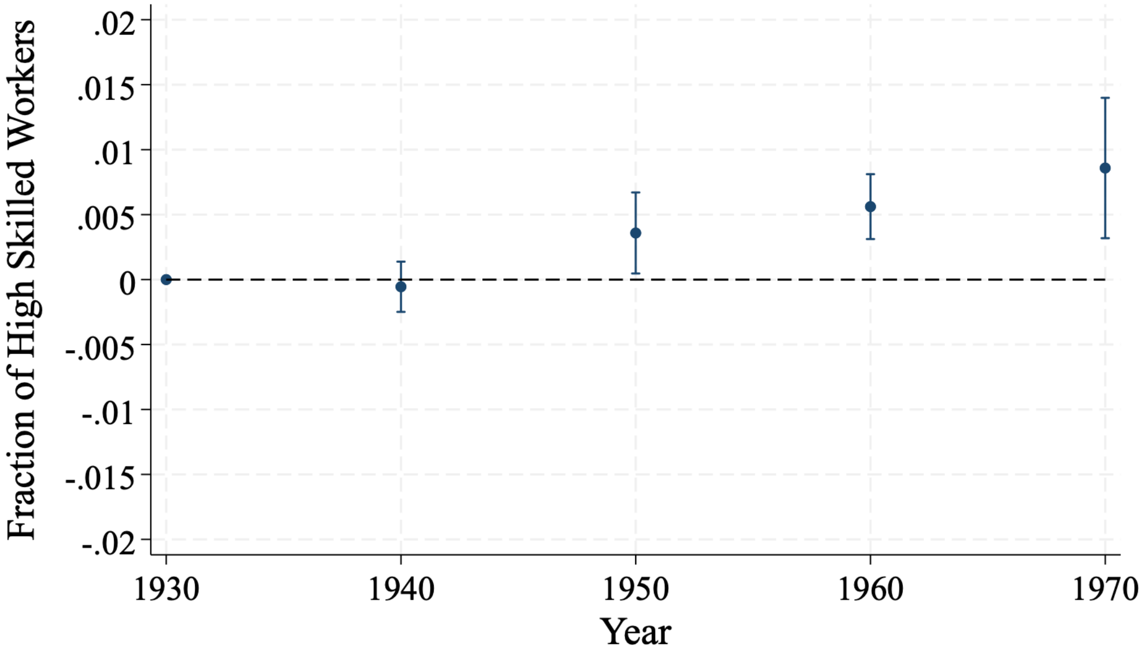


Figure 3.2: Effect on Fraction of High-Skilled Workers

Notes: This figure shows the decennial estimates of the effect os the OSRD shock on the fraction of high-skilled workers in a county. The vertical lines represent a 95 percent confidence interval.

attributed to the departure of low-skilled workers, but also to an overall population increase within the county.

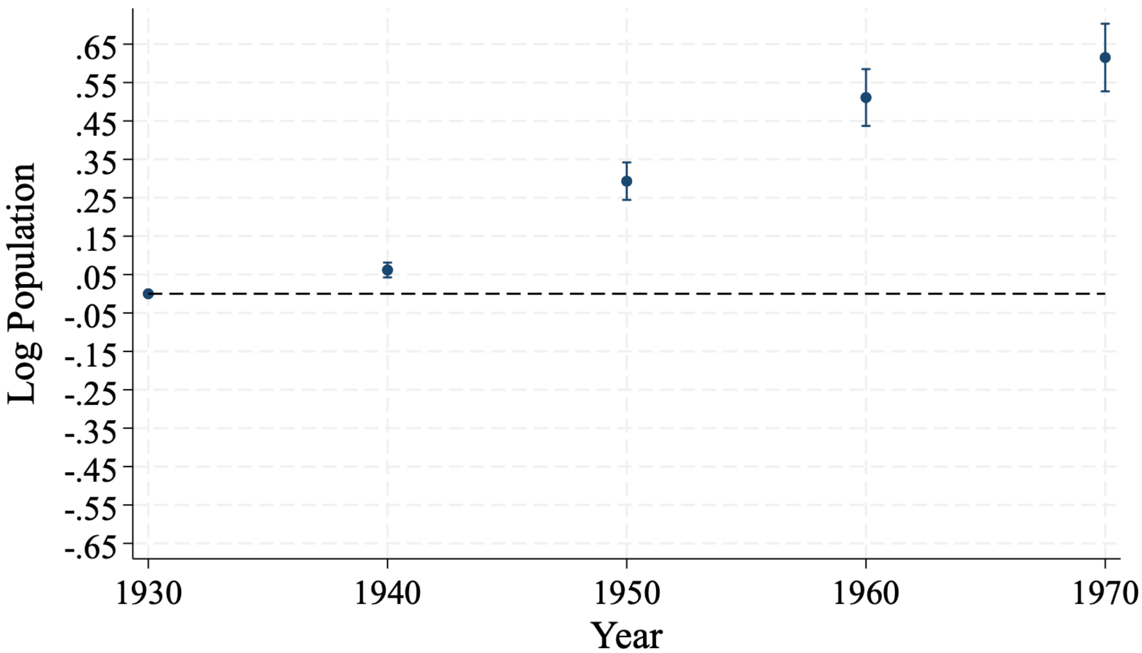


Figure 3.3: Effect on Population Growth

Notes: This figure shows the decennial estimates of the effect os the OSRD shock on the fraction of high-skilled workers in a county. The vertical lines represent a 95 percent confidence interval.

3.5 Conclusion

We develop a spatial growth model with a focus on mobility of innovators across space. The model highlights how inventor mobility, local knowledge externalities, and aggregate knowledge externalities jointly pin down the aggregate and distributional effects of innovation policy with a focus on regional and skill type heterogeneity. We use a quasi-natural experiment based on Gross and Sampat (2023) to discipline key parameters of the theory.

To do so we enrich their original dataset with information on inventor and skilled labor mobility, which we argue is a key ingredient to make progress toward a better understanding of how local innovation subsidies impact growth across space and time. We find that an innovation subsidy has long-lasting effects on the skill composition and size of a treated local labor market, which needs to be taken into account when inverting the estimates to uncover structural scale elasticities. Moreover, this reallocation of skilled labor has negative impacts on non-treated regions, which matters both for spatial inequality and aggregate research output. This short chapter is a first step in an ambitious project that we hope will contribute to a better understanding of the innovative process.

References

AGHION, P. and HOWITT, P. (1990). *A model of growth through creative destruction*. Tech. rep., National Bureau of Economic Research.

AGRAWAL, A., KAPUR, D. and MCHALE, J. (2008). How do Spatial and Social Proximity Influence Knowledge Flows? Evidence from Patent Data. *Journal of Urban Economics*, **64** (2), 258–269.

AZOULAY, P., GRAFF ZIVIN, J. S., LI, D. and SAMPAT, B. N. (2019). Public R&D Investments and Private-Sector Patenting: Evidence from NIH Funding Rules. *The Review of Economic Studies*, **86** (1), 117–152.

—, — and WANG, J. (2010). Superstar Extinction. *The Quarterly Journal of Economics*, **125** (2), 549–589.

BATTISTON, D., BLANES I VIDAL, J. and KIRCHMAIER, T. (2021a). Face-to-Face Communication in Organizations. *The Review of Economic Studies*, **88** (2), 574–609.

—, — and — (2021b). Face-to-Face Communication in Organizations. *The Review of Economic Studies*, **88** (2), 574–609.

BERNSTEIN, S., DIAMOND, R., JIRANAPHAWIBOON, A., MCQUADE, T. and POUSADA, B. (2022). The Contribution of High-Skilled Immigrants to Innovation in the United States. (30797).

BLOOM, N., HASSAN, T. A., KALYANI, A., LERNER, J. and TAHOUN, A. (2021). *The Diffusion of Disruptive Technologies*. Tech. rep., National Bureau of Economic Research.

—, JONES, C. I., VAN REENEN, J. and WEBB, M. (2020). Are ideas getting harder to find? *American Economic Review*, **110** (4), 1104–44.

—, LIANG, J., ROBERTS, J. and YING, Z. J. (2015). Does working from home work? evidence from a chinese experiment. *The Quarterly Journal of Economics*, **130** (1), 165–218.

—, SCHANKERMAN, M. and VAN REENEN, J. (2013). Identifying Technology Spillovers and Product Market Rivalry. *Econometrica*, **81** (4), 1347–1393.

BOLTE, L., IMMORLICA, N. and JACKSON, M. O. (2020). The Role of Referrals in Inequality, Immobility, and Inefficiency in Labor Markets. *Immobility, and Inefficiency in Labor Markets (January 1, 2020)*.

Bramoullé, Y., Currarini, S., Jackson, M. O., Pin, P. and Rogers, B. W. (2012). Homophily and Long-Run Integration in Social Networks. *Journal of Economic Theory*, **147** (5), 1754–1786.

Cabrales, A., Calvó-Armengol, A. and Zenou, Y. (2011). Social Interactions and Spillovers. *Games and Economic Behavior*, **72** (2), 339–360.

Calvó-Armengol, A. (2004). Job Contact Networks. *Journal of Economic Theory*, **115** (1), 191–206.

Calvo-Armengol, A. and Jackson, M. O. (2004). The Effects of Social Networks on Employment and Inequality. *American Economic Review*, **94** (3), 426–454.

Calvó-Armengol, A. and Jackson, M. O. (2007). Networks in Labor Markets: Wage and Employment Dynamics and Inequality. *Journal of Economic Theory*, **132** (1), 27–46.

Chillemi, O. and Gui, B. (1997). Team Human Capital and Worker Mobility. *Journal of Labor Economics*, **15** (4), 567–585.

Crescenzi, R., Nathan, M. and Rodríguez-Pose, A. (2016). Do Inventors Talk to Strangers? On Proximity and Collaborative Knowledge Creation. *Research Policy*, **45** (1), 177–194.

Cullen, Z. and Perez-Truglia, R. (2023). The Old Boys' Club: Schmoozing and the Gender Gap. *American Economic Review*, **113** (7), 1703–1740.

Cullen, Z. B. and Perez-Truglia, R. (2019). *The Old Boys' Club: Schmoozing and the Gender Gap*. Tech. rep., National Bureau of Economic Research.

Currarini, S., Jackson, M. O. and Pin, P. (2009). An Economic Model of Friendship: Homophily, Minorities, and Segregation. *Econometrica*, **77** (4), 1003–1045.

—, — and — (2010). Identifying the Roles of Race-Based Choice and Chance in High School Friendship Network Formation. *Proceedings of the National Academy of Sciences*, **107** (11), 4857–4861.

Deming, D. J. (2017). The Growing Importance of Social Skills in the Labor Market. *The Quarterly Journal of Economics*, **132** (4), 1593–1640.

Eaton, J. and Kortum, S. (2002). Technology, geography, and trade. *Econometrica*, **70** (5), 1741–1779.

Ellison, G. and Glaeser, E. L. (1997). Geographic Concentration in US Manufacturing Industries: a Dartboard Approach. *Journal of Political Economy*, **105** (5), 889–927.

— and — (1999). The Geographic Concentration of Industry: Does Natural Advantage Explain Agglomeration? *American Economic Review*, **89** (2), 311–316.

—, — and Kerr, W. R. (2010). What Causes Industry Agglomeration? Evidence from Coagglomeration Patterns. *American Economic Review*, **100** (3), 1195–1213.

Emanuel, N. and Harrington, E. (2020). *"Working" Remotely?: Selection, Treatment and the Market Provision Remote Work"*. Tech. rep., Harvard mimeo.

—, — and Pallais, A. (2023). The Power of Proximity to Coworkers: Training for Tomorrow or Productivity Today?

Forstall, R. L. (1996). *Population of States and Counties of the United States: 1790-1990*. US Department of Commerce.

Galeotti, A. and Merlino, L. P. (2014). Endogenous Job Contact Networks. *International Economic Review*, **55** (4), 1201–1226.

Greenstone, M., Hornbeck, R. and Moretti, E. (2010). Identifying Agglomeration Spillovers: Evidence from Winners and Losers of Large Plant Openings. *Journal of Political Economy*, **118** (3), 536–598.

Griliches, Z. (1998). *R&D and Productivity: The Econometric Evidence*. University of Chicago Press.

Griss, M. (1993). Hewlett-Packard Software Reuse: from Library to Factory.

Gross, D. P. and Sampat, B. N. (2023). America, Jump-Started: World War II R&D and the Takeoff of the US Innovation System. *American Economic Review*, **113** (12), 3323–3356.

— and — (2024). *The Government Patent Register: A New Resource for Measuring US Government-Funded Patenting*. Tech. rep., National Bureau of Economic Research.

Grossman, G. M. and Helpman, E. (1991). Quality ladders in the theory of growth. *The review of economic studies*, **58** (1), 43–61.

Hall, B. H., Jaffe, A. and Trajtenberg, M. (2005). Market Value and Patent Citations. *RAND Journal of Economics*, pp. 16–38.

—, Jaffe, A. B. and Trajtenberg, M. (2001). The nber patent citation data file: Lessons, insights and methodological tools. (8498).

Jacobs, A. Z. and Watts, D. J. (2021). A large-scale comparative study of informal social networks in firms. *Management Science*.

Jaffe, A. B., Jones, B. F. *et al.* (2015). *The Changing Frontier: Rethinking Science and Innovation Policy*. University of Chicago Press.

Jaravel, X., Petkova, N. and Bell, A. (2018). Team-Specific Capital and Innovation. *American Economic Review*, **108** (4-5), 1034–1073.

Jarosch, G., Oberfield, E. and Rossi-Hansberg, E. (2021a). Learning From Coworkers. *Econometrica*, **89** (2), 647–676.

—, — and — (2021b). Learning from Coworkers. *Econometrica*, **89** (2), 647–676.

Johnson, K. P. (2004). Redefinition of the bea economic areas. Acessed on January 20, 2023.

Jones, B. F. (2009). The Burden of Knowledge and the "Death of the Renaissance Man": Is Innovation Getting Harder? *The Review of Economic Studies*, **76** (1), 283–317.

Jones, C. I. (1995). R & d-based models of economic growth. *Journal of political Economy*, **103** (4), 759–784.

— (2016). The facts of economic growth. In *Handbook of macroeconomics*, vol. 2, Elsevier, pp. 3–69.

Kaltenberg, M., Jaffe, A. B. and Lachman, M. E. (2023). Invention and the Life Course: Age Differences in Patenting. *Research Policy*, **52** (1), 104629.

Kerr, S. P., Kerr, W., Özden, Ç. and Parsons, C. (2016). Global Talent Flows. *Journal of Economic Perspectives*, **30** (4), 83–106.

—, —, — and Parsons, C. (2017). High-skilled Migration and Agglomeration. *Annual Review of Economics*, **9**, 201–234.

— and Kerr, W. R. (2018). Global Collaborative Patents. *The Economic Journal*, **128** (612), F235–F272.

Kerr, W. R. (2008). Ethnic Scientific Communities and International Technology Diffusion. *The Review of Economics and Statistics*, **90** (3), 518–537.

Lerner, J. and Seru, A. (2021). The Use and Misuse of Patent Data: Issues for Finance and Beyond. *The Review of Financial Studies*, **35** (6), 2667–2704.

—, Sorensen, M. and Strömberg, P. (2011). Private Equity and Long-Run Investment: The Case of Innovation. *The Journal of Finance*, **66** (2), 445–477.

Lindenlaub, I. and Prummer, A. (2021). Network Structure and Performance. *The Economic Journal*, **131** (634), 851–898.

Lucas Jr, R. E. (2009). Ideas and Growth. *Economica*, **76** (301), 1–19.

— and Moll, B. (2014). Knowledge Growth and the Allocation of Time. *Journal of Political Economy*, **122** (1), 1–51.

Mailath, G. J. and Postlewaite, A. (1990). Workers versus Firms: Bargaining over a Firm's Value. *The Review of Economic Studies*, **57** (3), 369–380.

McPherson, M., Smith-Lovin, L. and Cook, J. M. (2001). Birds of a Feather: Homophily in Social Networks. *Annual review of sociology*, **27** (1), 415–444.

Merlino, L. P. (2014). Formal and Informal Job Search. *Economics Letters*, **125** (3), 350–352.

— (2019). Informal Job Search through Social Networks and Vacancy Creation. *Economics Letters*, **178**, 82–85.

Moretti, E. (2019). *The effect of high-tech clusters on the productivity of top inventors*. Tech. rep., National Bureau of Economic Research.

— (2021). The Effect of High-Tech Clusters on the Productivity of Top Inventors. *American Economic Review*, **111** (10), 3328–3375.

Moser, P., Voena, A. and Waldinger, F. (2014). German Jewish Émigrés and US Invention. *American Economic Review*, **104** (10), 3222–3255.

Myers, K. R. and Lanahan, L. (2022). Estimating Spillovers from Publicly Funded R&D: Evidence from the US Department of Energy. *American Economic Review*, **112** (7), 2393–2423.

Pallais, A. and Sands, E. G. (2016). Why the referential treatment? evidence from field experiments on referrals. *Journal of Political Economy*, **124** (6), 1793–1828.

Peters, M. (2021). *Market size and spatial growth-evidence from germany's post-war population expulsions*. Tech. rep., National Bureau of Economic Research.

Prato, M. (2022). The Global Race for Talent: Brain Drain, Knowledge Transfer, and Growth.

Romer, P. M. (1990). Endogenous technological change. *Journal of political Economy*, **98** (5, Part 2), S71–S102.

Roth, J. (2022). Pretest with Caution: Event-Study Estimates after Testing for Parallel Trends. *American Economic Review: Insights*, **4** (3), 305–322.

Sandvik, J. J., Saouma, R. E., Seegert, N. T. and Stanton, C. T. (2020). Workplace Knowledge Flows. *The Quarterly Journal of Economics*, **135** (3), 1635–1680.

Saxenian, A. (1994). *Regional Advantage: Culture and Competition in Silicon Valley and Route 128*. Harvard University Press.

Solow, R. M. (1956). A contribution to the theory of economic growth. *The quarterly journal of economics*, **70** (1), 65–94.

Stein, J. C. (2008). Conversations among Competitors. *American Economic Review*, **98** (5), 2150–2162.

Trouvain, F. (2022). Technology adoption, innovation, and inequality in a global world.

Van Reenen, J. (2021). *Innovation and Human Capital Policy*. Tech. rep., National Bureau of Economic Research.

Waldinger, F. (2012). Peer Effects in Science: Evidence from the Dismissal of Scientists in Nazi Germany. *The Review of Economic Studies*, **79** (2), 838–861.

Wuchty, S., Jones, B. F. and Uzzi, B. (2007). The Increasing Dominance of Teams in Production of Knowledge. *Science*, **316** (5827), 1036–1039.

Zacchia, P. (2018). Benefiting Colleagues but not the City: Localized Effects from the Relocation of Superstar Inventors. *Research Policy*, **47** (5), 992–1005.

— (2020). Knowledge Spillovers through Networks of Scientists. *The Review of Economic Studies*, **87** (4), 1989–2018.

Appendix A

Appendix to Chapter 1

A.1 Calculations

Deriving $p_i(s)$ Recall, $p_i(s)$ is the probability that an overloaded agent i gets help via the network. Denote by $|I| = (1-b)\,n$ the size of the set of idle workers, and by $|O| = bn$ the size of the set of overloaded workers.

Let i be an overloaded worker. The number of idle friends agent i has follows a binomial distribution $B(|I|, p)$, where $p = \frac{s}{n}$ is the probability a link between two agents is formed. Thus, the probability that i has m idle friends is given by

$$\mathbb{P}(N_i(I) = m) = \binom{|I|}{m} p^m (1-p)^{|I|-m}$$

Let $j \in N_i(I)$ be an idle friend of j. The number of overloaded friends j has follows a binomial distribution as well $B(|O|, p)$. Therefore, the probability that j has k overloaded friends, including i, is given by

$$\mathbb{P}\left(N_j(O) = k \middle| g_{ij} = 1\right) = \binom{|O|-1}{k-1} p^{k-1} (1-p)^{|O|-k}$$

Since j's help is distributed uniformly at random, the probability that i gets j's help is $\frac{1}{k}$. And we get that the expected probability that j helps i (under the assumption that a link between them exists) is

$$\mathbb{P}\left(j \text{ helps i} \middle| g_{ij}=1\right)=\sum_{k=1}^{|O|}\left[\frac{1}{k}\cdot\binom{|O|-1}{k-1}p^{k-1}\left(1-p\right)^{|O|-k}\right]$$

Now, since the probability that each of i's idle friends helps her is independent. Hence, if i has m links with idle workers, the probability that i does not get help is

$$\left\{1-\sum_{k=1}^{|O|}\left[\frac{1}{k}\cdot\binom{|O|-1}{k-1}p^{k-1}\left(1-p\right)^{|O|-k}\right]\right\}^{m}$$

And the expected probability i does not get help through the network is

$$\sum_{m=0}^{|I|}\left\{\underbrace{\binom{|I|}{m}p^{m}\left(1-p\right)^{|I|-m}}_{\mathbb{P}(n_i(I)=m)}\cdot\underbrace{\left\{1-\sum_{k=1}^{|O|}\left[\frac{1}{k}\cdot\binom{|O|-1}{k-1}p^{k-1}\left(1-p\right)^{|O|-k}\right]\right\}^{m}}_{\mathbb{P}\left(j \text{ helps i}\middle|g_{ij}=1\right)}\right\}$$

First, note that

$$\begin{aligned}\sum_{k=1}^{|O|}\left[\frac{1}{k}\cdot\binom{|O|-1}{k-1}p^{k-1}\left(1-p\right)^{|O|-k}\right]&=\sum_{k=1}^{|O|}\frac{1}{|O|\,p}\binom{|O|}{k}p^{k}\left(1-p\right)^{|O|-k}\\&=\sum_{k=1}^{|O|}\frac{1}{|O|\,p}\binom{|O|}{k}p^{k}\left(1-p\right)^{|O|-k}\\&=\frac{1}{p\,|O|}\left[1-\left(1-p\right)^{|O|}\right]\end{aligned}$$

The expression then becomes

$$\begin{aligned}\sum_{m=0}^{|I|}\left\{\binom{|I|}{m}p^{m}\left(1-p\right)^{|I|-m}\cdot\left[1-\frac{1-\left(1-p\right)^{|O|}}{p\,|O|}\right]^{m}\right\}&=\sum_{m=0}^{|I|}\binom{|I|}{m}\left(p-\frac{1-\left(1-p\right)^{|O|}}{|O|}\right)^{m}\cdot\\&\qquad\cdot\left(1-p\right)^{|I|-m}\\&=\left(p-\frac{1-\left(1-p\right)^{|O|}}{|O|}+\left(1-p\right)\right)^{|I|}\\&=\left[1-\frac{1-\left(1-p\right)^{|O|}}{|O|}\right]^{|I|}\end{aligned}$$

Where the third equality follows by the binomial identity.

So, we get that

$$\sum_{m=0}^{|I|}\left\{\underbrace{\binom{|I|}{m}p^{m}\left(1-p\right)^{|I|-m}}_{\mathbb{P}(n_i(I)=m)}\cdot\underbrace{\left\{1-\sum_{k=1}^{|O|}\left[\frac{1}{k}\cdot\binom{|O|-1}{k-1}p^{k-1}\left(1-p\right)^{|O|-k}\right]\right\}^{m}}_{\mathbb{P}\left(j \text{ helps i}\middle|g_{ij}=1\right)}\right\}=$$

$$=\left[1-\frac{1-\left(1-p\right)^{|O|}}{|O|}\right]^{|I|}$$

and since $|O|=abn$ and $|I|=\left(1-a\right)\left(1-b\right)n$ we have

$$\left[1-\frac{1-\left(1-p\right)^{bn}}{bn}\right]^{(1-b)n}$$

And, since $p=\frac{s}{n}$, if $n\to\infty$, we get that

$$p_i\left(s\right)=1-\lim_{n\to\infty}\left[1-\frac{1-\left(1-p\right)^{bn}}{bn}\right]^{(1-b)n}=1-\exp\left(-\frac{1-b}{b}\left(1-\exp\left(-sb\right)\right)\right)$$

If, instead we are looking at $p_i=\frac{s_i s}{(n-1)s+s_i}$, that is the probability that i forms links, we get that

$$\sum_{m=0}^{|I|}\left\{\underbrace{\binom{|I|}{m}p_i^{m}\left(1-p_i\right)^{|I|-m}}_{\mathbb{P}(n_i(I)=m)}\cdot\underbrace{\left\{1-\sum_{k=1}^{|O|}\left[\frac{1}{k}\cdot\binom{|O|-1}{k-1}p^{k-1}\left(1-p\right)^{|O|-k}\right]\right\}^{m}}_{\mathbb{P}\left(j \text{ helps i}\middle|g_{ij}=1\right)}\right\}$$

$$=\left[1-p_i\frac{1-\left(1-p\right)^{|O|}}{|O|\,p}\right]^{|I|}$$

And

$$p_i\left(s_i,s\right)=1-\lim_{n\to\infty}\left[1-p_i\frac{1-\left(1-p\right)^{bn}}{bnp}\right]^{(1-b)n}=1-\exp\left(-\frac{1-b}{b}\frac{s_i}{s}\left(1-\exp\left(-sb\right)\right)\right)$$

A.2 Proofs

Proof of Proposition 1.3.1. We begin with $\frac{\partial p}{\partial s}\left(s\right)$. When s increases, the expression $1-\exp\left(-sb\right)$ increases, and therefore $p\left(s\right)=1-\exp\left(-\frac{1-b}{b}\left(1-\exp\left(-sb\right)\right)\right)$ increases. Next, the derivative of $p(s)$ with respect to b has the same sign as the derivative of

$$\frac{1-b}{b}\left(1-\exp\left(-sb\right)\right)$$

has with respect to b. This derivative is equal to

$$-\frac{1}{b^2}\left[1-\exp\left(-sb\right)\left(1+sb\left(1-b\right)\right)\right]$$

which is negative if and only if

$$\exp\left(-sb\right)\left(1+sb\left(1-b\right)\right)<1$$

or

$$-sb+\ln\left(1+sb\left(1-b\right)\right)<0$$

Note that $-sb+\ln\left(1+sb\left(1-b\right)\right)$ at $s=0$ is equal to zero, and that

$$\frac{\partial}{\partial s}\left[-sb+\ln\left(1+sb\left(1-b\right)\right)\right]=-\frac{b}{1+sb\left(1-b\right)}\left[b+sb\left(1-b\right)\right]<0$$

Thus, we get that

$$\frac{\partial}{\partial b}p\left(s\right)<0$$

□

Proof of Proposition 1.3.2. Suppose an interior solution exists. Let $\mathbf{s}$ be a strategy profile such that $s_j=s$ for all $j\neq i$. The interior solution s^* is the solution to

$$\left(w-k_{\text{help}}\right)ab\frac{\partial p_i}{\partial s}\left(s\right)=c$$

where

$$p_i\left(s\right)=1-\exp\left(-\frac{1-b}{b}\frac{s_i}{s}\left(1-\exp\left(-sb\right)\right)\right)$$

Thus,

$$\frac{\partial p_i}{\partial s}\left(s^*\right)=\frac{1-b}{bs^*}\left[1-\exp\left(-s^*b\right)\right]\exp\left(-\frac{1-b}{b}\left(1-\exp\left(-s^*b\right)\right)\right)$$

Therefore, the solution is given by

$$\frac{\left(w-k_{\text{help}}\right)(1-b)}{s^*}\left[1-\exp\left(-s^*b\right)\right]\exp\left(-\frac{1-b}{ab}\left(1-\exp\left(-s^*b\right)\right)\right)=c$$

Now, the LHS decreases in s^* as

1. $$\frac{\partial}{\partial s^*}\left[\frac{1-\exp\left(-s^*b\right)}{s^*}\right]=\frac{\left[-1+\exp\left(-s^*b\right)\left(1+s^*b\right)\right]}{\left(s^*\right)^2}<0$$

   since the numerator is negative as seen in the proof of Proposition 1.3.1.

2. $\exp\left(-\frac{1-b}{b}\left(1-\exp\left(-s^*b\right)\right)\right)$ decreases in s^* as well.

Also note that

$$\lim_{s^*\to 0}\frac{\left(w-k_{\text{help}}\right)(1-b)}{s^*}\left[1-\exp\left(-s^*b\right)\right]\exp\left(-\frac{1-b}{b}\left(1-\exp\left(-s^*b\right)\right)\right)=$$
$$=\left(w-k_{\text{help}}\right)(1-b)\,b$$

And that

$$\lim_{s^*\to \infty}\frac{\left(w-k_{\text{help}}\right)(1-b)}{s^*}\left[1-\exp\left(-s^*b\right)\right]\exp\left(-\frac{1-b}{b}\left(1-\exp\left(-s^*b\right)\right)\right)=0$$

Thus, as long as $c<\left(w-k_{\text{help}}\right)(1-b)\,b$ we get that an interior solution exists and it is unique.

□

Proof of Proposition 1.3.3. Suppose an interior solution exists. Let $\mathbf{s}$ be a strategy profile such that $s_j=s$ for all $j\neq i$. The interior solution s^* is the solution to

$$\left(v-k_{\text{help}}\right)b\frac{\partial p_i}{\partial s}(s)=c$$

where

$$p(s)=1-\exp\left(-\frac{1-b}{b}\left(1-\exp\left(-sb\right)\right)\right)$$

Thus,

$$\frac{\partial p}{\partial s}\left(s^*\right)=(1-b)\exp\left(-s^*b\right)\exp\left(-\frac{1-b}{b}\left(1-\exp\left(-s^*b\right)\right)\right)$$

Therefore, the solution is given by

$$\left(v-k_{\text{help}}\right)(1-b)\, b \exp\left(-s^{*} b\right) \exp\left(-\frac{1-b}{b}\left(1-\exp\left(-s^{*} b\right)\right)\right)=c$$

Now, the LHS decreases in s^{*} as

1. $\exp\left(-s^{*} b\right)$ decreases in s^{*}
2. $\exp\left(-\frac{1-b}{b}\left(1-\exp\left(-s^{*} b\right)\right)\right)$ decreases in s^{*} as well

Also note that

$$\lim_{s^{*} \rightarrow 0}\left(v-k_{\text{help}}\right)(1-b)\, b\left[1-\exp\left(-s^{*} b\right)\right] \exp\left(-\frac{1-b}{b}\left(1-\exp\left(-s^{*} b\right)\right)\right)= \\ =\left(v-k_{\text{help}}\right)(1-b)\, b$$

And that

$$\lim_{s^{*} \rightarrow \infty}(1-b)\, b\left[1-\exp\left(-s^{*} b\right)\right] \exp\left(-\frac{1-b}{b}\left(1-\exp\left(-s^{*} b\right)\right)\right)=0$$

Thus, as long as $c<\left(v-k_{\text{help}}\right)(1-b)\, b$ we get that an interior solution exists and it is unique. □

Proof of Proposition 1.3.4. First, the IR constraint holds with equality. This implies that

$$f=c s^{*}+k b p\left(s^{*}\right)+w\left(1-b\left(1-p\left(s^{*}\right)\right)\right)$$

Let w^{*} be the solution to the first order condition to the firm's problem (1.3) after substituting f in, if exists. That is, w^{*} solves:

$$\frac{\partial s^{*}}{\partial w}\left[(v-k)\, b \frac{\partial p}{\partial s}\left(s^{*}\right)-c\right]=0$$

where s^{*} is the solution to the workers' problem and

$$\frac{\partial s^{*}}{\partial w}=\frac{1-\exp\left(-s^{*} b\right)}{(w-k)\left[\frac{1-\exp\left(-s^{*} b\right)\left(1+s^{*} b\right)}{s^{*}}+(1-b) \exp\left(-s^{*} b\right)\left(1-\exp\left(-s^{*} b\right)\right)\right]}$$

After plugging in the workers' solutions and rearranging it is equivalent to

$$\frac{\partial s^*}{\partial w}\cdot c\left[\frac{v-k}{w-k}\cdot\frac{\exp\left(-s^*b\right)s^*b}{1-\exp\left(-s^*b\right)}-1\right]$$

Since the firm internalizes the negative externalities workers impose on each other[1] Therefore,

$$\frac{\partial p}{\partial s}\left(s^*\right)=\left(1-b\right)\exp\left(-\frac{1-b}{b}\left(1-\exp\left(-s^*b\right)\right)\right)\exp\left(-s^*b\right)$$

next, note that for $w^*\leq k$ we get $s^*=0$ which is not optimal. Also, the firm strictly prefers choosing $w^*\leq v$ over choosing $w^*>v$, and therefore we can focus on the interval $[k,v]$.

We begin by showing that a solution exists:

1. For $w=v$, we get that the LHS is equal to $\frac{\exp(-s^*b)s^*b}{1-\exp(-s^*b)}<0$ since $1-\exp\left(-s^*b\right)\left(1+s^*b\right)>0$ as was shown in Proposition 1.3.1.
2. For $w\to k$, we get $\underbrace{\frac{v-k}{w-k}}_{\to\infty}\cdot\underbrace{\frac{\exp\left(-s^*b\right)s^*b}{1-\exp\left(-s^*b\right)}}_{>0}>0$

Thus, continuity implies that a solution w^* exists.

In order to show that this solution is unique, it is enough to show that the second order condition is negative at any point were the FOC is satisfied.[2] By the envelop theorem, it is equivalent to showing that

$$\frac{\partial}{\partial w}\left[\frac{v-k}{w-k}\cdot\frac{\exp\left(-s^*b\right)s^*b}{1-\exp\left(-s^*b\right)}-1\right]<0$$

The second order condition satisfies

$$\text{SOC}\left(w^*\right)\propto-\frac{v-k}{\left(w-k\right)^2}\cdot\frac{\exp\left(-s^*b\right)s^*b}{1-\exp\left(-s^*b\right)}+\frac{v-k}{w-k}\frac{\partial s^*}{\partial w}\left[\frac{\exp\left(-s^*b\right)b\left[-s^*b+1-\exp\left(-s^*b\right)\right]}{\left(1-\exp\left(-s^*b\right)\right)^2}\right]$$

Thus, since

$$-\frac{v-k}{\left(w-k\right)^2}\cdot\frac{\exp\left(-s^*b\right)s^*b}{1-\exp\left(-s^*b\right)}<0$$

[1] Having more friends crowds out helps from one's friends of friends (i.e., workers who are a distance two from each other.)

[2] This implies that at any point w such that $\text{FOC}(w)=0$ we have that the derivative is negative, and therefore it must be the cases that there exists a unique solution. The reason is that there cannot be multiple interior maxima, and no minima.

and

$$\frac{v-k}{w-k}\frac{\partial s^*}{\partial w}\frac{1}{\left(1-\exp\left(-s^*b\right)\right)^2}>0$$

it is sufficient to show that

$$\exp\left(-s^*b\right)b\left[-s^*b+1-\exp\left(-s^*b\right)\right]<0$$

However, when $s^*=0$ we get that the expression on the LHS is equal to 0, and

$$\frac{\partial}{\partial s^*}\left[\exp\left(-s^*b\right)b\left[-s^*b+1-\exp\left(-s^*b\right)\right]\right]=b\left[-1+\exp\left(-s^*b\right)\right]<0$$

And since $s^*>0$ we get that $\exp\left(-s^*b\right)b\left[-s^*b+1-\exp\left(-s^*b\right)\right]<0$.

Also, since w^* solves

$$\left(v-k\right)b\frac{\partial p}{\partial s}\left(s^*\right)-c=0$$

which is the social planner's first order condition, we know that $s^*\left(w^*\right)=s^*_{SP}$, and the first best is achieved. □

Proof of Proposition 1.3.5. We begin by totally differentiating equation (1.1) with respect to b. We get that

$$\frac{\partial s^*}{\partial b}=\frac{\exp\left(-s^*b\right)\left[1+\left(1-b\right)s^*\right]-1+\frac{(1-b)}{b^2}\left(1-\exp\left(-s^*b\right)\right)\left(1-\exp\left(-s^*b\right)-\left(1-b\right)bs^*\exp\left(-s^*b\right)\right)}{\left(1-b\right)\left[\frac{1-\exp(-s^*b)[1+bs^*]}{s^*}+\left(1-b\right)\left(1-\exp\left(-s^*b\right)\right)\exp\left(-s^*b\right)\right]}$$

Since the denominator is always positive it is enough to show that the numerator changes its sign at some point $\bar{b}$ and that this is the only point.

First note that, since s^* is strictly positive by the definition of a symmetric equilibrium,[3]

$$\left.\frac{\partial s^*}{\partial b}\right|_{b\approx 0}=s^*>0$$

and that

$$\left.\frac{\partial s^*}{\partial b}\right|_{a\approx 1}=\exp\left(-s^*\right)-1<0$$

Now, since the numerator is continuous, we get that $\frac{\partial s^*}{\partial b}$ crosses the x-axis at least once. In order to show that it crosses it exactly once it is sufficient to show that at any such point the

[3] At $b=0$ or $b=1$ the workers do not exert any effort since they know they either won't find help, or won't need it. On the other hand, when $b\approx 0$ or $b\approx 1$, we get that $s^*>0$, and these inequalities hold.

second derivative of the numerator is negative.

We therefore take the second derivative of the numerator at the point $\frac{\partial s^*}{\partial b}$ and we get

$$- \exp\left(-s^* b\right) s^* \left[2 + \left(1-b\right) s^*\right] \underbrace{\left[1 - \frac{\left(1-b\right)}{b}\left(1 - \exp\left(-s^* b\right)\right)\right]}_{>0} -$$

$$- \frac{\left(2-b\right)}{b^3} \underbrace{\left(1 - \exp\left(-s^* b\right)\left(1 + \left(1-b\right) b s^*\right)\right)}_{>0} \underbrace{\left[1 - \exp\left(-s^* b\right)\left(1 + \frac{\left(1-b\right) b s^*}{2-b}\right)\right]}_{>0} < 0$$

Denote by $\bar{b}$ the point at which $\frac{\partial s^*}{\partial b} = 0$. Thus, by continuity, s^* is increasing on $\left(0, \bar{b}\right)$ and decreases on $\left(\bar{b}, 1\right)$. $\square$

Proof of Proposition 1.4.1. Note that $p\left(s\right)$ decreases in δ for $b \leq \frac{1}{2}$, and follows an inverted U-shape otherwise. $\square$

Appendix B

Appendix to Chapter 2

B.1 Proofs

Proof of Proposition 2.2.1. The difference in the output over the two periods, ΔY_i is given by:

$$\begin{aligned}\Delta Y_i &= Y_i^2 - Y_i^1 \\ &= \left\{ I_i(2) - I_i(1) + \lambda \sum_{l=1}^{n} w_{il} \left\{ g_{il}(2) \left[\alpha_l + I_l(2) \right] - g_{il}(1) \left[\alpha_l + I_l(1) \right] \right\} \right\} \\ &= I_i(2) - I_i(1) + \lambda w_{ij} \left[g_{ij}(2) \left(\alpha_j + I_j(2) \right) - g_{ij}(1) \left(\alpha_j + I_j(1) \right) \right] \end{aligned} \tag{B.1}$$

where the third equality follows from the assumption that the relocation of inventor j affects inventor i only in a direct way.

When inventor i and inventor j cut all types of their links – $g_{ij}(2) = 0$ and $s_{ij}(1) = 0$ – inventors i does not acquire any information in the second period and therefore $I_i(2) = I_i(1)$. Therefore, equation (B.1) becomes

$$\Delta Y_i = -\lambda w_{ij} \left(\alpha_j + I_i(1) \right) < 0$$

On the other hand, as long as $s_{ij} = 1$, and if the information acquired in the second period is high enough, such that

$$I_i(2) - I_i(1) \geq \lambda w_{ij} \left(\alpha_j + I_i(1) \right)$$

it follows that $\Delta Y_i \geq 0$. $\square$

Proof of Proposition 2.2.2. Under the assumptions made in this propositions, the effect on the total output of inventor i is given by

$$\Delta Y_i \left(N_j\left(2\right)\right) = I_i\left(2\right) - I_i\left(1\right)$$

and therefore, Under the assumptions made in this propositions, the effect on the total output of inventor i is given by

$$\Delta Y_i \left(N_j\left(2\right)\right) - \Delta Y_i \left(\tilde{N}_j\left(2\right)\right) = \left[I_i\left(2\right) - I_i\left(1\right)\right] - \left[\tilde{I}_i\left(2\right) - \tilde{I}_i\left(1\right)\right] = I_i\left(2\right) - \tilde{I}_i\left(2\right)$$

where the last equality follows since the information gathered in the first period is equal across these two scenarios by definition.

Now, if information sharing between inventor j and inventor i does not take place, then $s_{ij}\left(2\right) = 0$ in both these cases, and no information is acquired in the second period, regardless of connections inventor j's forms. Hence,

$$\Delta Y_i \left(N_j\left(2\right)\right) - \Delta Y_i \left(\tilde{N}_j\left(2\right)\right) = 0$$

If, on the other hand, the inventors engage in information sharing, then $s_{ij}\left(2\right) = 1$, and

$$\begin{aligned} I_i\left(2\right) - \tilde{I}_i\left(2\right) &= \sum_{m \in N_j(2) \setminus N_i(1)} \left[\bar{w}_{im}\left(2\right) - \bar{w}_{im}\left(1\right)\right] k_m - \sum_{m \in \tilde{N}_j(2) \setminus N_i(1)} \left[\bar{w}_{im}\left(2\right) - \bar{w}_{im}\left(1\right)\right] k_m \\ &= \sum_{m \in N_j(2) \setminus \left(\tilde{N}_j(2) \cup N_i(1)\right)} \left[\bar{w}_{im}\left(2\right) - \bar{w}_{im}\left(1\right)\right] k_m \end{aligned}$$

Where the first equality follows by the assumption that inventor i's direct connections do not change (besides potentially that with inventor j), and Assumption 2.2.1 which implies that as long as the weight placed on the path leading from inventor i to their indirect collaborators did not increase, inventor i does not acquire information through them. This implicitly means that even if inventor j creates a direct link to one of inventor i's direct connection, further information is not acquired through that link.

And as long as $N_j(2) \setminus \left(\tilde{N}_j(2) \cup N_i(1)\right) \neq \varnothing$, the effect on inventor i's output is positive:

$$I_i(2) - \tilde{I}_i(2) > 0$$

□

Proof of Proposition 2.2.3. In this case, the innate ability of inventors i and j changes, but the information they acquire does not depend on the set of inventors they interact with after the move and

$$I_i(2) = \tilde{I}_j(2)$$

Therefore,

$$\Delta Y_i\left(N_j(2)\right) = \left\{\left[\alpha_i(2) - \alpha_i(1)\right] + \lambda w_{ij}\left[g_{ij}(2)\,\alpha_j(2) - g_{ij}(1)\,\alpha_j(1)\right]\right\} = \Delta Y_i\left(\tilde{N}_j(2)\right)$$

□

Proof of Proposition 2.2.4. By definition, the information acquired by inventor i in both periods, depends on the strength of the relationship between inventor i and inventor j. On the other hand, the information acquired by inventor j depends on w_{ji} that can be different than w_{ij} and hence does not change, by assumption. Therefore,

$$\Delta Y_i\left(w_{ij}\right) = I_i(2) - I_i(1) + \lambda w_{ij}\underbrace{\left\{g_{ij}(2)\left[\alpha_j + I_j(2)\right] - g_{ij}(1)\left(\alpha_j + I_j(1)\right]\right\}}_{X}$$

Similarly,

$$\begin{aligned}\Delta Y_i\left(\tilde{w}_{ij}\right) &= \tilde{I}_i(2) - \tilde{I}_i(1) + \lambda\tilde{w}_{ij}\left\{g_{ij}(2)\left[\alpha_j + I_j(2)\right] - g_{ij}(1)\left(\alpha_j + I_j(1)\right]\right\}\\ &= \tilde{I}_i(2) - \tilde{I}_i(1) + \lambda\tilde{w}_{ij}X\end{aligned}$$

Now, following equation (2.1) and the assumption that inventor j is the only inventors moving in the network, the new information can only be acquired through new connections inventor j forms, and as long as they are (and were) not directly connected to inventor i.

Thus,

$$I_i(2) = I_i(1) + \sum_{l=1}^{n} s_{ij}(2) \cdot s_{jl}(2) \cdot (1 - s_{il}(1)) \cdot \left(1 - s_{jl}(1)\right) w_{ij} \cdot w_{jl}$$
$$\tilde{I}_i(2) = \tilde{I}_i(1) + \sum_{l=1}^{n} s_{ij}(2) \cdot s_{jl}(2) \cdot (1 - s_{il}(1)) \cdot \left(1 - s_{jl}(1)\right) \tilde{w}_{ij} \cdot w_{jl}$$

And,

$$I_i(2) - I_i(1) = w_{ij} \underbrace{\sum_{l=1}^{n} s_{ij}(2) \cdot s_{jl}(2) \cdot (1 - s_{il}(1)) \cdot \left(1 - s_{jl}(1)\right) w_{jl}}_{Z}$$

And similarly for $\tilde{I}_i(2) - \tilde{I}_i(1)$.

This means we can express the effect on inventor i output as:

$$\Delta Y_i\left(w_{ij}\right) = w_{ij} \left\{ \sum_{l=1}^{n} s_{ij}(2) \cdot s_{jl}(2) \cdot (1 - s_{il}(1)) \cdot \left(1 - s_{jl}(1)\right) w_{jl} + \lambda X \right\}$$

Therefore,

$$\begin{aligned} \left|\Delta Y_i\left(w_{ij}\right)\right| - \left|\Delta Y_i\left(\tilde{w}_{ij}\right)\right| &= w_{ij} \left|Z + \lambda X\right| - \tilde{w}_{ij} \left|Z + \lambda X\right| \\ &= \left(w_{ij} - \tilde{w}_{ij}\right) \left|Z + \lambda X\right| \geq 0 \end{aligned}$$

with a strict inequality as long as the benefit from the information gained through inventor j and the cost of discontinuing co-patenting with them do not exactly cancel out. In the case where inventor i and inventor j cut all of their links (informational and co-patenting) or when they continue collaborating, the difference will also be strictly positive. □

B.2 Model Extensions

In this section, I present the model without the assumption imposed on the distance the information can travel. I spell the whole model again, although some of the parts did not change to maintain some level of continuity.

The main goal of these extension is to emphasize that the predictions do not depend on this assumption, and that everything else follows.

B.2.1 Basic Framework

Inventors' Network

A society of n inventors is connected via a directed and weighted network, which has an adjacency matrix $\mathbf{W} \in [0,1]^{n\times n}$. A general element $w_{ij} \in [0,1]$ represents the status and the strength of the relationship between inventor i and inventor j, where a higher w_{ij} is associated with a stronger connection.[1] Specifically, an entry $w_{ij} = 0$ implies that inventor i and inventor j do not collaborate, and therefore are not connected.

Inventors are also endowed with an innate ability level $\alpha_i \in \mathbb{R}_+$ and with a knowledge level $k_i \in \mathbb{R}_+$. These concepts play a central role in patent production and information sharing.

Information Acquisition

Inventors acquire information through their network, not just from their immediate connections, but also from inventors located farther away. To formalize this concept, it is helpful to introduce a notion that measures the distance between any to inventors in the network. Let d_{ij} be the length of the shortest path between inventor i and inventor j in the network $\mathbf{W}$. This distance metric signifies the smallest number of connected inventors forming a sequence that establishes an indirect link between inventors i and j. Formally, a path of

[1] One way to interpret the weights of the form w_{ij} is through the eyes of patents production relationship. In this case, a higher w_{ij} corresponds to a higher number of patents both inventor i and inventor j are jointly listed on.

length m between inventor i and inventor j is an ordered set $M = \{i_1, i_2, ..., i_{m+1}\}$ such that $w_{i_l i_{l+1}} > 0$ for all $l \in \{1, ..., m\}$, $i_1 = i$ and $i_{m+1} = j$. The length of the shortest path between inventor i and j is the minimal m that satisfies these conditions. The weight of this path w_{ij}^M is given by multiplying the weights of the links composing this path, which can be expressed as

$$w_{ij}^M = \Pi_{l=1}^{m} w_{i_l i_{l+1}}$$

The total information held by inventor i is a result of a combination between their initial knowledge and what they acquire through interactions with other inventors. Let $D \in \{1, ..., n-1\}$ be a bound on the distance information can travel on the network.[2] With this concept and the notation introduced earlier, the total information held by inventor i is given by:

$$I_i = k_i + \sum_{d=1}^{D} \sum_{\{j:\, d_{ji}=d\}} \bar{w}_{ji} k_j \tag{B.2}$$

where $\bar{w}_{ji}$ corresponds to the cumulative weight of the shortest paths connecting j and i, which also meet the condition of being of a minimal distance d.[3] These weights represent the strengths of the paths, and they capture the idea that the information inventor i acquires through inventor j is proportional to a measure of dependence between them.

Figure B.1 provides two examples that illustrate the bound D and the cumulative weight of the paths between inventor i and inventor j, $\bar{w}_{ij}$. The weights above the links represent the clockwise connections, while those below correspond to the counter-clockwise links.[4] In panel (a), the bound on the distance between inventors for information acquisition purposes

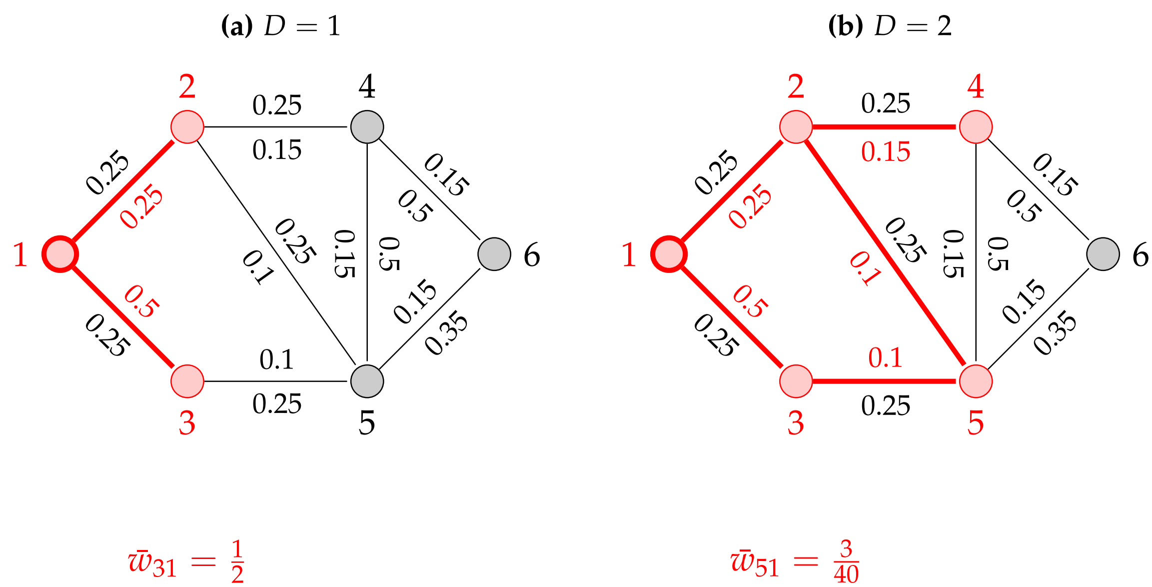
[2]This meant to express the idea that when two inventors are positioned far apart within the network, they are less likely to share information with each other. The parameter D specifies what "far apart" means in this context.

[3]When inventors i and j are directly connected the cumulative weight $\bar{w}_{ij}$ is equal to the strength of the link between inventors i and j, w_{ij}. When the shortest path between inventors i and j is of length two, $\bar{w}_{ij}$ is equal to the sum of the weights of the form $w_{il} \cdot w_{lj}$ where inventor l is directly connected to both inventor i and inventor j, but inventors i and j are not connected to each other. Formally, $\bar{w}_{ij} = \sum_{l=1}^{n} w_{il} w_{lj} \cdot \mathbb{1}\left\{w_{ij} = 0\right\}$. The formula for longer paths follows the same reasoning, but it becomes more cumbersome as the minimal distance d increases, since there are more conditions to verify a path is indeed the shortest.

[4]For example, in panel (a), the weight of the link going from inventor j to inventor i, denoted as w_{ji}, is 0.5, whereas the weight of the link from inventor i to inventor j, represented as w_{ij}, is 0.25.

is equal to one. In this scenario, inventor 1 gains information from inventors 2 and 3, and this information is scaled by $w_{21} = 0.25$ and $w_{31} = 0.25$, respectively. In panel (b), inventor 1 gains information from inventors 2, 3, 4 and 5, and the information acquired through inventor 5 is scaled by 0.5.

Figure B.1: Information Sharing on Network Example



Patent Production Function

Inventors produce patents in teams. Each inventor's total output relies on their individual output, which is determined by their innate ability and total information they hold, as well as the output contributed by their collaborators.

Inventor i's total output is, therefore, given by

$$Y_i = y_i + \sum_{j=1}^{n} w_{ji} y_j \tag{B.3}$$

$$y_i = \alpha_i + I_i$$

where y_i is inventor i's individual output.

This production function reflects the substitution between the different sources that drive output production. It emphasizes the tradeoff between the information accessed through the network (and one's fixed innate ability) and the direct benefit one gains from co-patenting, which comes in the form of the output contributed by their collaborators. This tradeoff will be the main focus of the next subsection.

B.2.2 Two Period Model

In general, patents are produce in various geographical locations. As a consequence, inventors may move around. In this section, I study the effect of a relocation on the total output of the mover's former collaborators in the eyes of the model. Specifically, the magnitude and direction of this effect will be contingent upon the specifics of the connections between these inventors, and on how their network changes in response to the relocation.

To start with, consider two time periods and two geographical locations. Let $t = 1$ represent the time before any relocation occurs, and $t = 2$ reflect the time after the relocations have take place, and assume that innate abilities α_i and the weighted adjacency matrix $\mathbf{W}$ are fixed across periods. Since the collaboration network can undergo changes between the two time period, I introduce new notations which capture the state of the network in each one of these periods. Denote by $\mathbf{G}\left(t\right) \in \{0,1\}^{n \times n}$ the undirected adjacency matrix at time t. The ij-th entry is the collaboration status between inventor i and inventor j at time t, with the entry equals one if they co-patent, and zero otherwise. This relationship is reciprocal. Additionally, let $\mathbf{S}\left(t\right) \in \{0,1\}^{n \times n}$ be the undirected information exchange network at time t. This is a symmetric matrix whose entries equal to one whenever the inventors are engaged in information sharing. In the initial period, information sharing occurs exclusively when inventors co-patent. However, after the relocation, in period $t = 2$, inventors who previously co-patented in $t = 1$ may cease their collaboration in period $t = 2$, and yet still engage in information sharing. Formally,

$$
\begin{aligned}
g_{ij}(1) &\iff s_{ij}(1) = 1 && \forall i, j \\
g_{ij}(2) &\implies s_{ij}(2) = 1 && \forall i, j
\end{aligned}
$$

Lastly, $I_i(t)$ is the level of the information held by inventor i at time t, where the paths are now measured on the network $\mathbf{S}(t)$.[5]

Information Acquisition

To accommodate some degree of continuity across the two periods, I assume that

Assumption B.2.1. The information inventors acquired in the first period cannot be forgotten, and therefore is not subject to relearning.

This implies that inventor i begins period $t = 2$ with knowledge that is equal to $I_i(1)$, rather than k_i, as it is at the beginning of period $t = 1$.

The idea behind this assumption is that once techniques and ideas are acquired, they can't be unlearned. Once learned, inventors can use them again without relearning.

In particular, equation (2.1) becomes

$$
\begin{aligned}
I_i(1) &= k_i + \sum_{d=1}^{D} \sum_{\{j:\, d_{ji}(1)=d\}} \bar{w}_{ji}(1)\, k_j && \forall i \\
I_i(2) &= I_i(1) + \sum_{\{j:\, d_{ji}(2)=d\}} \mathbb{1}\left\{\bar{w}_{ji}(2) > \bar{w}_{ji}(1)\right] \cdot \left[\bar{w}_{ji}(2) - \bar{w}_{ji}(1)\right] k_j && \forall i \qquad \text{(B.4)}
\end{aligned}
$$

where $d_{ji}(t)$ denotes the minimal distance on the matrix $\mathbf{S}(t)$ and $\bar{w}_{ji}(t)$ corresponds to the cumulative paths weights on matrix $\mathbf{S}(t)$.[6] The multiplication by the elements $\mathbb{1}\left\{\bar{w}_{ji}(2) > \bar{w}_{ji}(1)\right\}$ imposes the restriction that information can only be acquired in the

---

[5] In the first period, the elements of the matrix $\mathbf{G}(1)$ and $\mathbf{S}(1)$ are equivalent to the indicators $\mathbb{1}\left\{w_{ij} > 0\right\}$. Therefore, the paths on $\mathbf{W}$, $\mathbf{G}(1)$ and on $\mathbf{S}(1)$ are the same. However, in the second period, since information exchange can take place even when then inventors do not collaborate, it does not necessarily hold.

[6] Although the matrix $\mathbf{W}$ is fixed across the two time periods, the potential differences in the matrices $\mathbf{S}(1)$ and $\mathbf{S}(t)$ can lead to differences in $\bar{w}_{ji}(t)$ across the time periods, if new paths are formed and/or old paths are severed.

second period, and cannot be forgotten.

Patent Production

Production in both periods follows the same reasoning as in equation (2.2), with total output depending on both individual and scaled collaborative outputs. That is,

$$
\begin{aligned}
Y_i(t) &= y_i(t) + \sum_{j=1}^{n} g_{ji}(t)\, w_{ji} y_j && \forall i \\
y_i(t) &= \alpha_i + I_i(t) && \forall i
\end{aligned}
$$

B.3 Data

B.3.1 Description of Patent Data

The information about the patents I use is from the USPTO PatentsView. Besides supplying information about inventors and patents, they also conduct a disambiguation procedure where each inventor gets a unique identifier number, and all of their patents and the information provided at the time of the application are attached to it. This is not trivial as inventors might use different names at the time of the application and are not necessarily patenting under the same assignee or in the same location.

The data I use is patent applications for patents which were eventually granted between 1976 and 2022. The data is provided in a TSV format, where each inventor, location and patent has a unique identifier. Using these identifiers, I can merge all the information given in an application about a specific inventor. In particular, the inventor's name, the names of the other inventors who are listed with them on the patent, the date of the application as well as the date at which the patent was granted, the number of citations the patent received and which patents cited it, the citations granted by that patent, the CPC classes of the patents, the residential address of the inventor and the assignee and the headquarters's address.

B.3.2 Construction of the Sample

I restrict the sample of the USPTO patent application to include only inventors whose first patent was applied for on 1990 or after. The reason is that online professional profiles were introduced in 2008, and older people, mainly ones who have already retired or are close to retirement, are less likely to have an account due to its purpose being a device that makes it easier to learn about individuals' work history and potentially ease hiring and recruitment. That way I can ensure, with a higher likelihood, that the linking rates are not biased by the probability of opening an account.

I also exclude a very small number of patent numbers who were "withdrawn," which can be found in https://www.uspto.gov/patents/search/withdrawn-patent-numbers. Patent numbers are assigneed before the patents are granted. If between the date at which the patent number was assigned and the date of the issuance of the patent some information that indicates that the application is not ready to be issued is reveled, the patent is not granted at that date and the number will never be used again.[7] It is important to note, however, that the application can still be issued, and in this case, it will be issued under a different patent number. Therefore, dropping the withdrawn patent numbers have two purposes. The first, is to avoid including ideas that were not patented after all. And second, to avoid double counting patents, in cases where they were issued under different patent numbers.

Another point is about how to assign CPC classes to patents. When patents are granted and published, more than one CPC class is usually assigned. Following the general method in the literature, the CPC class I assign to the patent is the one listed first in the sequence.

B.3.3 Linking Algorithm to Construct the Dataset

Both the USPTO data and the data I receive through Revelio Labs have different advantages. While the USPTO dataset provides me with information about the patenting activity of

[7] This may happen if, for example, the fees were not paid in full or if some prior art was found.

inventors, and their physical location at the time of the application, the information that includes data from online professional profiles, adds to this demographics, work locations and the company workers are employed by at each point in time. Only the combination of both these datasets will allow me to construct the panel that I need in order to study the effect of inventor mobility on their co-inventor productivity.

I leverage as much information as I can from both these dataset to construct the linking in a way that minimizes mistakes. There are two types of possible mistakes. The first is about not linking an inventor that should have otherwise be linked. These mistakes, although costly in term of number of observations, are less likely to bias the data. The type of mistakes that I try most to avoid are the ones around linking an inventor and a user that should not be linked. And my methodology, focuses mainly on the second, although I try to account for both.

The linking is performed on three identifiers. The first is the inventor and the user names. As a first step I only consider the first and last name of the inventor and the user. An inventor is linked to a user if and only if their first and last name match exactly. This implies that there are some inventors I cannot match due to usage of different names on their patent application and their online professional profile. An example is an inventors who uses their full name on their patent and their nickname on their online professional profile.[8] The second element I link on is the state the inventor lived in at the time of the application, and the state the user lists as the location of their workplace at a six-month window around the application data.[9] Linking on the state level can create issues involving the difference between workplace location and residential address. At the application stage inventors are asked to provide their residential address. That is, they report their home address. On online professional profiles, users provide the state at which their work takes place. That is, the location of their office. When inventors live and work in different

[8]Some of the common names and nicknames are “Robert” and “Bob” or “William” and “Bill.”

[9]Some users take time to update their online professional profiles with their new employers. In order to allow for some flexibility, I allow the state to match even in an earlier or a later stage.

states, they will not be linked to their user. The last characteristics I link inventors and users on is their employer. I begin with standardizing the USPTO assignees and the online professional profiles companies following the standardizing method in the NBER's Patent Data Project (https://sites.google.com/ site/patentdataproject/). For example, "AMAZON" and "AMAZON USA" are the same company. I also do not use any abbreviations for companies. That is, whenever a company is known to be abbreviated, such as in "IBM," I change all of its instances to "INTERNATIONAL BUSINESS MACHINES" to avoid mismatched for that reason. I, then, use a fuzzy match, linking between the standardize assignee on the inventor's application and the standardize user's employer at the time of the application. More precisely, I use Jaro-Winkler distance measure and set the threshold for match at 0.99. This is a quite high threshold, ensuring that I so not mismatch between different companies and assignees. Finally, I include only inventors who are linked to exactly one user, and I use middle initials to break ties.

There are at least two issues which follow from the last step. The first is that the assignee is not necessarily the employer of all the inventors on the patent. It is usually the case that at least one inventor is employed by the company listed as an assignee, but it is not necessarily the case that all of the inventors are.[10] In these cases, the inventor and the user will not be matched. The second issue might arise in cases where mergers are acquisitions are involved. In these cases users in online professional profiles, might update their user to include the information about the acquiring firm rather than the name of the firm they were employed in prior to the acquisition. If at the time of the application the merger or acquisition has not taken place yet, there will be inconsistency between the information on the online professional profile and the information on their USPTO application.

Characteristics of Linked and Unlined Inventors

Table B.1 presents the summary statistics of the linked and the full samples. It shows that the linked inventors tend to be more productive, and that is due to the fact that linking is

[10] There are also cases where more than once assignee is listed, and in this case I check all the combinations.

easier the more observations there are as it increases the accuracy.

Table B.1: Summary Statistics on Full and Linked Samples

(a) Panel A: Full Sample

	Mean	Median	Std. Dev	# Obs.
Year of First Patent	2004	2004	8	1,150,368
Total Number of Patents	4.39	2.00	7.06	1,150,368
Total Citations Stock	62.80	11.00	159.30	1,150,368
Total Adjusted Citations Stock	3.19	0.96	6.24	1,150,368
Average Team Size	2.75	2.25	1.85	1,150,368
Male	0.85	1.00	0.35	1,048,732

(b) Panel B: Linked Sample

	Mean	Median	Std. Dev	# Obs.
Year of First Patent	2007	2009	9	229,290
Total Number of Patents	6.41	3.00	9.18	229,290
Total Citations Stock	78.82	8.00	200.24	229,290
Total Adjusted Citations Stock	4.34	1.26	7.75	229,290
Average Team Size	3.01	2.60	1.85	229,290
Male	0.85	1.00	0.35	207,203

Notes: The table provides summary statistics for two datasets. Panel A presents the attributes of inventors based in the United States within the patent data, encompassing inventors who initiated patenting activities for the first time after 1990 and never indicated a residential address outside of the United States. Panel B offers summary statistics for the inventors successfully matched to the Revelio Labs data. All variables represent cumulative values over the entire observed period. The Average Team Size is computed across all patents, with a solo patent being considered as a team size of one.

Adding to the information in Table B.1 the BEA's "economic areas" with the lowest linking rates are Great Falls MT, Aberdeen SD, Cape Girardeau-Jackson MO-IL, Lewiston ID-WA and Panama City-Lynn Haven FL. And the BEA's "economic areas" with the highest linking rates are Seattle-Tacoma-Olympia WA, San Jose-San Francisco-Oakland CA, San Diego-Carlsbad-San Marcos CA, Milwaukee-Racine-Waukesha WI, Minneapolis-St. Paul-St. Cloud MN-WI, Raleigh-Durham-Cary NC and Boston-Worcester-Manchester, MA-NH. The latter are more associated with patenting activity.

Moreover, as anticipated, younger inventors have more incentives and are more likely to setup an online account. For that reason, the years of the first patent which is associated with the lowest linking rates are 1990-2004, where the linking rates are monotonically decreasing with the years. And the highest linking rates correspond to inventors patenting for the first time in the years 2013-2020. Where, in this case, the linking rates is monotonically increasing in the years.

The last characteristic I looked at in the CPC classes the linked and and unlinked inventors patent in. The lowest rated of linking are associated with the following CPC classes:

Table B.2: Lowest Linking Rates CPC Classes

CPC Class	Title
B43	WRITING OR DRAWING IMPLEMENTS; BUREAU ACCESSORIES
B44	DECORATIVE ARTS
B63	SHIPS OR OTHER WATERBORNE VESSELS; RELATED EQUIPMENT
B68	SADDLERY; UPHOLSTERY
A45	HAND OR TRAVELLING ARTICLES
A46	BRUSHWARE

Table B.3: Highest Linking Rates CPC Classes

CPC Class	Title
B06	GENERATING OR TRANSMITTING MECHANICAL VIBRATIONS IN GENERAL
B33	ADDITIVE MANUFACTURING TECHNOLOGY
C07	ORGANIC CHEMISTRY
F15	FLUID-PRESSURE ACTUATORS; HYDRAULICS OR PNEUMATICS IN GENERAL
G06	COMPUTING; CALCULATING OR COUNTING
G16	INFORMATION AND COMMUNICATION TECHNOLOGY [ICT] SPECIALLY ADAPTED FOR SPECIFIC APPLICATION FIELDS
H03	ELECTRONIC CIRCUITRY
H04	ELECTRIC COMMUNICATION TECHNIQUE

Note that these CPC classes are aligned with the locations of the linked or unlinked inventors, respectively. To put differently, the main CPC classes in these areas match the CPC classes with the highest or lowest linking rates, respectively.

B.4 Additional Figures and Tables

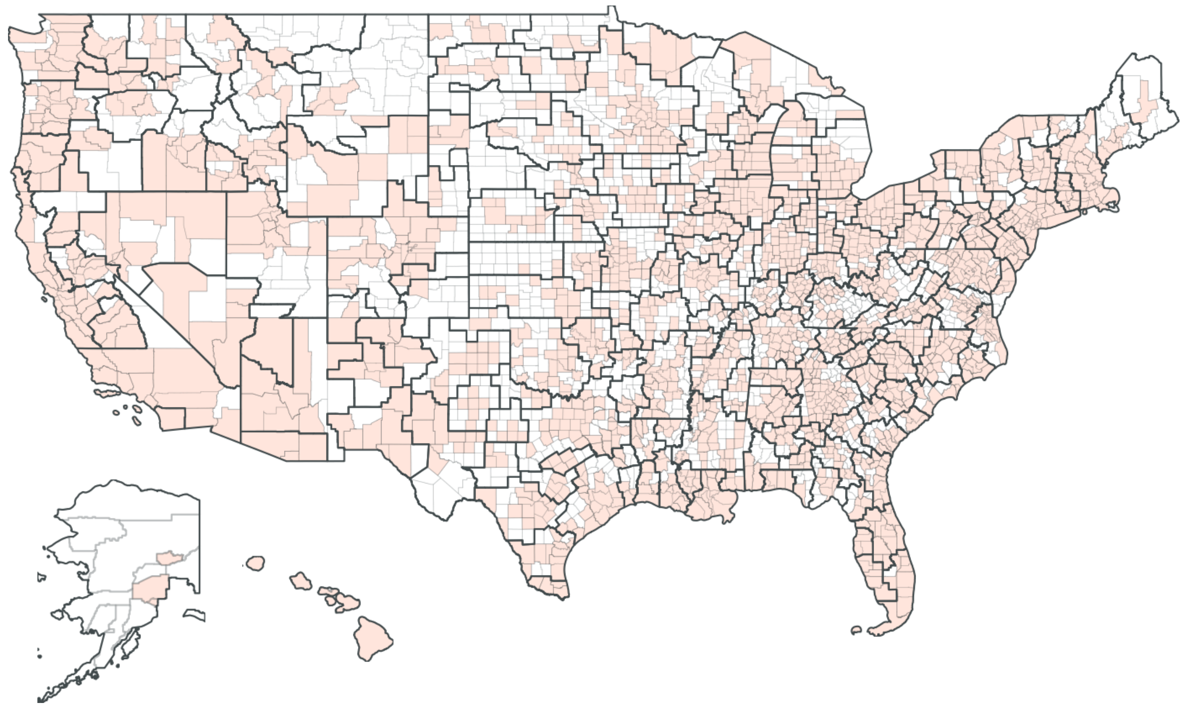


Figure B.2: Bureau of Economic Analysis' "Economic Areas" Map

Notes: This figure presents the BEA's economic areas map and compares them the MSA's, painted in pink. One can see that the size of the economic areas changes across locations, and it may coincide with the corresponding MSA, but it can also be larger.

Table B.4: Effect Size and Within vs. Between Firm Move

	Within Firm		Across Firms	
	(1) Annual Number of Patents	(2) Annual Number of Adjusted Citations	(3) Annual Number of Patents	(4) Annual Number of Adjusted Citations
$PostMove^{\text{Real}}$	0.076^{***}	0.085^{***}	0.061^{**}	0.055^{**}
	(0.021)	(0.021)	(0.029)	(0.026)
Control Post Mean	0.501	0.336	0.498	0.348
Percentage Change	+15.08%	+25.23%	+12.18%	+15.78%
P-Value H_0: Diff. $= 0$			0.67	0.36
Observations	223955	223955	331860	331860
Individual FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Experience FE	Yes	Yes	Yes	Yes

Notes: The information presented in this table presents the results from regression equation (2.5) on two different samples. Columns (1) and (2) correspond to cases where the mover moves within the same firm, and columns (3) and (4) cover the cases where the mover moves across firms. The unit of analysis in these regressions is still inventor-year. The dependent variable in columns (1) and (3) is the number of patents per year and in columns (2) and (4) is the number of adjusted citations per year, as defined in Section 2.3. Standard errors are clustered at the mover level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table B.5: Effect Size and the Characteristics of Across Firm Move

	Corp.-Corp.		Non-Corp.-Corp.	
	(1) Annual Number of Patents	(2) Annual Number of Adjusted Citations	(3) Annual Number of Patents	(4) Annual Number of Adjusted Citations
$PostMove^{\text{Real}}$	0.071^{***}	0.047^{*}	-0.015	-0.015
	(0.027)	(0.027)	(0.088)	(0.066)
Control Post Mean	0.441	0.326	0.517	0.345
Percentage Change	+16.15%	+14.29%	-2.95%	-4.42%
Observations	315511	315511	52956	52956
Individual FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Experience FE	Yes	Yes	Yes	Yes

Notes: The information presented in this table presents the results from regression equation (2.5) on two different samples. Columns (1) and (2) correspond to cases where the mover and the left behind are of the same sex, and columns (3) and (4) cover the cases where the mover and the left behind are of opposite sexes. The unit of analysis in these regressions is still inventor-year. The dependent variable in columns (1) and (3) is the number of patents per year and in columns (2) and (4) is the number of adjusted citations per year, as defined in Section 2.3. Standard errors are clustered at the mover level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table B.6: Heterogeneity based on Good vs. Bad Moves

	Bad Move		Good Move	
	(1) Annual Number of Patents	(2) Annual Number of Adjusted Citations	(3) Annual Number of Patents	(4) Annual Number of Adjusted Citations
$PostMove^{\text{Real}}$	-0.008	0.019	0.178***	0.128***
	(0.017)	(0.015)	(0.044)	(0.042)
Control Post Mean	0.442	0.289	0.639	0.464
Percentage Change	+-1.82%	+6.59%	+27.88%	+27.49%
P-Value H_0: Diff. $= 0$			00	0.01
Observations	149064	149064	406751	406751
Individual FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Experience FE	Yes	Yes	Yes	Yes

Notes: The information presented in this table presents the results from regression equation (2.5) applied to two distinct subsets of data. Specifically, Columns (1) and (2) pertain to scenarios where the mover relocates to a location where the density of inventors who patent in the same CPC class is higher, and columns (3) and (4) delve into the opposite case. The unit of analysis remains inventor-year in these regression analyses. Columns (1) and (3) encompass the number of patents per year as the dependent variable, while columns (2) and (4) utilize the number of adjusted citations per year, as elaborated in Section 2.3. Standard errors are clustered at the mover level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table B.7: The Effect of a Relocation and Race Differences

	Same Race		Different Races	
	(1) Annual Number of Patents	(2) Annual Number of Adjusted Citations	(3) Annual Number of Patents	(4) Annual Number of Adjusted Citations
$PostMove^{\text{Real}}$	0.049***	0.048***	0.038	0.057
	(0.018)	(0.017)	(0.043)	(0.039)
Control Post Mean	0.499	0.348	0.503	0.324
Percentage Change	+9.80%	+13.9%	+7.49%	+17.65%
P-Value H_0: Diff. $= 0$			0.80	0.833
Observations	366726	366726	189089	189089
Individual FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Experience FE	Yes	Yes	Yes	Yes

Notes: The information presented in this table presents the results from regression equation (2.5) applied to two distinct subsets of data. Columns (1) and (2) address situations where the mover and the left behind inventor have the same race. Columns (3) and (4) report the results when the mover and the left behind inventors have different races. The unit of analysis remains inventor-year in these regression analyses. Columns (1) and (3) encompass the number of patents per year as the dependent variable, while columns (2) and (4) utilize the number of adjusted citations per year, as elaborated in Section 2.3. Standard errors are clustered at the mover level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table B.8: The Effect of Relocation on Share of Collaborators to the Destination Patent Based Measure

	(1) Annual Percentage of Collaborators in Destination	(2) Annual Percentage of Collaborators in Destination
	All	Only When Patenting
$PostMove^{\text{Real}}$	1.172***	5.291***
	(0.096)	(0.222)
Control Post Mean	1.523	7.216
Percentage Change	+76.94%	+73.32%
Observations	555815	142296
Individual FE	Yes	Yes
Year FE	Yes	Yes
Experience FE	Yes	Yes

Notes: The information presented in this table presents the results from regression equation (2.5). In column (1), I consider all the observations in the panel, whether or not the inventor patented in that year, while column (2) covers only the years when the inventor patents. The unit of analysis in these regressions is still inventor-year. The dependent variable is the share of collaborators in the destination location. Standard errors are clustered at the mover level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table B.9: The Effect of a Relocation and Gender Differences

(a) Panel A: Same Sex

	Male-Male		Female-Female	
	(1) Annual Number of Patents	(2) Annual Number of Adjusted Citations	(3) Annual Number of Patents	(4) Annual Number of Adjusted Citations
$PostMove^{\text{Real}}$	0.042**	0.043**	-0.021	-0.054
	(0.018)	(0.018)	(0.056)	(0.068)
Control Post Mean	0.517	0.362	0.424	0.289
Percentage Change	+8.07%	+12%	-4.93%	-18.6%
P-Value H_0: Diff. $= 0$			0.28	0.16
Observations	343092	343092	14969	14969
Individual FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Experience FE	Yes	Yes	Yes	Yes

(b) Panel B: Opposite Sex

	Male-Female		Female-Male	
	(1) Annual Number of Patents	(2) Annual Number of Adjusted Citations	(3) Annual Number of Patents	(4) Annual Number of Adjusted Citations
$PostMove^{\text{Real}}$	0.039	0.064	-0.052	-0.021
	(0.038)	(0.039)	(0.060)	(0.043)
Control Post Mean	0.395	0.252	0.583	0.349
Percentage Change	+9.94%	+25.5%	-8.87%	-6.09%
P-Value H_0: Diff. $= 0$	0.95	0.62	0.13	0.16
Observations	54593	54593	54696	54696
Individual FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Experience FE	Yes	Yes	Yes	Yes

Notes: The information presented in this table presents the results from regression equation (2.5) applied to four distinct subsets of data. Each subset corresponds to a different combination of the sexes of both the mover and the left behind inventor. Panel A addresses situations where the mover and the left behind inventor share the same sex. Specifically, Columns (1) and (2) pertain to scenarios where both the mover and the left behind are males, and columns (3) and (4) delve into cases where both the mover and the left behind are females. On the other hand, Panel B delves into scenarios where the sexes of the mover and the left behind inventor differ. Columns (1) and (2) within this panel represent cases where the mover is male and the left behind is female, while columns (3) and (4) encapsulate the reverse situation—where the mover is female and the left behind is male. The unit of analysis remains inventor-year in these regression analyses. Columns (1) and (3) encompass the number of patents per year as the dependent variable, while columns (2) and (4) utilize the number of adjusted citations per year, as elaborated in Section 2.3. Standard errors are clustered at the mover level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table B.10: The Effect of Relocation and Continued Collaboration

	Continued Collaboration with Mover		No Collaboration with Mover	
	(1) Annual Number of Patents	(2) Annual Number of Adjusted Citations	(3) Annual Number of Patents	(4) Annual Number of Adjusted Citations
$PostMove^{\text{Real}}$	0.055^{**}	0.041^{**}	0.003	0.059
	(0.023)	(0.020)	(0.038)	(0.038)
Control Post Mean	0.34	0.22	0.95	0.66
Percentage Change	+16.01%	+18.15%	+0.3%	+8.87%
P-Value H_0: Diff. $= 0$	0.243	0.677		
Observations	426153	426153	129662	129662
Individual FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Experience FE	Yes	Yes	Yes	Yes

Notes: The information presented in this table presents the results from regression equation (2.5) applied to two distinct subsets of data. The first subset corresponds to left behind inventors who continue to collaborate with their respective mover, while the second one corresponds to cases where they stop collaborating with one another after the move. Columns (1) and (2) pertain to scenarios where the left behind and the mover continue to collaborate, and columns (3) and (4) delve into cases where they do not. The outcome variable in columns (1) and (3) is the number of patents per year, while columns (2) and (4) utilize the number of adjusted citations per year, as elaborated in Section 2.3. Standard errors are clustered at the mover level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.