Università Commerciale "Luigi Bocconi"
PhD School

PhD program in Economics and Finance

Cycle: 36th

Disciplinary Field (code): SECS-P/06

Essays in Human Capital and Innovation

Advisor: Nicolas Serrano-Velarde

Co-Advisor: Carlo Schwarz

PhD Thesis by

Enrico Stivella

ID number: 3142299

Year: 2026

Introduction

The diffusion of knowledge and information is a fundamental driver of technological progress and economic growth. In this dissertation, I investigate how such diffusion occurs and how it is shaped by the mobility of workers. Across three distinct but thematically linked studies, I explore the mechanisms through which human capital—embodied in skilled individuals—transmits knowledge within and across organizations, scientific domains and institutional frameworks. Each chapter focuses on a specific context where worker transitions play a critical role in facilitating or redirecting knowledge flows: the mobility of bankers, the careers of nuclear scientists, and the impact of public procurement on innovative activity.

Taken together, these studies aim to show how portable is knowledge and how it can be effectively harnessed to foster innovation and economic development. As bank managers are hired by competing banks, they are able to transfer their human capital (i.e. soft information about clients) to their new employers, thereby improving the overall allocation of credit in the economy. Viceversa, scientists who cannot do research in their original field due to funding cuts observe a significant decrease in their productivity, as their knowledge is not easily transferable to other domains. Finally, public procurement can be a powerful tool to steer the direction of innovation by creating demand for specific technologies, but it mainly affects the ability of firms to obtain further contracts, as opposed to generating new knowledge. Across the chapters, I employ a range of empirical methods, based on detailed administrative data (for the case of bankers) or large publicly available datasets (for the cases of scientists and procurement).

Worker Mobility and Capital Allocation: The Case of Bankers

Efficient capital allocation is crucial for economic growth (Rajan and Zingales, 1998). Banks play a central role in this process by providing loans (Stiglitz and Weiss, 1981) and by producing information about borrowers (Diamond, 1984). A well-established literature has distinguished between *hard information*—data that is verifiable and recordable—and *soft information*—judgmental insights derived from personal relationships (Stein, 2002). In my first chapter, I explore how the movement of bank employees affects the transmission of soft information and the resulting allocation of credit.

Using detailed administrative data on Italian bank managers and the firms they serve, I study what happens when a banker moves to a new institution. I show that the clients whom a banker has historically worked with are more likely to receive loans from the banker's new bank. These finding suggests that soft information is not simply lost when a banker moves; instead, it travels with the individual and can be leveraged in the new organization.

To identify this effect, I construct a dataset that links bank managers to their clients over time, tracking how loans evolve following the banker's relocation. The empirical strategy exploits variation in the timing and geography of banker moves, allowing for a difference-in-differences framework. This evidence underscores the idea that personal relationships are a key channel for knowledge diffusion in financial markets.

Scientific Human Capital and the Geography of Innovation

There is no economic growth without innovation (Romer, 1990; Aghion and Howitt, 1992; Akcigit, Pearce, and Prato, 2020). And there is no innovation without investments in scientific human capital (Barro, 2001; Prato, 2022). However, economists have not fully understood how human capital helps producing new technologies. Capital goods can only perform the limited range of tasks they are designed for. One cannot use an ax to plow or a hoe to cut wood. On the other hand, human capital is generally assumed to be more flexible, although labor economists have debated about its depreciation (Deming and Noray, 2020), specificity with respect to the firm (Becker, 1962; Lazear, 2009), to the team (Chen, 2021) or to the technology (Marx, Strumsky, and Fleming, 2009).

In the second chapter of this dissertation, I study how flexible scientific human capital can be. It is flexible when researchers are able to take some scientific principles and apply them out of the fields in which they were discovered. This is the story of many breakthrough innovations, such as penicilline or the X-rays, which are the successful - but unintended - results of projects started with different aims (Nelson, 1959).

I focus on the case of Italian nuclear scientists, whose field of study has been entirely de-funded after the Chernobyl disaster in 1986.

This setting offers a rare opportunity to observe the forced movement (across fields of study) of highly specialized human capital. By tracing the publication records and patenting activity of affected scientists, I quantify how their relocation influenced not only their own productivity.

Public Procurement and the Direction of Technical Change

In this third chapter, I turn to the role of public policy in shaping innovation outcomes. Most of the literature on innovation policy has focused on the spillovers of big push investments, such as the Manhattan project (Gross and Sampat, 2020), the Apollo program (Kantor and Whalley, 2022), or military spending (Moretti, Steinwender, and Van Reenen, 2019). Others have shown the efficacy of smaller and more targeted tools, such as research grants (Howell, 2017; Santoleri et al., 2022; Myers and Lanahan, 2022) or public procurement (Belenzon and Cioaca, 2022; Rassenfosse, Jaffe, and Raiteri, 2019). In this chapter I study a combination of grants and procurement, in a context in which funding priorities are set by the Department of Defense (DoD).

I study this question in the largest context in which the U.S. government can combine grants and procurement: the Small Business Innovation Research (SBIR) program at the Department of Defense (DoD). Combining data from SBIR grants, U.S. government procurement contracts (USAspending.gov) and U.S. patent records (Patstat), I analyze how public procurement affects the direction of technical change in small and medium-sized enterprises (SMEs). I show that when SBIR grant winners are awarded a procurement contract, they become more likely to obtain further (non-SBIR related) contracts, but not to file new patents. This suggests that public procurement can be a powerful tool to help firms grow and secure additional funding, but it does not necessarily lead to new technological breakthroughs.

Conclusion

Each of the three chapters in this dissertation examines a different facet of how knowledge and information diffuse in the economy, with a particular emphasis on the role of human capital and worker mobility. The first chapter highlights how soft information moves with bankers and affects credit allocation; the second traces how displaced scientists can struggle to apply their specialized knowledge in new fields, revealing the limits of human capital portability; and the third investigates how public procurement can influence firm behavior and innovation outcomes, albeit with a focus on contract acquisition rather than new knowledge creation.

These findings contribute to a deeper understanding of the mechanisms that underpin knowledge diffusion and innovation. Together, these studies underscore the importance of individuals—not just institutions or incentives—in the innovation process. Knowledge resides in people, and it travels with them. However, its transfer is not always straightforward.

From a policy perspective, these findings suggest that fostering efficient knowledge diffusion requires attention to the frictions and enablers of worker mobility. Organizational structures, labor market institutions, and the specificities of human capital all play crucial roles in determining how effectively knowledge can be shared and utilized.

Contents

List of Figures

List of Tables

Chapter 1

Loans and human capital: Manager mobility in the financial sector and credit allocation.

This chapter is co-authored with Angelo D'Andrea (Bank of Italy).

Abstract

Information is a core input in financial intermediation, yet little is known about how it is transmitted through labor mobility. In this paper - for the first time - we are able to link administrative data on loan contracts and employment histories to track the movements of bank managers and their portfolios of client firms across financial institutions. We show that when a manager switches employers, firms in their prior portfolio are five times more likely to initiate a lending relationship with the new bank. This effect reflects both a higher propensity of firms to apply for credit and a greater likelihood of loan approval. To isolate the causal role of manager mobility, we exploit variation in the timing of job switches and leverage exogenous shocks to mobility induced by branch closures. Our findings underscore the importance of individual-level human capital in shaping credit market outcomes and highlight a novel channel for information transmission in the financial sector.

1.1 Introduction

Efficient capital allocation is crucial for economic growth (R. Rajan and Zingales, 1998). Banks play a central role in this process by providing loans (Stiglitz and Weiss, 1981) and generating borrower information (Diamond, 1984). Economic literature has categorized *hard information*, which is systematically recorded in databases, and *soft information*, which arises from personal relationships between bank managers and borrowers (Stein, 2002).

While the significance of hard and soft information have been extensively studied, less is known about the financial workforce's role in this process—specifically, how bank managers generate, utilize, and disseminate soft information to allocate credit. This gap in the literature stems from a lack of microdata that directly capture the relationships between bank managers and borrowers. Existing studies have either focused on soft information within a single institution (Hertzberg, Liberti, and Paravisini, 2010), used geographical proximity as a proxy (Nguyen, 2019), or examined cultural homophily (Fisman, Paravisini, and Vig, 2017).

In this paper, we address this gap by combining for the first time data on financial workforce mobility and credit allocation. We examine firms' credit relationships in Italy from 2009 to 2018 and track bank managers' career movements. Using manager transitions as a source of variation in a generalized Difference-in-Differences framework, we find that a manager's new bank is significantly more likely to establish a credit relationship with a former client of the manager.

We then explore the underlying mechanisms using loan application data and show that managers transfer valuable information in two key ways. First, managers direct firms' search for credit: firms in a manager's portfolio are more likely to apply for a loan at the manager's new bank. Second, managers influence their bank's screening: these firms are also more likely to have their loan applications approved.

Italy provides an ideal setting for our analysis. First, Italian credit registry data are highly granular and detailed (Malgieri and Citino, 2025; De Marco, Sauvagnat, and Sette, 2023; Bolton et al., 2016; Rodano, Serrano-Velarde, and Tarantino, 2018; Gobbi and Sette, 2014), offering information on loan amounts, interest rates, and banks' loan inquiries about firms. Crucially, it includes the identifier of the bank-municipality pair issuing the loan, a key element in our analysis. Italian social security data cover the entire financial sector workforce, allowing us to identify bank managers and their characteristics. Second, the institutional context is particularly suitable. Italian firms maintain multiple banking relationships (Barone, Schivardi, and Sette, 2024; Kosekova et al., 2024), which enables us to disentangle credit supply from demand by augmenting our regressions with high-dimensional fixed effects, following Khwaja and Mian (2008). Moreover, from the 1990s through the 2010s, Italy's banking sector underwent significant consolidation, with many local banks merging into a few national champions (Focarelli and Panetta, 2003), and thus reorganizing (and sometimes laying off) their workforce. At the same time, labor laws facilitated mobility within the sector, leading to a large-scale reallocation of workers during our analysis period.

Our analysis employs a generalized difference-in-differences empirical strategy, leveraging two dimensions of variation: the bank branch with which a firm has credit and whether the bank manager responsible for that credit has moved to a different bank. Under the parallel trends assumption, this variation enables us to estimate the effect of a manager's move on the likelihood of a firm establishing a credit relationship with the manager's new bank. Our approach allows us to rule out several confounding factors. First, we control for time-invariant characteristics of the firm-bank relationship that influence their probability of matching, such as industry specialization (Paravisini, Rappoport, and Schnabl, 2023) or geographical proximity (Nguyen, 2019), through firm-bank fixed effects. Second, we account for credit supply shocks that affect a bank's overall lending propensity with bank-time fixed effects. Third, we address credit demand shocks that influence a firm's likelihood of applying for a loan by using firm-time fixed effects (as in Khwaja and Mian, 2008). Finally, we control for the possibility that a manager's move is endogenous to a firm's credit decision by focusing on cases where the move is triggered by a branch closure —an event exogenous to the firm's choice to shift credit relationships.

The main finding of the paper is that bank managers' information about their clients is portable. Five years after a manager moves, the probability of a firm establishing a credit relationship with the manager's new bank rises from a baseline of 1.2% to 7%. We provide two benchmarks for this result. First, we compare transitions between banks to transitions within the same bank group (i.e., loan officer rotations, as in Hertzberg, Liberti, and Paravisini, 2010). We find that managers who switch to a different bank group bring more than twice as many clients with them compared to those who move within the same group. Second, we contrast our estimate with client portability in the context of sales managers (Patault and Lenoir, 2024), showing that the effect of managerial information is particularly strong in banking—approximately an order of magnitude greater than in sales.

To understand the mechanisms driving this result, we analyze loan applications. We find that managers directly affect firms' search: after a move, the probability of a portfolio firm applying for a loan at the manager's new bank increases by 8 percentage points, compared to a baseline of 2%. Additionally, managers influence banks' screening: conditional on applying, the likelihood of a portfolio firm receiving a loan rises from 35% to 42%.

We then examine how our results vary across three key dimensions: firm, bank, and manager characteristics. First, we find that the impact of a manager's move is stronger for smaller firms and those with fewer credit relationships. Second, we show that managers' information is particularly valuable when they move from a large bank group to a smaller one, as larger banks tend to rely on more standardized credit allocation processes. Finally, we find that the effect is stronger for managers with greater experience and those from branches with fewer managers, as they have had more time and interactions to build relationships with their clients.

Taken together, our findings provide direct evidence that bank managers generate and utilize soft information to allocate credit. The academic debate has often centered on the effects of de-branching (Amberg and Becker, 2024; Nguyen, 2019) or banking crises (Chodorow-Reich, 2014) on credit allocation. Meanwhile, policy discussions—particularly following the 2008 financial crisis—have focused on strengthening bank stability while ensuring credit access to the real economy (Philippon, 2015). Diamond (2001) argued that banks with personal information about borrowers should be protected through government intervention during crises. Our findings suggest an additional channel through which bank stability can be influenced: the reallocation of human capital within the banking sector.

Our results are robust to several tests. We show that our results are valid even if we aggregate branches at the provincial level. This operation serves two purposes: first, it allows us to reduce the dimensionality of our dataset, which is particularly important when we consider the large number of branches in our sample. Second, and more important, it allows us to rule out the fact that the moving firm had a pre-existing relationship with the new bank, but in a different municipality. Finally, we test the robustness of our specification by allowing for cohort-specific treatment effects, following Sun and Abraham (2021). Results show that our main findings are robust to this alternative specification.

Our paper contributes to several strands of literature. First, we add to the extensive research on bank branching and relationship lending, which — since the seminal work of Stiglitz and Weiss (1981) — has examined how soft information in banking relationships influences credit allocation (Amberg and Becker, 2024; Babina, Buchak, and Gornall, 2022; Fisman, Paravisini, and Vig, 2017) and loan pricing (Beraldi, 2025). Scholars have used mainly two approaches to measure the effects of soft information on credit: geographical proximity (Bonfim, Nogueira, and Ongena, 2021; Degryse and Ongena, 2005; Petersen and R. G. Rajan, 2002) and cultural homophily (Fisman, Paravisini, and Vig, 2017; Cornell and Welch, 1996). Nguyen, 2019 and Duquerroy et al., 2022 exploit branch closures to assess the impact of losing a local branch on credit allocation and sectoral specialization. Fisman, Paravisini, and Vig (2017) find that in India, cultural homophily between loan officers and borrowers affects both the likelihood of loan approval and loan success. Hertzberg, Liberti, and Paravisini (2010) use internal bank data to show how loan officer rotations influence screening and monitoring incentives, ultimately shaping credit allocation. We contribute to this literature in two key ways. We adopt methods from labor and trade economics (Patault and Lenoir, 2024; Kramarz and Skans, 2014) to construct a portfolio of client firms for each bank manager, allowing us to directly measure individual relationships. In addition, the richness of our dataset enables us to track managers and their clients across banks — not just within the same banking group — allowing us to quantify the portability of managerial information across the financial sector. Additionally, we show that personal relationships, rather than physical distance, are the primary determinant of a firm's bank selection, as clients follow managers up to a long distance.

A second strand of literature explores the intersection of labor market dynamics and finance. Acabbi, Panetti, and Sforza (2024) and Jasova et al. (2021) combine workforce administrative data with credit registry data to examine how labor responds to credit and monetary shocks. We contribute to this literature by considering the opposite direction of causality—how labor reallocation within the financial sector influences credit allocation. To do so, we build on prior research examining occupation and wages in finance. Philippon and Reshef (2012) initiated a line of work documenting key features of the financial labor force, including the drivers of wage premia (Bell and Van Reenen, 2014) and concerns about "brain drain" following financial sector deregulation (Boustanifar, Grant, and Reshef, 2018; Böhm, Metzger, and Strömberg, 2023). Gao, Wang, and Wu (2025) develop a career model for bank employees, arguing that much of their human capital is not portable, leading to penalties when they switch banks. Our contribution lies in linking these stylized facts about the financial workforce to core banking functions such as credit allocation. We show that, contrary to some prior findings, bank managers' human capital is portable because it is rooted in personal relationships with clients.

Personnel economics has extensively documented the impact of managers on firm performance (Bloom et al., 2013; Bandiera et al., 2020; Lazear, Shaw, and Stanton, 2015). Building on Minni (2025), who shows that manager rotations can have long-term effects on subordinates' careers, we show that bank managers similarly contribute to the long-term performance of their client firms. The study that has most closely examined the portability of managerial relationships - although in the context of French exporting firms - is Patault and Lenoir (2024), who defines "customer capital" for sales managers, showing that they can retain clients even after switching employers. We extend this insight to the banking sector, revealing that the portability of managerial relationships is more pronounced in finance than in sales. Moreover, we show that this phenomenon not only benefits the receiving bank but also enhances outcomes for the client firms.

The remainder of the paper is organized as follows. In Section 1.2, we describe our data sources and sample construction. Section 1.3 outlines our empirical strategy. Section 3.3 presents our main results. In Section 1.5, we analyze information requests to uncover the mechanisms driving our findings. Section 1.6 examines heterogeneity across different dimensions. Section 1.7 provides robustness checks. Finally, Section 1.8 concludes.

1.2 Data and sample construction

In the following paragraphs we first present the datasets we use, then we describe how we perform our matching procedure, and finally we describe the resulting dataset that we use in our empirical analysis.

1.2.1 Main datasets

We combine three different datasets, which are provided by the Bank of Italy and refer to the period 2009-2018. The first one is the credit registry data, which contains all loans granted by Italian banks to Italian firms. The second one is the social security data, which contains all the workforce of the Italian financial sector. The third one is the Cerved data, with balance sheet information on all incorporated Italian firms (around 500 thousand unique firms).

Credit information The credit registry data is a panel of all outstanding loans greater than 30 thousand euros granted by Italian banks to Italian firms in the years 2009-2018, provided by the Bank of Italy. Each loan is identified by the receiving firm, the amount, the year and the granting bank branch.[1] From two separate sections of the credit registry we also retrieve data on interest rates and loan applications. Interest rates are coded at the bank group - firm level, so that we can tell the aggregate average interest rate that a firm has paid to a bank group in a given year. For loan applications, we use a dataset containing the requests for information (”Richieste di prima informazione” in Italian) that banks make to the credit registry in order to screen potential borrowers or monitor existing ones (Branzoli and Fringuellotti, 2020; D'Andrea, Pelosi, and Sette, 2023). We consider these requests for information as an intermediate step between the first contact between a firm and a bank, and the actual granting of a loan. The availability of such data is unique in the context of credit registry data, and allows us to investigate the mechanisms behind our main results. As for the interest rates, requests for information are coded at the bank group - firm level, so that we can tell if a bank group has requested information about a firm in a given year.

We consolidate the bank group identifier at the end of the sample, in order to avoid considering any mechanical credit relocation due to M&As.[2] The dataset contains 8 million observations (one for each firm-branch-year triplet), 440 thousand unique firms, and 31 thousand unique branches.

Workforce information We access a subset of Italian Social Security data, which contains the universe of workers in the Italian financial sector for the years 2009-2018. The dataset is provided to the Bank of Italy by INPS (Istituto Nazionale della Previdenza Sociale, the Italian National Institute of Social Security). We observe for

[1]In this paper we identify bank branches in the credit registry by the unique combination of municipality and bank group codes.

[2]Consider indeed the case in which Bank group A acquires Bank group B in 2015, and the two banks have a branch each in the same municipality. Right after the M&A operations the two branches would be considered as the same branch in the credit registry. Therefore we would observe credit (and also people, as we will see later) moving mechanically from one branch to the other. As we are interested only in firm's decision to move credit, we use a version of the credit registry that sets the bank group at the end of the period (2018) as the identifier of the branch for the whole period. In the previous case we then would have a unique branch of Bank group A in the municipality for the years 2009-2018.

each worker, at yearly level, her bank group, the municipality where she works,her type of contract (managerial or not), her wage and a set of demographic characteristics (such as age or gender). As for the credit registry data, we identify a branch by the unique combination of municipality and bank group. The dataset comprises around 3.5 million observations, resulting in 350 thousand workers per year.

Firm characteristics The third dataset, which contains firm-level charactersistics on all incorporated Italian firms, is administered by Cerved Group and licensed to the Bank of Italy. The unique number of firms in the dataset is around 500 thousand. The dataset provides balance sheets and income statements, that we use to obtain standard firm-level variables, such as size, industry, assets, investments, credit score, and profits. Sole proprietorships, small household producers, or unincorporated partnerships are not covered. A firm is identified by a unique fiscal code, which we then match to the credit registry through a crosswalk provided by the Bank of Italy.

1.2.2 Sample construction

We construct our sample by adding to the credit registry information on workforce transitions. Therefore our final dataset will be built with the same structure of the credit registry (firm-branch-year triplets), and an additional column that keeps track of manager moves, matching firms that the managers have known in the past to the branches where they move. Our matching procedure relies on the fact that we can identify where a manager is employed in the social security data, and where a firm has credit in the credit registry data. We measure knowledge flow through transitions, of which Figure 1.1 provides a schematic representation. Consider two banks, that we just label as old and new bank (with respect to the manager's move). At time $t = 0$, the manager is in her old bank, where she gives credit to firms A and B. At the same time, the new bank she will move to has credit relationships with firms C and D. At time $t = 1$, the manager moves to the new bank, and the credit relationships change. Both firm A and B are part of the manager's portfolio as she moves. Firm A actually follows the manager to the new bank, while firm B does not. In order to measure this phenomenon in the data, we need to define who the branch managers are, which firms are in their portfolio, and what a move is.

Branch managers In both the credit registry and the social security data, a branch is identified by the unique combination of municipality and bank group codes. In order

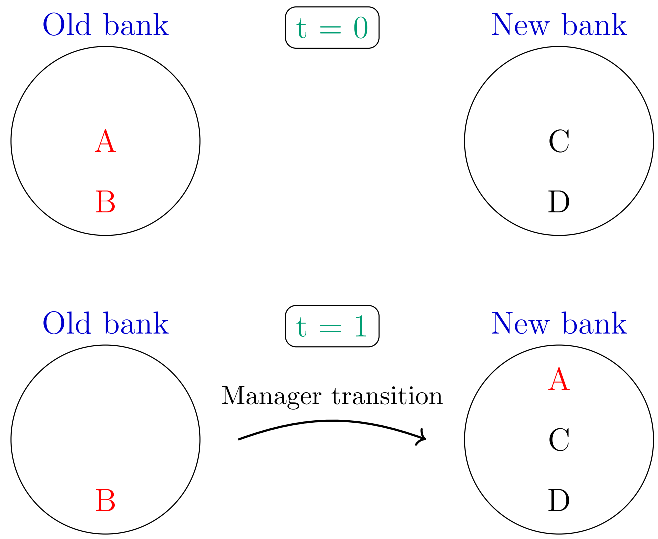


Figure 1.1. Representation scheme of the manager transition.

Notes: Figure represents the impact of a manager transition from the old bank to the new bank. Firms A and B were part of its portfolio at the old bank, A follows her.

to define who the branch managers are, we select only the workers that are labelled as top or middle managers in the social security data. At the same time, we need to identify within managerial workers, those who actually are in charge of their branch and make decisions about lending. As a consequence, we focus only on a particular subset of branches. We select only branches where the number of managers is less than 5, which correspond to more than 80% of the branches in our sample. The rationale behind this choice is threefold: first, we want to be sufficiently sure that the manager has had direct contact with the firms which had loans with the branch (i.e. we want to exclude branches where the are too many managers, and we cannot reliably map the manager to the firms). Second, we want to exclude headquarters, i.e. branches where there is abundance of managerial personnel, but those managers are involved also in other tasks and are not in direct contact with the firms. Third, our goal is to match as closely as possible the dataset definitions with the physical presence of branches. Indeed, in reality a bank may have multiple branches in a single municipality, but they all appear as one in the credit registry and in the social security data. Filtering for the number of managers is a way to reasonably exclude multiple branches in the same municipality. Our threshold choice allows us to still consider a large number of branches (14 thousand), but be more precise in our matching. Figure 1.10 in the Appendix shows the distribution of the number of managers per branch (the average number of managers is 2.3).

Following the same rationale, we exclude the main municipality per each administrative district, or province.[3] These restricting assumptions, that leave out major cities in Italy, such as Rome or Milan, reduce the size of the credit registry to around 50% of the original size, but they are crucial to make sure that we are actually capturing the credit allocation decisions of the branch managers. However, as it is shown in the Appendix, the distribution of firms, branches and loan types is not affected by the filtering. In our robustness exercises, we repeat our main exercise using the full sample (but still considering only small branches as starting points for moves and portfolio construction), and the results are robust.

Portfolio Once we have established who the branch managers are, and what are the relevant branches, we need to pin down the portfolio of clients each branch manager has. We label as portfolio firms all the firms that have had active loans with the branch where the manager works in both the two years before the manager moves. We do it to accomplish two results: first, we want to make sure that the manager was in direct contact with those firms up until she moved (i.e. we want to exclude firms that left the branch before the manager moved). Second, we want to consider as portfolio firms only the ones who have had a sufficiently long relationship with the branch, so that the manager could have actually acquired soft information on them. The resulting average portfolio size is around 20 firms per manager. Notice that we are not able to tell which firm was interacting with which manager within the branch. Therefore, our choice of restricting to small branches is crucial, as it increases the likelihood that the manager was actually in contact with the firms that we consider as her portfolio. Consider also that the possible noise in our portfolio construction would bias our results towards zero, as we introduce in a manager's portfolio firms that she may not have known. Any measure of the portability of credit relationships that we find is therefore conservative. We test the robustness of our results by focusing on subset of branches that had a smaller number of managers (only one being the most restrictive case), and our results are - as expected - stronger.

[3] Italy is divided in 110 provinces. Each province has a main city, where all administrative offices, such as Courts or Chambers of commerce, are concentrated. Panetta, Schivardi, and Shum (2009), Crawford, Pavanini, and Schivardi (2018) and Barone, Schivardi, and Sette (2024) use the province as the main dimension of the credit market. Our data confirm that 60% of firms have credit in a single province (see Figure 1.11 in the Appendix). Therefore it is reasonable to assume that banks locate their headquarters in capoluoghi di provincia, where the managerial personnel is mostly involved in non-lending related tasks.

Moves The third part of our matching procedure is how to define a move. In social security data, we observe where the branch manager works in each year. We know both her municipality and the bank group she works for. Our baseline definition of a move is a change in the bank group of the manager from year t to year $t+1$. However, there is also another type of move, which is a change in the municipality of the manager within the same bank group. The two types are inherently different. In the baseline case, which accounts for 4.1% of the yearly moves in our data, the manager transfers information from one bank to another, and create a new credit access for the client firms (Gao, Wang, and Wu, 2025). The second case, which accounts for 95.9% of the yearly moves, corresponds to a more traditional definition of loan officer rotation, or may be generally due to personnel management within the bank (Minni, 2025; Gao, Wang, and Yu, 2024). Even in this latter case there is room for information transmission, as personal information is specific to the branch manager, and the relationship lending literature has shown that even within the same bank, loan officers have incentives to hoard information about borrowers (Hertzberg, Liberti, and Paravisini, 2010). Our baseline specification will consider only moves between bank groups, as they constitute a unique feature of our dataset. We will use the second type of moves in our heterogeneity exercises, to benchmark the results with the more traditional definition of loan officer rotation.

1.2.3 Resulting dataset

As a result of the refinements discussed in the previous pages, we identify in the credit registry data 14 thousand small branches. By filtering the credit registry with respect to those branches, we obtain in our baseline specification a dataset of 4 million observations (one for each firm-branch-year triplet) and 160 thousand unique firms. Notice that in this filtered version of the credit registry (which will be used for our main specification) we include only firms that have been always active, i.e. present in the credit registry in all years 2009-2018. We do it to obtain a balanced panel sample in the regressions, and avoid any bias that may derive from the fact that some firms may have left or arrived in the credit registry. At the same time, by filtering the social security data with respect to the branch managers, we obtain a dataset of 20 thousand branch managers per year, and around 5 thousand moves per year. Within those moves, 95.9% are within the same bank group, and 4.1% are between bank groups. All the 160 thousand firms in our filtered sample are present in Cerved data.

Once we have defined our sample filterings, we need to do an extra balancing that increases significantly the dimensionality of our data. The credit registry is a panel of firm-branch-year triplets, but contains only the pairs that have realized (i.e. the firm has a loan with the branch). In order to have a proper representation of the credit market, we need to consider all the firm-branch potential pairs in the years 2009-2018. From the perspective of the firm, we need to consider all the branches it could have asked credit from, consider the ones that have realized, and measure if the presence of a manager that knew the firm has affected this matching realization. Similarly, from the perspective of the branch, we need to consider all the possible firms it could have given credit to. Panetta, Schivardi, and Shum (2009), Crawford, Pavanini, and Schivardi (2018) and Barone, Schivardi, and Sette (2024) define credit market at the provincial level,[4] so that we consider only firm-branch potential pairs that are in the same province. Then our final dataset is the aggregation of the 110 provincial markets, where a market is defined as the set of all potential firm-branch pairs in a province. The dataset has a size of 231 million observations, which makes computationally expensive the estimation of our model. In our main regressions we sample around 10% of the data, and we repeat our exercises to validate our results.

1.3 Empirical Strategy

Our main goal is to quantify how portable are the relationships that bank managers build with their clients. We measure it by computing the probability of a firm establishing a credit relationship with the new bank of the manager, as opposed to any other bank in the same province. As explained in section 1.2, in order to measure this probability, we have to create matching pairs of firms and banks. With respect to banks, we can identify the branch through the unique combination of municipality and bank group. Then our unit of analysis is a potential match between a branch and a firm. It is treated when the branch hires a manager that has given credit to the firm in the past (i.e. the firm is in the manager's portfolio).[5] We rely on two sources of variation: the cross-section (i.e. movements from one branch to another) and the timing of the moves. We employ a difference-in-differences strategy, explained in Equation (1.1).

$$I(\text{credit})_{bft} = \beta I(\text{Inflow})_{bft} + \alpha_{bf} + \delta_{bt} + \gamma_{ft} + \epsilon_{bft} \tag{1.1}$$

[4]See the discussion in the previous paragraph.

[5]See the discussion in subsection 1.2.2 for the definition of the portfolio.

$I(\text{credit})_{bft}$ is an indicator equal to one if the firm f has been granted a loan by branch b in year t, and zero otherwise. $I(\text{Inflow})_{bft}$ is an indicator equal to one if the branch b has hired a manager that had firm f in her portfolio. It is equal to one for all the years after the pair has been treated for the first time (i.e. for the first time a manager knowing firm f has moved to branch b). Notice that since we are using worker relocations, it can be possible that a pair is treated more than once, as two or more managers having the same firm in portfolio can move to the same branch, contemporaneously or in different years. We check in the data that it is a residual case (around 3% of the treated pairs are treated more than once). Notice also that the timing of the event is specific to the firm-branch pair, i.e. some units are treated at the beginning of the sampling period, and some in later years. In our robustness checks we take into account the staggered nature of the treatment, using estimators that are robust to negative weights or cohort-specific effects (Sun and Abraham, 2021). We include a vast array of fixed effects, that control for the time-invariant characteristics of the firm-branch relationship, and for the time-varying policies of the branch and the firm. More extensive discussion on the identification strategy and the use of fixed effects is provided in subsection 1.3.1. We cluster standard errors at the bank-firm level. Our control group is composed of all the firm-branch potential pairs within the same province, that have not been involved in any manager transition. [6] Hence, the coefficient β measures the increase in the probability of a firm establishing a credit relationship with the new branch of the manager, as opposed to any other branch in the same province. In our main specification we will only rely on variation coming from moves between bank groups, that have a more profound impact on information transmission throughout the financial system. However, we will use the same empirical strategy to detect the effect of moves within the same bank group, both to benchmark our results and to inform the literature on loan officer rotation.

We estimate Equation (1.1) also through an event study specification. It allows us to provide a visual representation of our main results, and to rule out any anticipation of the move by the firms. If indeed the branch where the manager was working had poor performance, firms may have reacted to such poor performance by moving their credit relationship elsewhere. And if the manager was aware of the poor performance, she may have decided to relocate exactly to the branch where the firm had already moved. Our setup allows us to test for this anticipation effect, in two ways. First, if the firm had relocated by shutting down the credit relationship with the manager's

[6] For more details on the sample construction, see subsection 1.2.3.

old branch before her move, it would not be part of the portfolio, and hence not treated. Second, if the firm had relocated but kept the credit relationship with the manager's old branch at least until the time of the move, we would observe a higher probability of the firm establishing a credit relationship with the new branch before the event. This would be a sign of anticipation of the move, and would suggest that the manager's move was not the cause of the firm's relocation. Equation (1.2) provides a formal representation of the event study specification.

$$I(\text{credit})_{bft} = \sum_{\tau=-4}^{4} \beta_\tau \times I\{t = t_{bf} + \tau\} + \alpha_{bf} + \delta_{bt} + \gamma_{ft} + \epsilon_{bft} \tag{1.2}$$

The term t_{bf} is the year in which the manager that knows firm f moves to branch b. $I\{t = t_{bf} + \tau\}$ is an indicator equal to one if the year is τ years before or after the manager's move. The coefficient β_τ measures the year-specific variations in the matching probability between the firm and the branch. We standardize β_{-1} to zero, so that the other coefficients represent variation with respect to the baseline year. All fixed effects, sample and control group are the same as in Equation (1.1).

1.3.1 Identification discussion

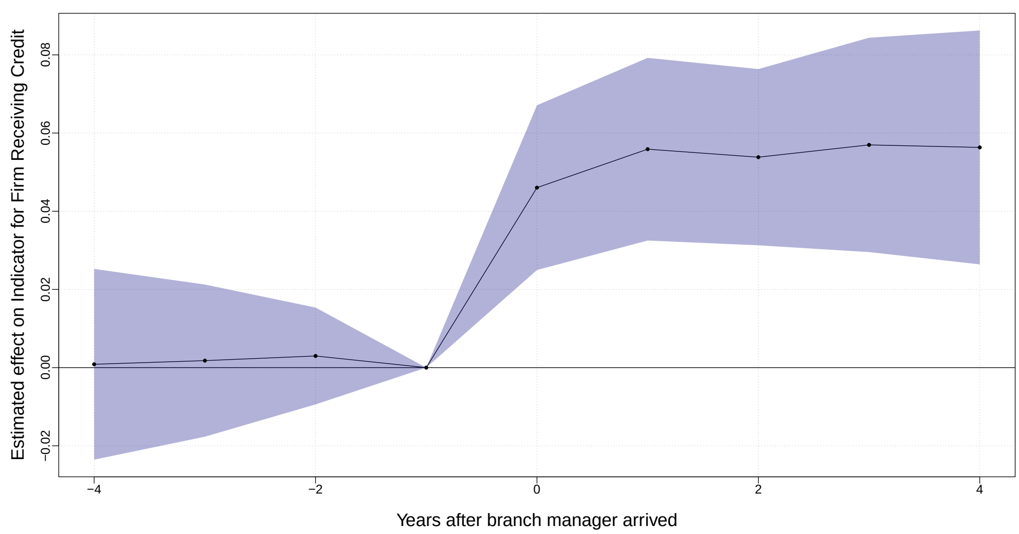
We exploit the dyadic structure of our data, and include a vast array of fixed effects. In particular, α_{bf} are firm-branch fixed effects, that capture the time invariant characteristics of the relationship between the firm and the branch. In the context of our analysis, they represent the assortative matching probability between firms and branches. They control for geographical characteristics - i.e. firms that are close to the branch are more likely to have a loan with the branch, as in the relationship lending literature (Bonfim, Nogueira, and Ongena, 2021; Keil and Ongena, 2024; Nguyen, 2019) - or industry components - in the case of lending specialization, as in Paravisini, Rappoport, and Schnabl (2023), Duquerroy et al. (2022), and Goedde-Menke and Ingermann (2023). We include branch-time fixed effects δ_{bt}, that take into account for branch-level time-varying policies, such as credit supply, branch size (hirings or layoffs), deposit inflows or outflows. As Khwaja and Mian (2008), we include firm-time fixed effects γ_{ft}, which account for firm-level time-varying characteristics, such as credit demand, investments, or financial health of the firm. A unique feature of the Italian banking sector is the presence of multi-bank relationships, even for small firms (the average number of bank relationships per firm in the credit registry is 2.76), as documented also by Gobbi and Sette (2014) and Kosekova et al. (2024). Therefore, the inclusion of firm-time fixed effects appears to be even more crucial in this context, as firms may direct their credit demand to different banks and branches over time. Then the parameter β is identified by any variation that, at the event time, changes the matching probability between the firm and the branch.

Absent any confounders, we attribute the change in the matching probability to the manager's move, as Patault and Lenoir (2024). However, there is still a threat to our identification strategy, coming from the potential endogeneity of the manager's move. Indeed, the decision of the manager to move may be driven by shocks that also change the matching probability between the firm and the branch. An example of that is some poaching activity, where the new bank actively tries to attract the firms of the manager (and maybe wants to expand in the same industry), and as a consequence the manager decides to move. An event study regression can help us control if the managers' new branch was already attracting the firms in the portfolio before the move, but cannot rule out simultaneity. β would not correctly identify the portability of the relationship if, at the same time, but separately, both the manager and the firm decided to move. Since we are interested in the firm's decision to relocate as a dependent variable, we address the manager's move endogeneity by restricting our sample of moves to those that are caused by branch closures, as in Eliason et al. (2023). Italian labor law tries to protect workers from the consequences of branch closures, and facilitates their relocation within the financial sector. Most of the relocations happen within the same bank group, but we discard them for two reasons. First, in our main specification we are interested in bank-to-bank moves. Second, when a bank group decides to close a branch, all borrowers of the branch are moved to another branch of the same bank group. We want to avoid any mechanical credit relocation, so we consider only moves in which the manager has moved to a bank group that was not involved in the decision to close the branch. By including fixed effects, and using an instrument for the manager's move, we are thus able to identify the effect of the manager's knowledge on the matching probability between the firm and the branch.

1.4 Main Results

Figure 1.2 shows the event study representation for Equation (1.2). It represents the time variation of the probability that the branch b that hires the manager establishes a credit relationship with the firm f that was in the manager's portfolio. As one can notice, there is no significant variation in the probability of creating such relationship before the manager moves. To provide a benchmark for our estimates, consider that the average probability of a firm establishing a credit relationship with any branch in the same province is around 1.2%. In the year right after the manager moves, the probability increases by 4 percentage points, further increasing by another percentage point in the following years. One can notice two main facts. First, the increase is sizable, as it represents a 3.5 times increase with respect to the baseline probability. Second, the dynamics show that managers are able to bring their clients with them right after they move, as most of the increase happens in the first two years.[7]

Figure 1.2. Increase in probability of giving credit to portfolio firms after the move



Notes: Event study estimates for the probability of a firm receiving credit from the branch where its previous loan officer has moved. Plot shows OLS estimates and 95% confidence intervals for the coefficients β_τ in equation (1.2). Branch - firm, branch - time, and firm - time fixed effects are included. Standard errors are clustered at the bank group - firm level. Sample includes all potential firm - branch pairs within a province, for the period 2009-2018. Only firms and branches that are present in the data in all years are included. Moves are defined as relocations of managers from one bank group to another. To keep consistency with the sample definition, only moves that occur within the same province are considered.

Table 1.1 shows the DiD estimates for β across different specifications. The first column shows the baseline specification, with just the inclusion of branch-firm fixed effects, to capture the assortative matching time invariant characteristics. In columns

[7] Notice that obtaining credit in Italy is rather a fast process, with the average time between the application and the disbursement of the loan being around 30 days.

(2) and (3) we include branch-time and firm-time fixed effects, while column (4) shows the results with the whole set of fixed effects. Estimates are stable across all specifications, and the coefficient hovers around 0.04.

Table 1.1. DiD estimates for β

	Credit indicator			
	(1)	(2)	(3)	(4)
Inflow × Post	0.046***	0.044***	0.047***	0.045***
	(0.011)	(0.013)	(0.011)	(0.013)
R^2	0.782	0.788	0.783	0.789
Observations	25,364,714	25,364,714	25,364,714	25,364,714
Branch-Firm fixed effects	✓	✓	✓	✓
Branch-Time fixed effects		✓		✓
Firm-Time fixed effects			✓	✓

Notes: Estimates for the probability of a firm receiving credit from the branch where its previous loan officer has moved. Column (1) includes only branch fixed effects, column (2) includes branch and time fixed effects, column (3) includes branch and firm fixed effects, and column (4) includes all fixed effects. Standard errors are clustered at the bank group - firm level. The sample includes all potential firm - branch pairs within a province, for the period 2009-2018. Only firms and branches that are present in the data in all years are included. Moves are defined as relocations of managers from one bank group to another. To keep consistency with the sample definition, only moves that occur within the same province are considered. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

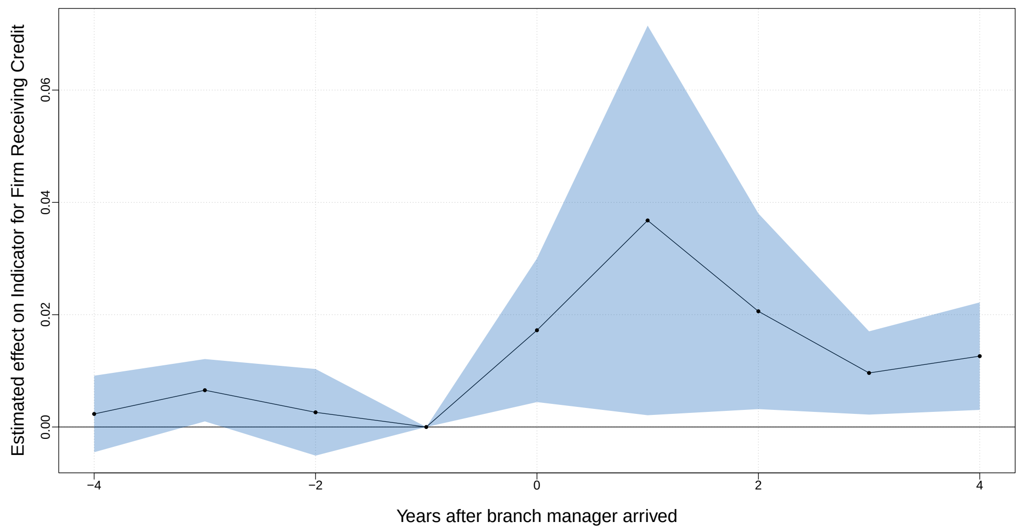
1.4.1 Branch closures identification

In order to address the potential endogeneity of the manager's move, we estimate Equation (1.1) considering only moves that are caused by branch closures. As in the previous case, we consider only bank-to-bank moves. It is noteworthy that considering only bank-to-bank moves allows us to rule out any form of mechanical credit relocation. Indeed, if a bank group decides to close a branch, it will move both the managers and the credit relationships to a different branch. By considering only managers that have moved to a different bank group, we can be sure that the increase in the matching probability is due to the manager's knowledge, and not to the bank's decision to move the credit relationships.

Using branch-closure induced moves as a source of identification helps us to address the potential endogeneity of the manager's move. However, we expect the effect to be lower than in the previous case, for two reasons. First, we can picture relocated managers as requiring more time to establish themselves in the new branch, while in the case of regular moves the new branch is already prepared to create a space for the new manager. As a consequence, we expect the matching probability to increase more slowly. Second, after a branch closure the manager may decide to temporarily accept any available position (and so relocate), but then take advantage of better opportunities in the future. So, we expect managers to stay in their new branch for a shorter time. Firms may anticipate that the manager will not stay for long, and may decide not to follow her, thus decreasing the matching probability. Another possible case - still coherent with our model - is that not only firms follow the manager once she is relocated after the branch closure, but they also decide to move to a different branch after the manager has moved. Since in our econometric specification we do not allow branch-firm pairs to switch out of treatment (i.e. once a manager brings arrives in a branch, all the firm-branch pairs in her portfolio are treated and stay treated), the matching probability would increase after the manager has moved, peak at a certain point, and then decrease.

Figure 1.3 shows the event study representation. As for the previous case, there is no anticipation of the move by the firms, that do not move to the new branch before the manager has moved. The right-hand side of the graph confirms our hypotesis: the matching probability increases more slowly than in the case of all bank moves, peaking at 4 percentage points in the second year. Then, the probability decreases (meaning that some firms have left the branch), but it remains significantly higher than the baseline probability even 4 years after the move.

Figure 1.3. Increase in credit probability to portfolio firms after branch closures



Notes: Event study estimates for the probability of a firm receiving credit from the branch where its previous loan officer has moved due to a branch closure. Plot shows OLS estimates and 95% confidence intervals for the coefficients β_τ in equation (1.2). Branch - firm, branch - time, and firm - time fixed effects are included. Standard errors are clustered at the bank group - firm level. Sample includes all potential firm - branch pairs within a province, for the period 2009-2018. Only firms and branches that are present in the data in all years are included. Moves are defined as relocations of managers from one bank group to another, and the originating branch has closed. To keep consistency with the sample definition, only moves that occur within the same province are considered.

In Table 1.2 we repeat the previous exercise, i.e. we test the robustness of our DiD estimates across different specifications, with the alternative exclusion of fixed effects. The point estimates are lower than the previous case, capturing the decrease from year 2 to year 4, but still remain significant and stable across specifications. Taking the estimate from the fully specified model in column (4), we can conclude that the knowledge of the manager causes an increase in the matching proability of 1.8 percentage points, more than doubling the baseline probability.

Table 1.2. DiD estimates for β, identified via branch closures

	Credit indicator			
	(1)	(2)	(3)	(4)
Inflow × Post	0.016***	0.017***	0.018***	0.018***
	(0.005)	(0.006)	(0.006)	(0.007)
R^2	0.801	0.805	0.802	0.805
Observations	24,125,500	24,125,500	24,125,500	24,125,500
Branch-Firm fixed effects	✓	✓	✓	✓
Branch-Time fixed effects		✓		✓
Firm-Time fixed effects			✓	✓

Notes: Estimates for the probability of a firm receiving credit from the branch where its previous loan officer has moved, after her branch has closed. Column (1) includes only branch fixed effects, column (2) includes branch and time fixed effects, column (3) includes branch and firm fixed effects, and column (4) includes all fixed effects. Standard errors are clustered at the bank group - firm level. The sample includes all potential firm - branch pairs within a province, for the period 2009-2018. Only firms and branches that are present in the data in all years are included. Moves are defined as relocations of managers from one bank group to another, and the originating branch has closed. To keep consistency with the sample definition, only moves that occur within the same province are considered. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

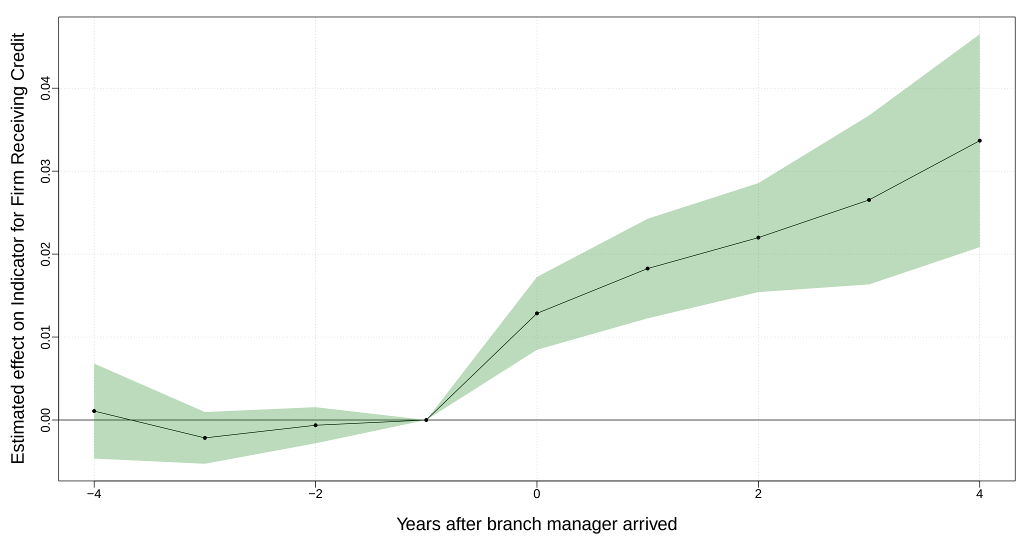
1.4.2 Within bank moves

We now turn to the case of moves within the same bank group. Moves between bank groups are most likely caused by career choices of the manager, who actively sought a new position, or by external circumstances such as branch closures. Here we consider a different type of transition: within group moves are often decided by the bank. Managers can be rotated across branches, or promoted to a different position, maybe in a larger or more strategic branch. The case of loan officer rotation has been vastly studied by theoretical (Stein, 2002) and empirical literature (Hertzberg, Liberti, and Paravisini, 2010), and is a policy designed by the bank to ensure a more trasparent monitoring, reducing the incentives for loan officers to grant credit to undeserving borrowers. However, notice that the presence of personal ties between the loan officer and the borrower can also have a positive impact on the bank's profits. Indeed, more profound knowledge between the two parts may increase the probability that the borrower stays with the bank, and can even increase the loan performance outcomes

(Fisman, Paravisini, and Vig, 2017). So, while an effective design of such policies would rely on the fact that the loan officer is not able to bring her clients with her, it is not straightforward that the bank would reduce this possibility to zero.

In Figure 1.4 we show, indeed, that even in the case of within group moves, the manager is able to bring some of her clients with her. Table 1.8 in the Appendix shows the DiD estimates for the within group moves. The probability of moving is lower than in the case of between group moves, but still significant, and it exhibits a different dynamic pattern with the case of between group moves, where the welcoming bank has no incentive to delay the transfer of new clients (if valuable). Also in this case firms do not anticipate the move. But the matching probability, instead of increasing sharply right after the move, increases constantly, and peaks 5 years after, at a level that is 3 times the baseline probability. We can interpret it as the interplay of two forces: while the manager may like to bring her clients on, the bank may have policies that prevent internal relocations. So, for the manager it may take a longer time to progressively bring her clients with her.

Figure 1.4. Credit probability increases, within bank moves



Notes: Event study estimates for the probability of a firm receiving credit from the branch where its previous loan officer has moved. Only moves within the same bank group are considered. Plot shows OLS estimates and 95% confidence intervals for the coefficients β_τ in equation (1.2). Branch - firm, branch - time, and firm - time fixed effects are included. Standard errors are clustered at the bank group - firm level. Sample includes all potential firm - branch pairs within a province, for the period 2009-2018. Only firms and branches that are present in the data in all years are included. Moves are defined as relocations of managers within the same bank group. To keep consistency with the sample definition, only moves that occur within the same province are considered.

This result provides a benchmark for the previous estimates: bank moves cause a increase in the matching probability that is 2.5 times as large as the one caused by within bank moves. We can at least partially attribute this difference in size to the bank's incentives. It also complements the existing literature on loan officer rotation. In the period covered by our sample, the number of bank groups has decreased. In our data we set the bank group at the end of the period, so that our results are net of any merger and acquisition activity. As a consequence, some of the within bank moves that we observe have happened in the context of consolidation processes. In such events, banks may have an incentive to exploit their managerial human capital to retain clients (which is what we observe in our results), but at the same time this fact reduces the monitoring effectiveness (Hertzberg, Liberti, and Paravisini, 2010).

In the following section we focus on the process by which the bank that hires the manager decides to grant credit to the firms that were in the manager's portfolio. We use data from loan applications to see if the manager is able to affect the screening process of the bank, and if the bank is able to capture the value of the manager's portfolio.

1.5 Mechanism: credit applications

In the previous section we provided evidence on the fact that when a manager moves, the firms that were in her portfolio are more likely to establish a credit relationship with her new branch. In this section we uncover the mechanism behind this phenomenon. We still focus on bank-to-bank moves, and we rely on a different data source from the credit registry, which is a request for information that the bank sends to the Bank of Italy. The object of the request is a firm, and the Bank of Italy provides the requesting bank with the firm's credit history, or its non-performing loans. It is a tool that can be used both in screening a potential borrower, so to acquire more information on the firm, or in monitoring the firm's behavior in the following years (D'Andrea, Pelosi, and Sette, 2023; Branzoli and Fringuellotti, 2020). We focus on requests for information that were filed from banks to firms that did not have previous credit relationships with the bank, so to isolate the screening component of the request. Noting that filing request for information is a rather standard procedure in the screening process, we can also interpret those requests as a proxy for loan applications. Indeed, since requests for information come from the bank, we can reasonably assume that they come after the screened firm has previously applied for a loan.

The Bank of Italy stores requests for information data at the firm - bank group - year level. We are not able to identify the branch from which the request comes, but we can still reliably assume that the request is sent by bank where the manager has moved, and where the firm will eventually have credit, if the request is accepted. Notice also that we are interested in differences in the probability of the request being sent (or accepted), so measurement error is constant throughout the sample, and does not affect our estimates. We make two conjectures on the manager's role in the screening process. First, we expect branch managers to influence search, i.e. to induce the firms that were in her portfolio to apply for a loan. Second, we expect managers to influence the screening process, i.e. to increase the probability of the request being accepted.

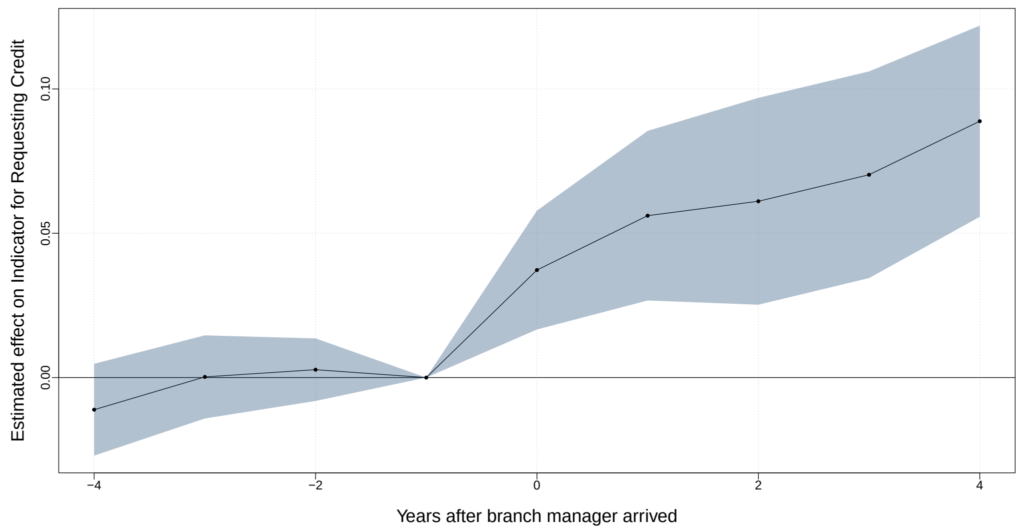
1.5.1 Search

Proxing loans applications with requests for information, we can test if the manager is able to affect the probability of the firm applying for a loan. If this is the case, we should see an increase in the requests of portfolio firms after the manager moves. Equation (1.3) shows the empirical strategy we use to test this conjecture.

$$I(\text{applied})_{bft} = \sum_{\tau=-4}^{4} \beta_\tau \times I\{t = t_{bf} + \tau\} + \alpha_{bf} + \delta_{bt} + \gamma_{ft} + \epsilon_{bft} \tag{1.3}$$

The dependent variable $I(\text{applied})_{bft}$ is an indicator equal to one if the bank b has requested information on firm f in year t or earlier, and zero otherwise. The right-hand side of the equation is the same as in Equation (1.2), with the only difference that observations are collapsed at the firm-bank group level. The unit of observation becomes then the bank-firm pair, as opposed to the branch-firm pair in the previous analysis. Therefore fixed effects capture the underlying matching probability between a bank group and a firm, and the time-varying characteristics of the bank group (as well as the firm). Figure 1.5 shows the event study representation of the results. The baseline probability of a bank requesting information on a firm is 2.5%. Five years after the manager moves, the probability increases by 8 percentage points, reaching a level that is 4 times higher than the baseline probability.

Figure 1.5. Increase in probability of applying for credit to the new bank



Notes: Event study estimates for the probability of a firm applying for credit in the bank where its bank manager has moved. Plot shows OLS estimates and 95% confidence intervals for the coefficients β_τ in equation (1.3). Bank - firm, bank - time, and firm - time fixed effects are included. Standard errors are clustered at the bank group - firm level. Sample includes all potential firm - bank pairs, for the period 2009-2018. A pair is considered potential if in the province where the firm was present, there was at least one branch of the bank. Only firms and branches that are present in the data in all years are included. Moves are defined as relocations of managers from one bank group to another. To keep consistency with the sample definition, only moves that occur within the same province are considered.

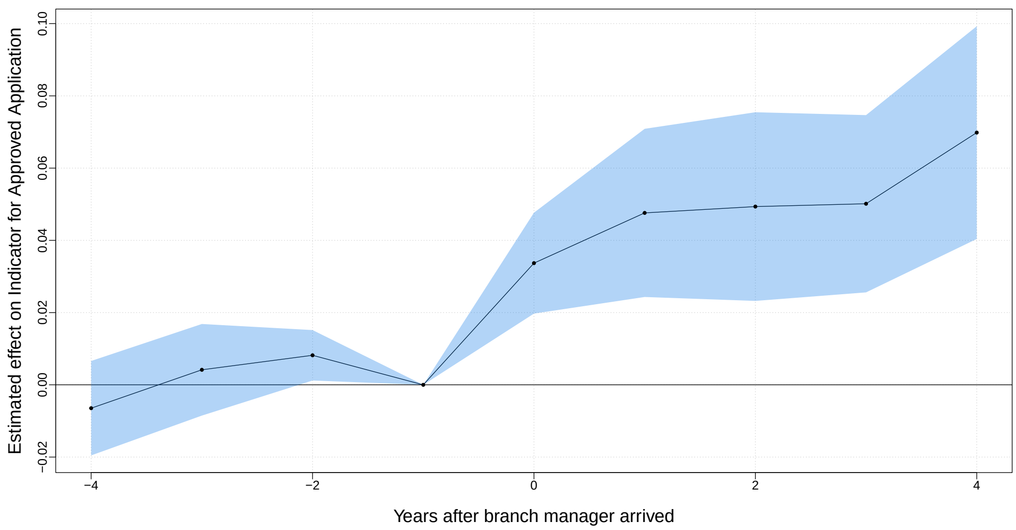
1.5.2 Screening

Our second conjecture is that the manager not only affects the probability of the firm applying for a loan, but also the probability of the request being accepted. In order to test this conjecture, we estimate Equation (1.4), by regressing the conditional probability of the request being accepted on the manager's move. We build the conditional acceptance probability $I(\text{accepted}|\text{applied})_{bft}$ in our dataset by checking the rows in our dataset where the firm has requested for information, and the bank has granted the loan. The equation reads as follows:

$$I(\text{accepted}|\text{applied})_{bft} = \sum_{\tau=-4}^{4} \beta_\tau \times I\{t = t_{bf} + \tau\} + \alpha_{bf} + \delta_{bt} + \gamma_{ft} + \epsilon_{bft} \qquad (1.4)$$

The conditional acceptance probability mean in our sample is 0.35, which means that the bank accepts a loan request in 35% of the cases. Figure 1.6 shows the event study plot. Five years after the manager has moved, the probability of the request being accepted increases by 7 percentage points, reaching a level that is 25% higher than the baseline probability. Therefore, not only the manager's firms are more likely to apply for a loan, but also the bank is more likely to accept the request.

Figure 1.6. Increase in probability of being approved for credit in the new bank



Notes: Event study estimates for the probability of a firm being approved for credit in the bank where its bank manager has moved. Plot shows OLS estimates and 95% confidence intervals for the coefficients β_τ in equation (1.4). Bank - firm, bank - time, and firm - time fixed effects are included. Standard errors are clustered at the bank group - firm level. Sample includes all potential firm - bank pairs, for the period 2009-2018. A pair is considered potential if in the province where the firm was present, there was at least one branch of the bank. Only firms and branches that are present in the data in all years are included. Moves are defined as relocations of managers from one bank group to another. To keep consistency with the sample definition, only moves that occur within the same province are considered.

In the following sections we will explore the features of the match between the manager and the firm, and will investigate what kind of value the manager brings to the bank that hires her.

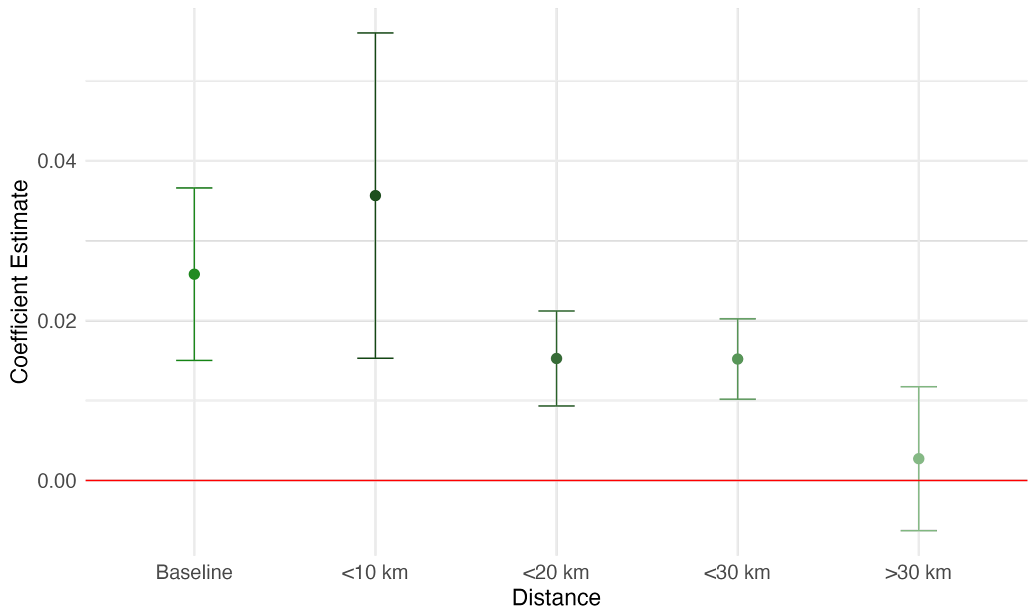
1.6 Heterogeneity

1.6.1 Distance

Banking literature has often used distance between the borrower and the lender as a way to define relationship lending (Bonfim, Nogueira, and Ongena, 2021; Keil and Ongena, 2024; Amberg and Becker, 2024). In our context, we are able to use a more direct measure of the relationship between a loan officer and a borrower, as we construct a portfolio for each manager.

We can then exploit the structure of our dataset to understand the interplay betweeen our measure of relationship and the physical distance of the managers' move. We define the distance between the old and the new branch as the distance between the centroids of the municipalities where the branches are located. We then test if the effect of the manager's move on the probability of the firm establishing a credit relationship with the new branch is affected by the distance of the move. We do it by interacting the β coefficient in Equation (1.1) with a set of dummies that capture the distance of the move. Figure 1.7 shows the results.

Figure 1.7. Estimates of β for different distances of the move



Notes: The figure plots estimates of β across different distances. The first column plots the average treatment effect, as in column (4) of Table 1.1. Columns (2) to (5) plot estimates of β for different distances, respectively less than 10 km, between 10 and 20 km, between 20 and 30 km, and more than 30 km. The estimates are obtained by running the same regression as column (4) in Table 1.1 (i.e. including both pair, firm-time and bank-time fixed effects) on the subsample of moves that fall within the distance range. The estimates are then plotted against the average distance of the moves in the subsample. The lines represent 95% confidence intervals. The red line represents a treatment effect of zero.

The first column shows the baseline effect of the manager's move, as computed in the previous sections. The second column shows the effect of the manager's move when the distance is less than 10 km. As one can notice, managers that move closer to their old branch are more likely to bring their clients with them. The trend is confirmed by the third and the fourth columns, that show the effects for moves closer than 20 km and 30 km, respectively. However, although reduced in magnitude, the effect is still significant for moves up to 30 km. This result is particularly interesting, as it shows that the manager's portability of her clients exceeds the 5/10 kilometers range, which is what the literature has often used as a threshold for relationship lending: Degryse and Ongena, 2005 set this threshold at 1.4 miles, Petersen and R. G. Rajan, 2002 at 4 miles, Crawford, Pavanini, and Schivardi, 2018 at 1.8 miles.

1.6.2 Span of control

When defining our dataset, we confined ourselves at small branches (with less than 5 managers) in order to be able to have a reliable matching between the manager and the firms. However, within the small branches, the number of managers can vary. We then test for the span of control of the manager. Indeed, if a manager was working in a larger branch, with more managers, although possibly interacting with all the firms that were receiving credit from the branch, she may have had less chances to build a relationship with all of them, and maybe she focused on interacting with a subset of them. We interact the coefficient β in Equation (1.1) with a set of dummies that capture the number of managers in the branch (one, one or two, up to three). Notice that this regression can be also interpreted as a robustness check of our results, as we are able to test the sensitivity of our estimates to the choice of the threshold for the number of managers. There is a clear trade-off between increasing the precision of our estimates (i.e. reducing our threshold for the number of managers) and considering a sizable number of branches. Table 1.3 shows the results.

Table 1.3. Heterogeneity by number of managers

	Credit indicator		
	(1)	(2)	(3)
From one manager branch	0.109*		
	(0.057)		
From one/two managers branch		0.077***	
		(0.028)	
From one/two/three managers branch			0.059***
			(0.021)
R^2	0.786	0.786	0.786
Observations	25,422,560	25,422,560	25,422,560
Branch-Firm fixed effects	✓	✓	✓
Firm-Time fixed effects	✓	✓	✓
Branch-Time fixed effects	✓	✓	✓

Notes: DiD estimates for β in equation 1.1 for different number of managers in the branch the loan officer moved from. Column (1) includes moves from branches with one manager, column

(2) includes moves from branches with one or two managers, and column (3) includes moves from branches with one, two, or three managers. Only firms and branches that are present in the data in all years are included. Moves are defined as relocations of managers from one bank group to another. To keep consistency with the sample definition, only moves that occur within the same province are considered. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

The first row shows the results for branches that had just one manager when she moved. The estimate for the portability of the manager's clients increases from the baseline of 0.04 to 0.10 when the manager is the only one in the branch. Put it differently, managers that were the only ones in the branch are able to bring 10% of their clients with them. When we consider also branches with two managers, the effect decreases to 0.08 and then up to 0.06 for branches with three managers. We can rationalize this result both with the measurement argument (i.e. the more managers, the less likely the manager was in contact with all the firms), and with the span of control argument.

1.6.3 Bank size

We then turn to another dimension of heterogeneity: the bank. As Stein (2002) explains, the structure of information within a firm has a crucial role on the internal organization. He argues that his model is mostly suited for the banking sector, where there is the contemporaneous presence of hard and soft information. Some banks - the large ones in the model - may rely more on hard information, and so have a centralized credit awarding process, where the role of the branch is limited to the collection and the "hardening" of credit information, so that it can be used by the central office. On the other hand, another type of banks - typically smaller - may rely more on soft information, and so have a decentralized credit awarding process, where the branch manager has a more prominent role in the credit allocation.

We test this conjecture by interacting the manager's move with a dummy that captures the size of the bank. We define large banks as banks that have more than 5 percent of the total number of employees in the banking sector, and small banks as the remaining ones. There are then 4 categories, interacting the previous bank size (large or small) with the size of the manager's new bank. Theory would predict that large banks are less likely to let the manager bring her clients with her, especially if the manager is coming from a small bank. Indeed, if large banks rely on a more centralized credit allocation system, they would not place high value in the manager's soft information, and so would not be willing to let her bring her clients with her. On the other hand, small banks may be more willing to let the manager bring her clients with her, as they may rely more on the manager's soft information. It is not a priori clear which kind of moves would bring more clients to a small bank, if manager is coming from a large or a small bank. One could argue that managers coming from large banks may have more clients that are dissatisfied with the centralized credit allocation process, and so moving to a small bank may be a way for them to make use of the soft information they have. Viceversa, a manager coming from a small bank may have already used her soft information to benefit her clients, and so they would be more willing to move with her to another small bank. Table 1.4 shows the results. For each regression we estimate the baseline effect (first row of the table), and then the interaction term between the manager's move and the bank size.

Table 1.4. Heterogeneity by type of bank move

	Credit indicator			
	(1)	(2)	(3)	(4)
Inflow × Post	0.042***	0.049***	0.058***	0.054**
	(0.016)	(0.015)	(0.015)	(0.024)
Inflow × Post × Large to small	0.052			
	(0.035)			
Inflow × Post × Large to large		-0.008		
		(0.016)		
Inflow × Post × Small to large			-0.049***	
			(0.017)	
Inflow × Post × Small to small				-0.010
				(0.034)
R^2	0.786	0.786	0.786	0.786
Observations	25,422,560	25,422,560	25,422,560	25,422,560
Branch-Firm fixed effects	✓	✓	✓	✓
Firm-Time fixed effects	✓	✓	✓	✓
Branch-Time fixed effects	✓	✓	✓	✓

Notes: DiD estimates for β in equation 1.1 for different types of bank moves. Column (1) includes an interaction term with moves from a large bank to a small bank, column (2) includes an interaction term with moves from a large bank to a large bank, column (3) includes an interaction term with moves from a small bank to a large bank, and column (4) includes an interaction term with moves from a small bank to a small bank. Small banks are defined as the ones which have less than 5 percent of the total number of employees in the banking sector. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Consistently with the theory, movements from a small bank to a large bank are almost ineffective, as the interaction term counterbalances the baseline effect. Managers moving from a small bank to a large bank are the least likely to bring their clients with them, as they would not be able to use their soft information in the new bank. Movements within the bin size (i.e. large to large and small to small) add almost zero variation to the main effect, suggesting in one case that small banks clients were already benefiting from the manager's soft information, while in the other case that large banks clients were already used to the centralized credit allocation process. On the other hand, movements from a large bank to a small bank are the most effective, as the interaction term is positive and doubles the baseline effect. We can label those movements as the ones in which the information is "softened" by the manager, and the new bank is willing to use it.

1.6.4 Firm characteristics

The last dimension of heterogeneity we investigate on relates to the characteristics of the firms in the manager's portfolio. We examine three aspects of the firms: size, age, and credit risk. One could expect that smaller firms are the ones that rely more on the direct knowledge of the manager, since they do not have the same access to credit as larger firms. We follow Eurostat cathegorization and define four bins of firm size: micro firms, with less than 10 employees, small firms, that have less than 50 employees (thus including the micro firms), medium firms, that have less than 250 employees (thus including small and micro firms), and large firms, that have more than 250 employees. We then interact the treatment with dummies that capture the size of the firms. Table 1.5 shows the results.

Table 1.5. Heterogeneity by firm size

	Credit indicator			
	(1)	(2)	(3)	(4)
Inflow × Post	0.043***	0.052***	0.055***	0.054***
	(0.015)	(0.014)	(0.015)	(0.013)
Inflow × Post × Micro	0.020			
	(0.015)			
Inflow × Post × Small		-0.001		
		(0.013)		
Inflow × Post × Medium			-0.025	
			(0.018)	
Inflow × Post × Big				-0.058
				(0.062)
R^2	0.789	0.789	0.789	0.789
Observations	24,645,380	24,645,380	24,645,380	24,645,380
Branch-Firm fixed effects	✓	✓	✓	✓
Firm-Time fixed effects	✓	✓	✓	✓
Branch-Time fixed effects	✓	✓	✓	✓

Notes: DiD estimates for β in equation 1.1 for different types of firms. The dependent variable is the credit indicator. Standard errors are clustered at the branch-firm level. Column (1) includes an interaction between β and a dummy for the firm being micro (less than 10 employees), column (2) includes an interaction with a dummy for the firm being small (less than 50 employees), column (3) includes an interaction with a dummy for the firm being medium (less than 250 employees), and column (4) includes an interaction with a dummy for the firm being big (more than 250 employees). Moves are defined as relocations of managers from one bank group to another. To keep consistency with the sample definition, only moves that occur within the same province are considered. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Although with a noisy estimate, the interaction term for micro firms is positive, suggesting that when firms are so little, they need a more direct relationship with the bank manager, and so they are more likely to follow her. As firms grow, the effect decreases, and it becomes (although not significantly) negative for medium and large firms.

We then turn to age and credit risk. With respect to age, two contrasting theories can be formulated. On the one hand, younger firms may be more dependent on the manager's soft information, as they may not have a long credit history, and so the manager's information may be more valuable. On the other hand, older firms may have built a long lasting relationship with the bank manager, who can transfer them more easily when she moves. We define two bins of firm age: young firms, that had been founded in the 5 years before the manager moved, and old firms, that had been founded more than 10 years before the manager moved.

Similarly, with credit risk, the a priori effect is not clear. On the one hand, bank managers can engage in cream-skimming (Di Maggio and Yao, 2021), and so they may be more likely to bring with them less risky firms. On the other hand, firms that are labelled as risky may be the ones that need the manager's soft information the most, and so they may be more likely to follow her. The Italian credit registry provides a credit risk score for each firm, which is computed every year by Cebi-Cerved (Rodano, Serrano-Velarde, and Tarantino, 2018). The score ranges from 1 to 10, where 1 is the lowest risk and 10 is the highest risk. We take the Cebi-Cerved definitions and label as low risk firms that have a score lower or equal than 4, and as high risk firms that have a score higher or equal to 7. Table 1.6 shows the results for the age and credit risk of the firms. No interaction term appears to be significant, suggesting that the two effects may cancel each other out, both for age and credit risk.

Table 1.6. Heterogeneity by firm age and credit risk

	Credit indicator			
	(1)	(2)	(3)	(4)
Inflow × Post	0.052***	0.053***	0.055***	0.053***
	(0.014)	(0.016)	(0.012)	(0.017)
Inflow × Post × Young	-0.019			
	(0.033)			
Inflow × Post × Old		-0.002		
		(0.010)		
Inflow × Post × Safe			-0.007	
			(0.012)	
Inflow × Post × Risky				-0.007
				(0.025)
R^2	0.789	0.789	0.789	0.789
Observations	24,645,380	24,645,380	24,645,380	24,645,380
Branch-Firm fixed effects	✓	✓	✓	✓
Firm-Time fixed effects	✓	✓	✓	✓
Branch-Time fixed effects	✓	✓	✓	✓

Notes: DiD estimates for β in equation 1.1 for different types of firms. The dependent variable is the credit indicator. Standard errors are clustered at the branch-firm level. Column (1) includes an interaction with a dummy for the firm being young (less than 5 years old), column (2) includes an interaction with a dummy for the firm being old (more than 10 years old), column (3) includes an interaction with a dummy for the firm being safe (credit score below 4 in a 1-10 scale), and column (4) includes an interaction with a dummy for the firm being risky (credit score above 6 in a 1-10 scale). Moves are defined as relocations of managers from one bank group to another. To keep consistency with the sample definition, only moves that occur within the same province are considered. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

1.7 Robustness

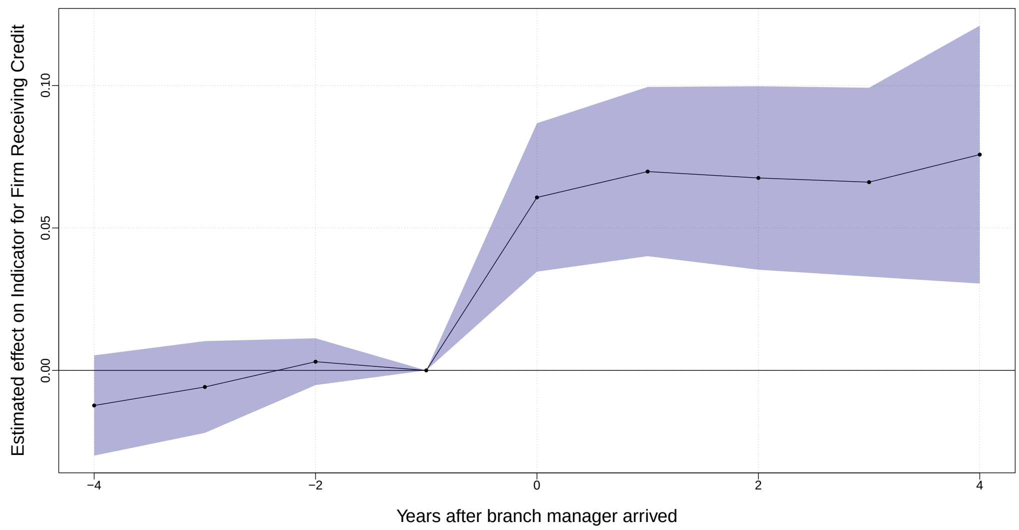
1.7.1 Alternative sample

We perform a robustness check on the sample construction. As we explain in section 1.2, we have constructed our sample by considering all the potential firm - branch pairs in a province. We made this choice in order to have a balanced panel, but it increased the dimensionality of our data and reduced the computational feasibility of our regressions. Indeed, as explained in section 1.2, we run the regressions of sections 3.3 and 1.6 on a random sample of 25 million observations. Furthermore, as it is explained in the data construction section, we made some restrictive assumptions on the branches, excluding the ones that were located in the largest municipality of each province.

We then test the robustness of our results by aggregating branches at the province level, that is the geographical definition of the credit market (Panetta, Schivardi, and Shum, 2009; Crawford, Pavanini, and Schivardi, 2018; Barone, Schivardi, and Sette, 2024). Instead of considering all the branches in a province, we consider only one bank - province pair for each bank operating in the province. We still define the manager's portfolio as the set of firms that were receiving loans in the municipality and bank group where she was working. But instead of considering all the possible branches (bank - municipality pairs) where she could have moved, we only consider the bank group where she moved (if she moved within the province). Therefore the unit of observation in this exercise is the bank - province - firm triplet. Then β in this exercise will capture the effect of the manager's move on the probability of the firm establishing a credit relationship with the new bank group in the province, as opposed to just the branch where the manager moved.

The rationale behind this sample definition is twofold. First, it allows us to estimate a variation of Equation (1.2) on the full sample, rather than on a subsample of the more granular data. Second, it allows us to test for a different kind of pre-trends. If in figure 1.2 we were controlling that the firm did not have a pre-existing relationship with the branch where the manager moved, in this specification we are controlling for no pre-existing relationship with her new bank. The broader implication of this exercise is that manager moves are actually reshaping credit allocation, and not only affecting reallocation of credit within the bank group.

Figure 1.8. Robustness: Bank-Firm Dataset



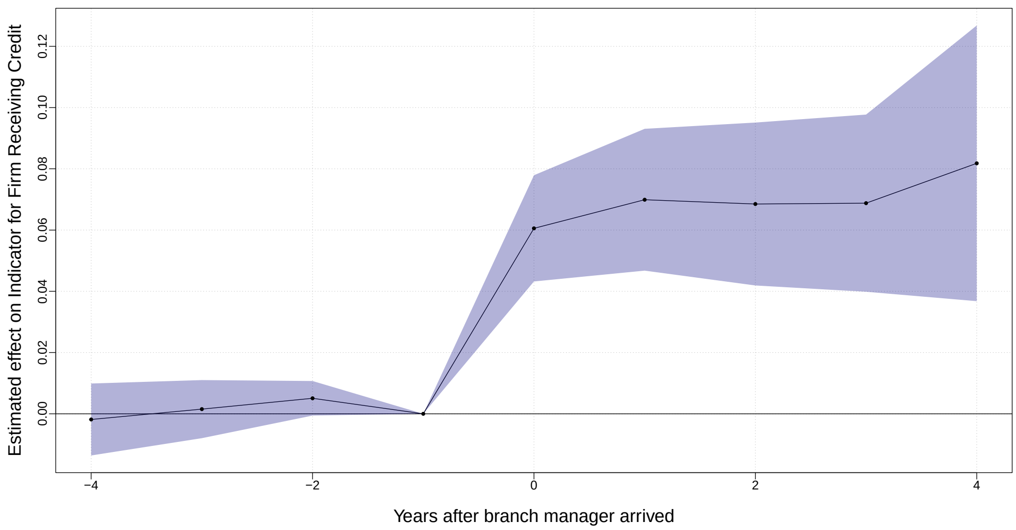
Notes: Event study estimates for the probability of a firm receiving credit from the branch where its previous loan officer has moved, as in Figure 1.2. Here branches are not defined as bank-municipality pairs, but rather as bank-province pairs (i.e. one branch per province).

We then estimate Equation (1.2) on the aggregated sample, and we test the robustness of our results. Figure 1.8 shows the results of the regression, which are extremely consistent with the ones in Figure 1.2. The arrival of the manager increases the probability of the firm establishing a credit relationship with the new bank group by 8 percentange points after 4 years.

1.7.2 Robust estimators

The third robustness check we perform is on the estimation method. The recent debate on Two-Way Fixed Effects estimators has shown that they may be biased in the presence of unobserved heterogeneity (De Chaisemartin and D'Haultfœuille, 2020), as it can generate negative weights for some observations. One of the concerns with the OLS estimates we have provided is that different years may be subject to heterogeneous macroeconomic conditions, that may alter the probability of firm creating new credit relationships (e.g. a period of high interest rates may reduce demand for credit, and so the probability of a firm establishing a credit relationship with a bank). To avoid those factors to impact our estimates, we estimate Equation (1.2) with the estimator proposed by Sun and Abraham (2021), which allows for cohort-specific heterogeneity. Figure 1.9 shows the results for the province-level aggregation. Appendix figure 1.12 reproduces the results for the main specification in the previous sections. In both cases the results are consistent with the ones we have found in the previous sections.

Figure 1.9. Robustness: Sun and Abraham (2021) estimator, province level sample



Notes: Event study estimates for the probability of a firm receiving credit from the branch where its previous loan officer has moved. Plot shows estimates and 95% confidence intervals for the coefficients β_τ in equation (1.2), estimated using the TWFE estimator proposed by Sun and Abraham (2021). Branch - firm, branch - time, and firm - time fixed effects are included. Standard errors are clustered at the bank group - firm level. Sample includes all potential firm - branch pairs, for the period 2009-2018. Branches are defined as bank-province pairs (i.e. one branch per province). Only firms and branches that are present in the data in all years are included. Moves are defined as relocations of managers from one bank group to another. To keep consistency with the sample definition, only moves that occur within the same province are considered.

1.8 Conclusion

In this paper, we show that when bank managers move, they bring their clients with them. By building personal relationships and acquiring soft information, managers use this knowledge to benefit the bank that hires them.

A manager's move increases a firm's likelihood of establishing a credit relationship with the new bank by 5 percentage points — more than four times the baseline matching probability. Using loan application data, we identify two key channels: first, managers increase the likelihood that firms they know apply for loans at their new bank; second, they improve firms' chances of loan approval. These effects are stronger when the manager is the only one in the branch, is older, moves from a large bank to a smaller one, or when her portfolio consists of micro and older firms. Our results highlight the role of soft information in credit allocation and show that while geographical distance is important, it is not the sole determinant of borrower-lender relationships.

We interpret our findings as a preliminary but necessary step toward answering broader questions about the role of soft information in credit allocation. Does the manager benefit from her portfolio in terms of career advancement or compensation? Is she incentivized to bring clients along? For the hiring bank, does the manager improve the screening process, helping acquire better clients and increasing profitability? Finally, how valuable is the manager to the firms? Do firms that follow her experience better survival, growth, or profitability outcomes? We plan to explore these and other questions in future research.

To conclude, in an era of debranching and increasing reliance on algorithms, our results suggest that human capital remains vital in credit allocation. To effectively transition to a more depersonalized process, the unique value of soft information must be considered.

References

Acabbi, Edoardo Maria, Ettore Panetti, and Alessandro Sforza (2024). "The Financial Channels of Labor Rigidities: Evidence from Portugal". en. In.

Amberg, Niklas and Bo Becker (2024). "Banking Without Branches". en. In.

Babina, Tania, Greg Buchak, and Will Gornall (2022). "Customer Data Access and Fintech Entry: Early Evidence from Open Banking". en. In: *SSRN Electronic Journal*. ISSN: 1556-5068. DOI: 10.2139/ssrn.4071214. URL: https://www.ssrn.com/abstract=4071214 (visited on 02/05/2024).

Bandiera, Oriana et al. (2020). "CEO Behavior and Firm Performance". en. In: *Journal of Political Economy*.

Barone, Guglielmo, Fabiano Schivardi, and Enrico Sette (2024). "Interlocking Directorates and Competition in Banking". en. In: *Journal of Finance* Forthcoming. ISSN: 1556-5068. DOI: 10.2139/ssrn.4153387. URL: https://www.ssrn.com/abstract=4153387 (visited on 02/11/2025).

Bell, Brian and John Van Reenen (Feb. 2014). "Bankers and Their Bonuses". en. In: *The Economic Journal* 124.574, F1–F21. ISSN: 0013-0133, 1468-0297. DOI: 10.1111/ecoj.12101. URL: https://academic.oup.com/ej/article/124/574/F1-F21/5078048 (visited on 02/06/2025).

Beraldi, Francesco (2025). "Banking Relationships and Loan Pricing Disconnect". en. In.

Bloom, Nicholas et al. (Feb. 2013). "Does Management Matter? Evidence from India*". en. In: *The Quarterly Journal of Economics* 128.1, pp. 1–51. ISSN: 0033-5533, 1531-4650. DOI: 10.1093/qje/qjs044. URL: https://academic.oup.com/qje/article/128/1/1/1838606 (visited on 02/06/2025).

Böhm, Michael J, Daniel Metzger, and Per Strömberg (Sept. 2023). ""Since You're So Rich, You Must Be Really Smart": Talent, Rent Sharing, and the Finance Wage Premium". en. In: *Review of Economic Studies* 90.5, pp. 2215–2260. ISSN: 0034-

6527, 1467-937X. DOI: 10.1093/restud/rdac079. URL: https://academic.oup.com/restud/article/90/5/2215/6884037 (visited on 01/20/2025).

Bolton, Patrick et al. (2016). "Relationship and Transaction Lending in a Crisis". en. In: 29.10.

Bonfim, Diana, Gil Nogueira, and Steven Ongena (July 2021). ""Sorry, We're Closed" Bank Branch Closures, Loan Pricing, and Information Asymmetries". en. In: *Review of Finance* 25.4, pp. 1211–1259. ISSN: 1572-3097, 1573-692X. DOI: 10.1093/rof/rfaa036. URL: https://academic.oup.com/rof/article/25/4/1211/5985541 (visited on 02/11/2025).

Boustanifar, Hamid, Everett Grant, and Ariell Reshef (Mar. 2018). "Wages and Human Capital in Finance: International Evidence, 1970–2011*". en. In: *Review of Finance* 22.2, pp. 699–745. ISSN: 1572-3097, 1573-692X. DOI: 10.1093/rof/rfx011. URL: https://academic.oup.com/rof/article/22/2/699/3075313 (visited on 02/06/2025).

Branzoli, Nicola and Fulvia Fringuellotti (2020). "The Effect of Bank Monitoring on Loan Repayment". en. In: *SSRN Electronic Journal*. ISSN: 1556-5068. DOI: 10.2139/ssrn.3601944. URL: https://www.ssrn.com/abstract=3601944 (visited on 02/11/2025).

Chodorow-Reich, Gabriel (Feb. 2014). "The Employment Effects of Credit Market Disruptions: Firm-level Evidence from the 2008–9 Financial Crisis*". en. In: *The Quarterly Journal of Economics* 129.1, pp. 1–59. ISSN: 0033-5533, 1531-4650. DOI: 10.1093/qje/qjt031. URL: https://academic.oup.com/qje/article/129/1/1/1899226 (visited on 10/25/2023).

Cornell, Bradford and Ivo Welch (June 1996). "Culture, Information, and Screening Discrimination". en. In: *Journal of Political Economy* 104.3, pp. 542–571. ISSN: 0022-3808, 1537-534X. DOI: 10.1086/262033. URL: https://www.journals.uchicago.edu/doi/10.1086/262033 (visited on 02/11/2025).

Crawford, Gregory S., Nicola Pavanini, and Fabiano Schivardi (July 2018). "Asymmetric Information and Imperfect Competition in Lending Markets". en. In: *American Economic Review* 108.7, pp. 1659–1701. ISSN: 0002-8282. DOI: 10.1257/aer.20150487. URL: https://pubs.aeaweb.org/doi/10.1257/aer.20150487 (visited on 02/11/2025).

D'Andrea, Angelo, Marco Pelosi, and Enrico Sette (2023). "When Broadband Comes to Banks: Credit Supply, Market Structure, and Information Acquisition". en. In:

SSRN Electronic Journal. ISSN: 1556-5068. DOI: 10.2139/ssrn.4351548. URL: https://www.ssrn.com/abstract=4351548 (visited on 02/05/2025).

De Chaisemartin, Clément and Xavier D'Haultfœuille (Sept. 2020). "Two-Way Fixed Effects Estimators with Heterogeneous Treatment Effects". en. In: *American Economic Review* 110.9, pp. 2964–2996. ISSN: 0002-8282. DOI: 10.1257/aer.20181169. URL: https://pubs.aeaweb.org/doi/10.1257/aer.20181169 (visited on 08/23/2023).

De Marco, Filippo, Julien Sauvagnat, and Enrico Sette (2023). "Lending to Overconfident Borrowers". en. In.

Degryse, Hans and Steven Ongena (Feb. 2005). "Distance, Lending Relationships, and Competition". en. In: *The Journal of Finance* 60.1, pp. 231–266. ISSN: 0022-1082, 1540-6261. DOI: 10.1111/j.1540-6261.2005.00729.x. URL: https://onlinelibrary.wiley.com/doi/10.1111/j.1540-6261.2005.00729.x (visited on 02/11/2025).

Di Maggio, Marco and Vincent Yao (Sept. 2021). "Fintech Borrowers: Lax Screening or Cream-Skimming?" en. In: *The Review of Financial Studies* 34.10. Ed. by Itay Goldstein, pp. 4565–4618. ISSN: 0893-9454, 1465-7368. DOI: 10.1093/rfs/hhaa142. URL: https://academic.oup.com/rfs/article/34/10/4565/6054822 (visited on 04/24/2023).

Diamond, Douglas W (1984). "Financial intermediation and delegated monitoring". In: *The review of economic studies* 51.3. Publisher: Wiley-Blackwell, pp. 393–414.

— (2001). "Should Japanese banks be recapitalized?" In: *Monetary and Economic Studies* 19.2. Publisher: Institute for Monetary and Economic Studies, Bank of Japan, pp. 1–19.

Duquerroy, Anne et al. (2022). "Bank Local Specialization". en. In: *SSRN Electronic Journal*. ISSN: 1556-5068. DOI: 10.2139/ssrn.4034634. URL: https://www.ssrn.com/abstract=4034634 (visited on 10/20/2023).

Eliason, Marcus et al. (2023). "Social connections and the sorting of workers to firms". In: *Journal of Econometrics* 233.2. Publisher: Elsevier, pp. 468–506.

Fisman, Raymond, Daniel Paravisini, and Vikrant Vig (Feb. 2017). "Cultural Proximity and Loan Outcomes". en. In: *American Economic Review* 107.2, pp. 457–492. ISSN: 0002-8282. DOI: 10.1257/aer.20120942. URL: https://pubs.aeaweb.org/doi/10.1257/aer.20120942 (visited on 01/29/2024).

Focarelli, Dario and Fabio Panetta (Aug. 2003). "Are Mergers Beneficial to Consumers? Evidence from the Market for Bank Deposits". en. In: *American Economic*

Review 93.4, pp. 1152–1172. ISSN: 0002-8282. DOI: 10.1257/000282803769206241. URL: https://pubs.aeaweb.org/doi/10.1257/000282803769206241 (visited on 03/30/2023).

Gao, Janet, Wenyu Wang, and Yufeng Wu (2025). "Human capital mobility and careers of M&A advisors". en. In: *Review of Financial Studies* Forthcoming.

Gao, Janet, Wenyu Wang, and Xiaoyun Yu (2024). "Big Fish in Small Ponds: Human Capital Mobility and the Rise of Boutique Banks". en. In: *Management Science* 70.11, pp. 7829–7850. ISSN: 1556-5068. DOI: 10.1287/mnsc.2021.03829. (Visited on 06/24/2023).

Gennaioli, Nicola, Andrei Shleifer, and Robert Vishny (Feb. 2015). "Money Doctors". en. In: *The Journal of Finance* 70.1, pp. 91–114. ISSN: 0022-1082, 1540-6261. DOI: 10.1111/jofi.12188. URL: https://onlinelibrary.wiley.com/doi/10.1111/jofi.12188 (visited on 10/12/2023).

Gobbi, Giorgio and Enrico Sette (Apr. 2014). "Do Firms Benefit from Concentrating their Borrowing? Evidence from the Great Recession*". en. In: *Review of Finance* 18.2, pp. 527–560. ISSN: 1573-692X, 1572-3097. DOI: 10.1093/rof/rft017. URL: https://academic.oup.com/rof/article/18/2/527/1580190 (visited on 02/05/2025).

Goedde-Menke, Michael and Peter-Hendrik Ingermann (Dec. 2023). "Loan Officer Specialization and Credit Defaults". en. In: *Journal of Banking & Finance*, p. 107077. ISSN: 03784266. DOI: 10.1016/j.jbankfin.2023.107077. URL: https://linkinghub.elsevier.com/retrieve/pii/S0378426623002728 (visited on 01/29/2024).

Hertzberg, Andrew, Jose Maria Liberti, and Daniel Paravisini (June 2010). "Information and Incentives Inside the Firm: Evidence from Loan Officer Rotation". en. In: *The Journal of Finance* 65.3, pp. 795–828. ISSN: 0022-1082, 1540-6261. DOI: 10.1111/j.1540-6261.2010.01553.x. URL: https://onlinelibrary.wiley.com/doi/10.1111/j.1540-6261.2010.01553.x (visited on 10/09/2023).

Jasova, Martina et al. (2021). "Monetary Policy, Labor Income Redistribution and the Credit Channel: Evidence from Matched Employer-Employee and Credit Registers". en. In: *SSRN Electronic Journal*. ISSN: 1556-5068. DOI: 10.2139/ssrn.3930286. URL: https://www.ssrn.com/abstract=3930286 (visited on 10/02/2024).

Keil, Jan and Steven Ongena (Jan. 2024). "The demise of branch banking – Technology, consolidation, bank fragility". en. In: *Journal of Banking & Finance* 158, p. 107038. ISSN: 03784266. DOI: 10.1016/j.jbankfin.2023.107038. URL: https:

//linkinghub.elsevier.com/retrieve/pii/S0378426623002297 (visited on 02/11/2025).

Khwaja, Asim Ijaz and Atif Mian (Sept. 2008). "Tracing the Impact of Bank Liquidity Shocks: Evidence from an Emerging Market". en. In: *American Economic Review* 98.4, pp. 1413–1442.

Kosekova, Kamelia et al. (2024). "Firm-bank relationships: a cross-country comparison". en. In: *CEPR Discussion Paper*. CEPR Press, Paris & London 19464.

Kramarz, F. and O. N. Skans (July 2014). "When Strong Ties are Strong: Networks and Youth Labour Market Entry". en. In: *The Review of Economic Studies* 81.3, pp. 1164–1200. ISSN: 0034-6527, 1467-937X. DOI: 10.1093/restud/rdt049. URL: https://academic.oup.com/restud/article-lookup/doi/10.1093/restud/rdt049 (visited on 10/17/2024).

Lazear, Edward P, Kathryn L Shaw, and Christopher T Stanton (2015). "The Value of Bosses". en. In: *Journal of Labor Economics* 33.4.

Malgieri, Cedomir and Luca Citino (2025). "Wage Contracts and Financial Frictions". en. In.

Minni, Virginia (2025). "Making the invisible hand visible: managers and the allocation of workers to jobs". en. In.

Nguyen, Hoai-Luu Q. (Jan. 2019). "Are Credit Markets Still Local? Evidence from Bank Branch Closings". en. In: *American Economic Journal: Applied Economics* 11.1, pp. 1–32. ISSN: 1945-7782, 1945-7790. DOI: 10.1257/app.20170543. URL: https://pubs.aeaweb.org/doi/10.1257/app.20170543 (visited on 03/17/2023).

Panetta, Fabio, Fabiano Schivardi, and Matthew Shum (June 2009). "Do Mergers Improve Information? Evidence from the Loan Market". en. In: *Journal of Money, Credit and Banking* Vol. 41.No. 4, pp. 673–709.

Paravisini, Daniel, Veronica Rappoport, and Philipp Schnabl (Aug. 2023). "Specialization in Bank Lending: Evidence from Exporting Firms". en. In: *The Journal of Finance* 78.4, pp. 2049–2085. ISSN: 0022-1082, 1540-6261. DOI: 10.1111/jofi.13254. URL: https://onlinelibrary.wiley.com/doi/10.1111/jofi.13254 (visited on 10/24/2023).

Patault, Bérengère and Clémence Lenoir (Oct. 2024). "Customer Capital Spillovers: Evidence from Sales Managers in International Markets". en. In: *American Economic Journal: Applied Economics* 16.4, pp. 404–443. ISSN: 1945-7782, 1945-7790. DOI: 10.1257/app.20220572. URL: https://pubs.aeaweb.org/doi/10.1257/app.20220572 (visited on 02/11/2025).

Petersen, Mitchell A. and Raghuram G. Rajan (Dec. 2002). "Does Distance Still Matter? The Information Revolution in Small Business Lending". en. In: *The Journal of Finance* 57.6, pp. 2533–2570. ISSN: 0022-1082, 1540-6261. DOI: 10.1111/1540-6261.00505. URL: https://onlinelibrary.wiley.com/doi/10.1111/1540-6261.00505 (visited on 01/19/2024).

Philippon, Thomas (Apr. 2015). "Has the US Finance Industry Become Less Efficient? On the Theory and Measurement of Financial Intermediation". en. In: *American Economic Review* 105.4, pp. 1408–1438. ISSN: 0002-8282. DOI: 10.1257/aer.20120578. URL: https://pubs.aeaweb.org/doi/10.1257/aer.20120578 (visited on 10/09/2023).

Philippon, Thomas and Ariell Reshef (Nov. 2012). "Wages and Human Capital in the U.S. Finance Industry: 1909–2006*". en. In: *The Quarterly Journal of Economics* 127.4, pp. 1551–1609. ISSN: 0033-5533, 1531-4650. DOI: 10.1093/qje/qjs030. URL: https://academic.oup.com/qje/article/127/4/1551/1843161 (visited on 10/09/2023).

Rajan, Raghuram and Luigi Zingales (1998). "Financial Dependence and Growth". en. In: *American Economic Review* 88.3, pp. 559–86. DOI: 10.3386/w5758. URL: http://www.nber.org/papers/w5758.pdf (visited on 02/06/2025).

Rodano, Giacomo, Nicolas Serrano-Velarde, and Emanuele Tarantino (Aug. 2018). "Lending Standards over the Credit Cycle". en. In: *The Review of Financial Studies* 31.8, pp. 2943–2982. ISSN: 0893-9454, 1465-7368. DOI: 10.1093/rfs/hhy023. URL: https://academic.oup.com/rfs/article/31/8/2943/4975834 (visited on 03/06/2025).

Stein, Jeremy C. (Oct. 2002). "Information Production and Capital Allocation: Decentralized versus Hierarchical Firms". en. In: *The Journal of Finance* 57.5, pp. 1891–1921. ISSN: 0022-1082, 1540-6261. DOI: 10.1111/0022-1082.00483. URL: https://onlinelibrary.wiley.com/doi/10.1111/0022-1082.00483 (visited on 10/12/2023).

Stiglitz, Joseph E and Andrew Weiss (1981). "Credit Rationing in Markets with Imperfect Information". en. In: *The American Economic Review* 71.3, pp. 393–410.

Sun, Liyang and Sarah Abraham (Dec. 2021). "Estimating dynamic treatment effects in event studies with heterogeneous treatment effects". en. In: *Journal of Econometrics* 225.2, pp. 175–199. ISSN: 03044076. DOI: 10.1016/j.jeconom.

2020.09.006. URL: https://linkinghub.elsevier.com/retrieve/pii/S030440762030378X (visited on 08/23/2023).

Appendix A: Descriptive figures and tables

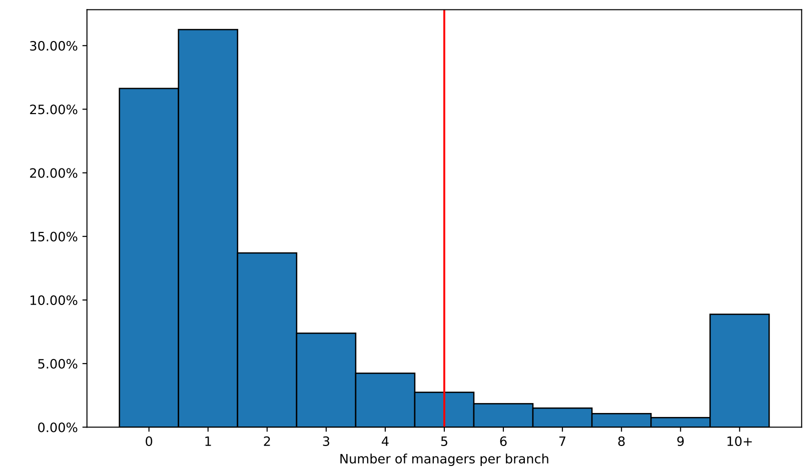


Figure 1.10. Distribution of the number of managers per branch.

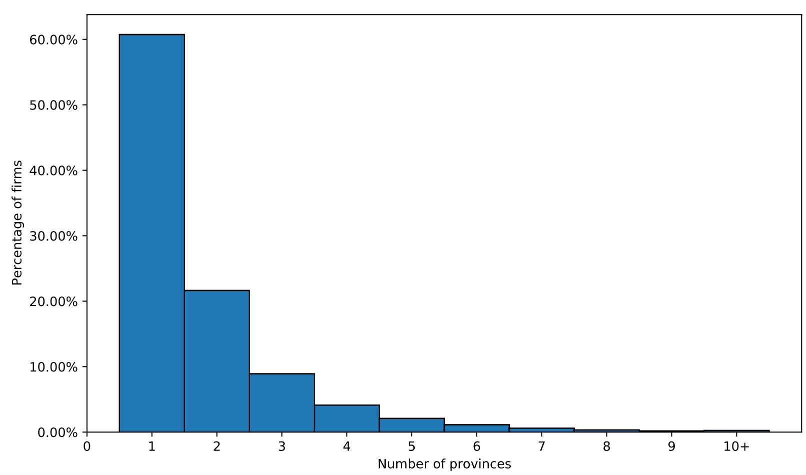


Figure 1.11. Distribution of the number of provinces where firms have credit.

Average number of workers per year		345,938
Move type	Average nr. movers	Percentage
All moves	40,638	11.7 %
only municipality	38,960	95.9 %
only bank group	967	2.4 %
both	712	1.8 %

Table 1.7. Summary of worker movements by type.

Appendix B: Additional results

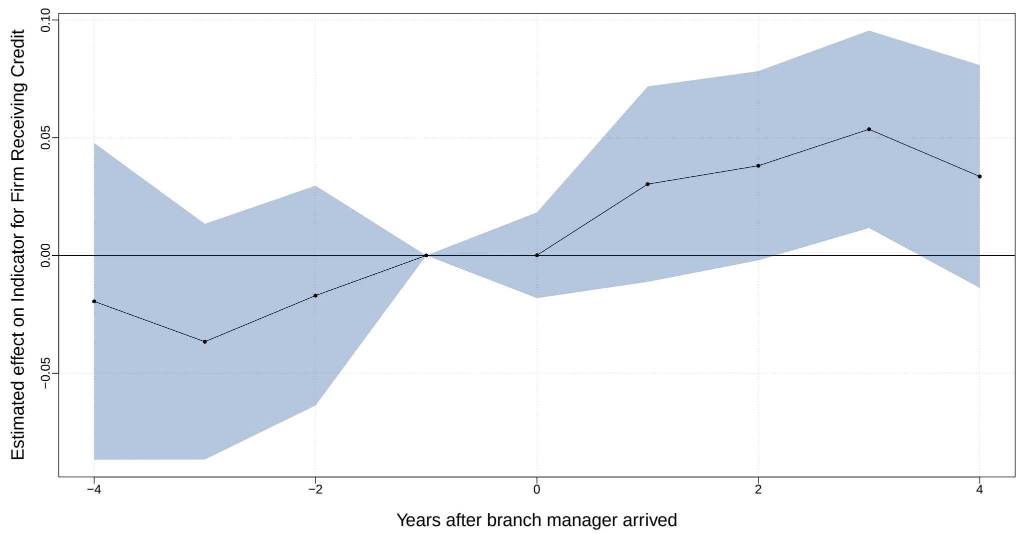
B.1 Robustness checks and additional results

Table 1.8. DiD estimates for β, within the same bank

	Credit indicator			
	(1)	(2)	(3)	(4)
Inflow × Post	0.021***	0.022***	0.021***	0.023***
	(0.006)	(0.005)	(0.006)	(0.005)
R^2	0.802	0.806	0.803	0.807
Observations	23,988,940	23,988,940	23,988,940	23,988,940
Branch-Firm fixed effects	✓	✓	✓	✓
Branch-Time fixed effects		✓		✓
Firm-Time fixed effects			✓	✓

Notes: Estimates for the probability of a firm receiving credit from the branch where its previous loan officer has moved, within the same bank group. Column (1) includes only branch fixed effects, column (2) includes branch and time fixed effects, column (3) includes branch and firm fixed effects, and column (4) includes all fixed effects. Standard errors are clustered at the bank group - firm level. The sample includes all potential firm - branch pairs within a province, for the period 2009-2018. Only firms and branches that are present in the data in all years are included. Moves are defined as relocations of managers within the same bank group. To keep consistency with the sample definition, only moves that occur within the same province are considered. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Figure 1.12. Robustness: Sun and Abraham (2021) estimator, full sample



Notes: Event study estimates for the probability of a firm receiving credit from the branch where its previous loan officer has moved. Plot shows estimates and 95% confidence intervals for the coefficients β_τ in equation (1.2), estimated using the TWFE estimator proposed by Sun and Abraham (2021). Branch - firm, branch - time, and firm - time fixed effects are included. Standard errors are clustered at the bank group - firm level. Sample includes all potential firm - branch pairs, for the period 2009-2018. Branches are defined as municipality - bank group pairs. Only firms and branches that are present in the data in all years are included. Moves are defined as relocations of managers from one bank group to another. To keep consistency with the sample definition, only moves that occur within the same province are considered.

B.2 Heterogeneity by manager age

We then turn to another aspect of the manager's characteristics: her age. There are two contrasting theories on the effect of the manager's age on the portability of her clients. On the one hand, younger managers may have a stronger incentive to build a personal (and thus portable) relationship with the firms, as they may receive career benefits from it. If managerial contracts contain a performance-based component, the manager may have an incentive to build a portfolio of clients that she can bring with her in case she moves. On the other hand, if we consider older managers, they may have had more time to build a relationship with the firms, as well as being more trusted (Gennaioli, Shleifer, and Vishny, 2015). Furthermore, conditional on having a managerial contract, older managers may actually have a more prominent role in the bank, be so more impactful on credit-related decisions, and for this reason being perceived as more valuable by the borrowers. We first test the effect of the manager's age on the portability of her clients, by interacting our treatment with two dummies, one for the manager being older than 40 years, and one for the manager being younger than 40 years. We then try to understand if the manager's age has a differential effect on some firms, depending on their age, size and diversification in the credit market. We define opaque firms as firms that have credit relationships with a single bank, have less than 10 employees, and have been founded in the last 5 years. We measure those characteristics at the time the manager moves. We want to test if a more trusted (or incentivized) manager can be more beneficial to those firms. Table 1.9 shows the results. For each regression we estimate the baseline effect (first row of the table), and then the interaction terms for the manager's age and the characteristics of the firms.

Table 1.9. Heterogeneity by managers' age

	Credit indicator			
	(1)	(2)	(3)	(4)
Inflow × Post	0.061***	0.027***	0.061***	0.028***
	(0.020)	(0.009)	(0.019)	(0.009)
Inflow × Post × Young	-0.030		-0.030	
	(0.024)		(0.023)	
Inflow × Post × Old		0.035		0.034
		(0.021)		(0.021)
Inflow × Post × Opaque			-0.007	-0.018
			(0.074)	(0.011)
Inflow × Post × Young × Opaque			0.008	
			(0.081)	
Inflow × Post × Old × Opaque				0.020
				(0.066)
R^2	0.789	0.789	0.789	0.789
Observations	24,645,380	24,645,380	24,645,380	24,645,380
Branch-Firm fixed effects	✓	✓	✓	✓
Firm-Time fixed effects	✓	✓	✓	✓
Branch-Time fixed effects	✓	✓	✓	✓

Notes: DiD estimates for β in equation 1.1 with heterogeneity by managers' age. Column (1) includes an interaction between β and a dummy for the manager being young (less than 40 years old), column (2) includes an interaction between β and a specular dummy for the manager being old (more than 40 years old). Column (3) and (4) interact the manager's age dummies with firm-level characteristics, such as opacity. A firm is defined as opaque if it is small (less than 10 employees), young (less than 5 years old), and having loans with only one bank. Moves are defined as relocations of managers from one bank group to another. To keep consistency with the sample definition, only moves that occur within the same province are considered. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Columns (1) and (2) show the results for the manager's age. Although noisy, the estimates for the interaction term show that older managers are more likely to bring their clients with them, be it for the experience they have in building relationships, or for the trust they have built with the firms. Columns (3) and (4) show the results for the interaction of the manager's age with the characteristics of the firms. Opaque firms seem to be slightly less willing to follow the manager, but the interaction with the manger's age counterbalances the effect, although not significantly. Older managers seem to be more beneficial for opaque firms. Those firms have weaker connections in the credit market, and have weaker fundamentals, so soft information should matter more for them. What our results seem to suggest is that a manager that has more "leverage" in bringing her clients with her (as proxied by the manager's age) is more beneficial for those firms for which getting additional credit would be harder (not because of lower credit score, but because of less information available to the bank).

Chapter 2

Scientific resilience: How Italian nuclear physics changed after the Chernobyl disaster

Abstract

Scientists' human capital is the main factor in the production of knowledge. I study how flexible field-specific human capital is, trying to understand if researchers can bring valuable contributions to innovation out of their main field of studies. I focus on the careers of Italian nuclear scientists before and after the Chernobyl disaster of 1986. In 1987 in Italy a referendum stopped the production of nuclear energy, and strongly reduced fundings to research in that field. Using data from Microsoft Academic Graph, I show that after Chernobyl the amount of Italian papers published in nuclear fission decreased by 50%. Researchers who had already published in fission experienced a reduction of 24% in their citations, and 7% in published papers 15 years after the shock. Compared to other physicists, they neither moved more frequently, nor contributed permanently to more new fields.

2.1 Introduction

There is no economic growth without innovation (Romer, 1990; Aghion and Howitt, 1992; Akcigit, Pearce, and Prato, 2020). And there is no innovation without investments in scientific human capital (Barro, 2001; Prato, 2022). However, economists have not fully understood how human capital helps producing new technologies. Capital goods can only perform the limited range of tasks they are designed for. One cannot use an ax to plow or a hoe to cut wood. On the other hand, human capital is generally assumed to be more flexible, although labor economists have debated about its depreciation (Deming and Noray, 2020), specificity with respect to the firm (G. S. Becker, 1962; Lazear, 2009), to the team (Chen, 2021) or to the technology (Marx, Strumsky, and Fleming, 2009).

In this paper I study how flexible scientific human capital can be. It is flexible when researchers are able to take some scientific principles and apply them out of the fields in which they were discovered. This is the story of many breakthrough innovations, such as penicilline or the X-rays, which are the successful - but unintended - results of projects started with different aims (Nelson, 1959). Yet, there is limited evidence on the transition of scientists to new fields, and their contribution to innovation. One would need to observe a population of scientists who - at some point - were unable to publish in their main field of studies, and had to relocate elsewhere.

In this article I use a rich database provided by Microsoft Academic Graph, which records information on more than 250 billion scientific publications, to study the case of Italian nuclear scientists. A year after the Chernobyl disaster of April 1986, Italy held a referendum that stopped the production of nuclear energy, and rapidly de-funded research in nuclear fission. By its nature, the shock affected only a single sector and a single country. Indeed, every other country involved in a nuclear program neither made such a political decision nor experienced a reduction in the number of publications in nuclear fission.

I document three main findings. First, after Chernobyl Italian nuclear fission lost more than 50% of its potential papers. Second, not only newcomers were kept out of the field, but also incumbent researchers experienced slowdowns in their careers. In particular, I find that authors who had published in fission pre-1986 faced a reduction of 10% in papers produced, and spoke to a narrower audience, being their citations reduced by 25%. Third, Italian fission scientists mainly stayed in Italy, and did not contribute to new fields more than other scientists.

We already know that human capital takes longer to build, and brings much more long-run value to the research output of university departments than physical capital (Waldinger, 2016). From the findings here presented I argue that scientific human capital has limited flexibility. Scientists' expertise is relatively narrow, as there are significant "losses in translation" in transferring knowledge to new fields.

I contribute to three strands of literature.

Within the field of Economics of Science, the paper that is the closest to mine is K. Myers (2020), that tries to identify the "elasticity of science". He measures how responsive are biomedical scientists in changing the topics of their research when the National Institutes of Health issue a new grant. However, he only focuses on temporary changes in research topics; indeed, NIH-funded scientists are not required to shift their whole research agenda after receiving the grant. In this paper's setting Italian nuclear scientists cannot go back to nuclear research unless they emigrate (which has not been the case): their field of studies has been permanently shut down.

My work also relates to other articles in the field of Economics of Science, that have analysed collaboration, networks and competition in scientific careers. Collaboration among scientists advances the frontier of knowledge (Iaria, Schwarz, and Waldinger, 2018; Jia et al., 2022), as they generate spillovers among their peers (Waldinger, 2010; Ductor et al., 2014) and experienced researchers promote the careers of PhD students (Waldinger, 2012). On the other hand, more productive scientists can also slow down their competitors' careers (Borjas and Doran, 2012) or act as gate keepers (Azoulay, Zivin, and Wang, 2010), steering research towards their interests and blocking unexplored avenues. Such papers have often exploited migration shocks, such as the Jewish diaspora from Nazi Germany (Moser, Voena, and Waldinger, 2014; S. O. Becker et al., 2021), the Russian outflow after the collapse of the Soviet Union (Ganguli, 2015; Borjas and Doran, 2015) or US migration policies (Agarwal et al., 2021; Prato, 2022). Compared to previous literature, I consider scientists who switched field, instead of moving abroad. I show that Italian fission scientists became less relevant and did not impact the direction of future research.

Second, this paper speaks to a branch of literature in Labor Economics. Starting from the seminal work by G. S. Becker (1962), labor economists have studied how and why human capital can exhibit various degrees of specificity. In some sense, workers are locked to their technology (Neal, 1995), to the task they perform (Gathmann, 2010), to their network (Jäger, 2022) or their education (Aghion, Akcigit, et al., 2022). In all those cases, workers' mobility might hurt firm growth, because it would imply some loss of their specific human capital.

A particular case in which human capital specificity can limit firm growth is the case of non-compete agreements, as shown in Marx, Strumsky, and Fleming (2009) and Marx (2011) and Marx (2022). In some industries knowledge workers are not allowed to move to competing firms. They are forced to change industry, thus losing a bigger part of their human capital. On the other hand, as Arts and Fleming

(2018) argue, those workers may be tempted to explore new areas bringing their past knowledge, which in some cases can lead to innovative activities.

In this paper I consider a particular case of high-skilled workers - i.e. academics - who are forced to exit their field. Results of my analysis show that they did not innovate significantly, but rather they lost human capital.

I also contribute to a third, growing, strand of literature that studies innovation spillovers of large-scale public projects, in defence (Moretti, Steinwender, and Van Reenen, 2019; Bhattacharya, 2021), pharmaceutics (Azoulay, Graff Zivin, et al., 2019), energy (K. R. Myers and Lanahan, 2022), or with state nationalisation programs (Akcigit, Hanley, and Serrano-Velarde, 2021). Recent works have shown positive local spillovers from two large-scale projects conducted by the US government, the Space Race (Kantor and Whalley, 2022) and the Office of Scientific Research and Development, deployed during WWII (Gross and Sampat, 2020a; Gross and Sampat, 2020b), which also lead to the development of nuclear reactors. I argue that the Italian investment in nuclear fission generated no spillovers in contiguous fields.

The rest of the paper proceeds as follows. Section 2.2 provides historical background on nuclear research and the Chernobyl disaster. In Section 2.3, I describe the construction of the sample of Italian fission scientists. In Section 2.4, I document the aggregate impact of Chernobyl on nuclear fission publications. I report results on individual careers in Section 2.5. Section 2.6 concludes.

2.2 Historical context

During World War II, the Allied countries, and in particular the US, invested in developing a new technology: nuclear fission. While its first scopes were military (the bombs on Hiroshima and Nagasaki), scientists in the Manhattan project already saw its potential as an energy source (Gross and Sampat, 2020b). Research in nuclear physics rapidly grew to be one of the most intriguing scientific fields in the second half of the 20th century.

In the mid 50s, the UK, the Soviet Union and the US built their first nuclear fission power plants. New generations of power plants were built through the 60s, as the number of adopting countries increased (including Italy). When in the 70s the oil crisis caused energy prices to ramp up, most of Western countries decided to expand their nuclear program. Italy already had 4 power plants in its territory, and planned to build 10 more by the end of the 80s.[1]

At the end of 1985 nuclear fission accounted for $5,60\%$ of the world production of energy.[2] The night of April 26th 1986, the fourth reactor of the Chernobyl power plant exploded. Toxic fumes spread all over Europe. Many people got worried and joined anti-nuclear movements, urging governments to stop nuclear programs.[3]

In the first months of 1987, the Italian Radical Party (Partito Radicale) issued a referendum on the use of nuclear power. On Nov. 9th 1987 65% of Italian electorate participated, and with a majority of 80% decided to put an end to the Italian nuclear program.[4]

In a matter of a few months, the Italian authority on nuclear research - ENEA - had to dramatically revise its plans. The composition of its budget rapidly changed. By 1989 all projects on nuclear fission were de-funded. Professors who worked in the nuclear sector had to relocate, either by changing research topics, or by moving abroad.

In Italy, public research on nuclear energy was conducted by universities and, most importantly, by ENEA. Originally founded as a solely-nuclear institution, in 1982 it was reformed in order to include research on renewable sources (mainly solar power). However, before the Chernobyl disaster, the greatest part of its research activity was still in the nuclear sector.

As it can be seen in Figure **??**, in 1985 nuclear fission accounted for 70% of the ENEA budget. Starting in 1986, fundings to projects in nuclear fission were rapidly reduced. In 1989 the share of fundings to nuclear fission dropped to 20% (no new project was financed; money spent on nuclear fission was for maintenance costs).

On the other hand, the overall budget of ENEA neither increased after 1986, nor dropped dramatically. As it can be seen in Figure **??**, the budget reached a peak of around 1020 billion liras[5] in 1985, then it slightly decreased in the following years. Following the red line, it is possible to see that the reduction of fundings to nuclear fission was dramatic both in relative and absolute terms.

[1] For detailed information on Italian nuclear expansion programs, see Appendix C.

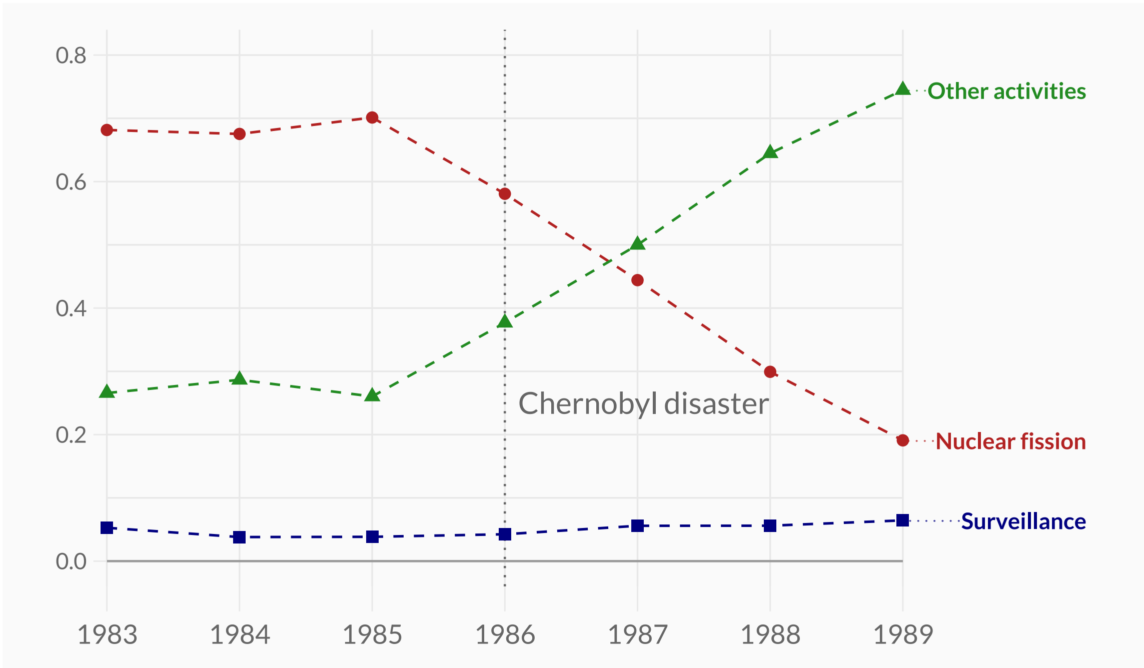
[2] See Hannah Ritchie and Rosado (2020).

[3] See Bini and Londero (2017).

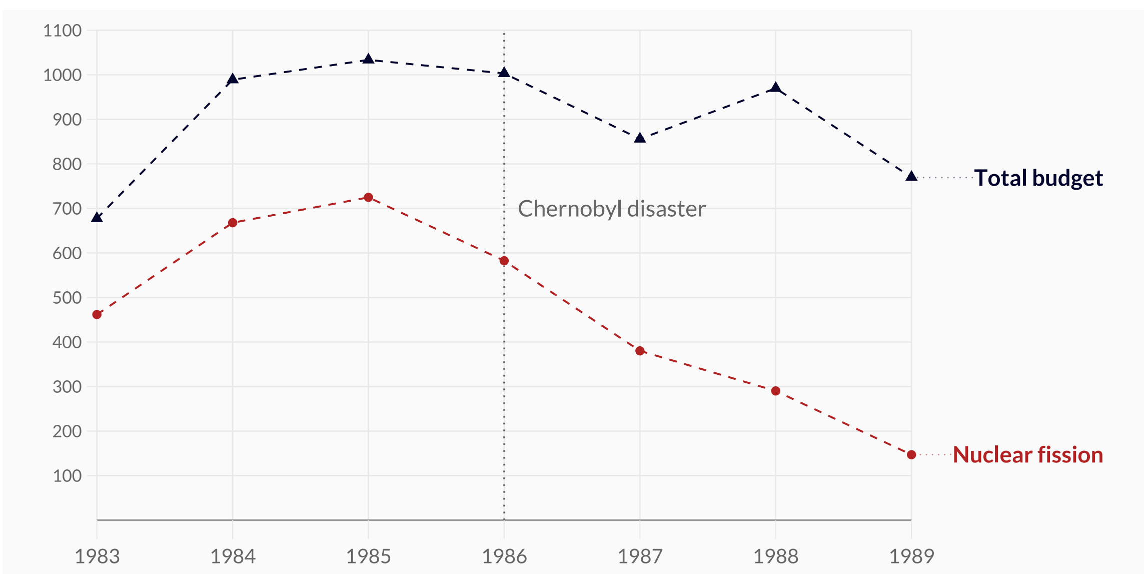
[4] Italians had to vote on three issues. First, they were asked to repeal a national interest law that allowed the government to overrule local municipalities on the permissions to build a nuclear power plant. Second, they repealed a law that gave municipalities money transfers if they hosted nuclear power plants on their territory. The third vote prohibited Enel - the National Institution for Electric Energy - to build power plants in foreign countries.

[5] Equivalent to 1,31 billion Euros in 2022.

Figure 2.1. ENEA budget



(a) Relative shares



(b) Absolute figures

Time series of the three expenditure destinations of the ENEA budget in the years 1983–1989. As it may be noticed, right after the Chernobyl disaster, the sector of nuclear fission lost its prominence in the ENEA budget, dropping from 70% of the budget in 1985 to less than 20% in 1989.

After the 1987 referendum, investment on new nuclear plants was completely stopped. Still, some maintenance costs on ongoing projects had to be sustained, so figures do not completely drop to zero.

2.3 Data

The main data source for this article is Microsoft Academic Graph (MAG), a freely-accessible online database on scientific publications. Each entry of the database corresponds to a paper. Every paper is associated to several pieces of information, such as authors (and their affiliations at the time), date of publication, field(s) of studies it belonged to and list of references.[6]

I select 16 fields of studies related to nuclear fission.[7] Papers associated to at least one of them are labelled as "fission papers". I identify Italian authors as being affiliated to universities whose ISO code was "IT" (MAG provides this information for all affiliations).

2.3.1 Summary statistics

I present some summary statistics about nuclear fission in Italy and in other developed countries.

First, I show that Italy was not on a frontier country in nuclear fission. As the first row of Table 2.1 shows, in years 1980-1995 in Italy only 67 papers belonging to nuclear fission have been published, by a total number of 126 authors. The country which contributed the most to nuclear research was the US, both in terms of authors and papers. The US led also in number of active institutions (an institution is labelled as active if at least a paper whose authors were affiliated to that institution was published in the considered timespan). Furthermore, in many countries private companies[8] produced around 50% of research in nuclear fission. In Italy the figure was far lower (18%), meaning that a reduction in public funds to nuclear research could have a big impact on the scientific sector.

Second, I show that patenting activity did not rely that much on nuclear research. Using an extension to Microsoft Academic Graph (Marx and Fuegi, 2022), in Table 2.2 I show that very few (only 3% worldwide) of papers published in nuclear fission were eventually cited in patents. In Italy just a single paper was cited in patents, suggesting that there was little technological transfer from academic research to innovation activity.

Third, I present data on some internal features of Italian fission research. In Table

[6] The project has been last updated in 2021. This article uses the last version of the database, downloaded on 2021-09-13.

[7] Coding methodology and definition of nuclear fission are explained in Appendix D.

[8] See Appendix D for the proper definition.

Table 2.1. Academic production in nuclear fission

Country	Authors	Papers	Institutions	Private inst.	Private papers
IT	126	67	17	3	12 (18%)
FR	272	158	23	1	86 (54%)
US	3823	2998	278	56	1364 (45%)
JP	1053	594	65	13	157 (26%)
DE	539	335	35	6	200 (60%)
GB	297	194	56	8	35 (18%)
RU	137	60	8	0	0 (0%)

The table reports the number of authors and research institutions that operated in nuclear fission. Years 1980-1995 are considered.
It also reports the number of institutions that were privately funded (i.e., private companies research labs), and the amount of papers that were published by those institutions.
Source: Microsoft Academic Graph database

2.3 I show that after Chernobyl, Italian research became less relevant worldwide. Italian papers were cited by more articles that had all-Italian authors (moving from 14% to 19%). The share of Italian authors in citing papers increased from 29 to 42 percent, and Italian papers got more intensively cited by papers in the same field rather than out-of-field papers. This preliminary evidence may suggest that the shock restricted the audience Italian scientists spoke to.

2.4 Aggregate impact on scientific output

In this section I show that after Chernobyl, Italian research in nuclear fission did not keep pace with other fields of studies.

Figure 2.2 plots the yearly number of publications in two fields of studies: nuclear fission and semiclassical physics. Both fields belong to the general realm of physics, but they are considerably different. Research in nuclear fission is applied, and aims to improve the efficiency of energy production. On the other hand, semiclassical physics mostly deals with theoretical research. Therefore neither the Chernobyl disaster nor the subsequent referendum should have any impact on it.

As it can be seen in the figure, both fields were expanding before 1986. Then, after Chernobyl, the two fields started to diverge. Yearly publications in semiclassical physics continued to grow, while publications in nuclear fission reached their 1985 level only in 1993.

I test whether the Chernobyl disaster has slowed down Italian research in nuclear

Table 2.2. Papers cited in patents

Country	Papers	Cited in patents	Produced by privates
IT	67	1	0
FR	158	3	1 (33%)
US	2998	92	28 (30%)
JP	594	22	10 (45%)
DE	335	4	2 (50%)
GB	194	12	6 (50%)
RU	60	0	0

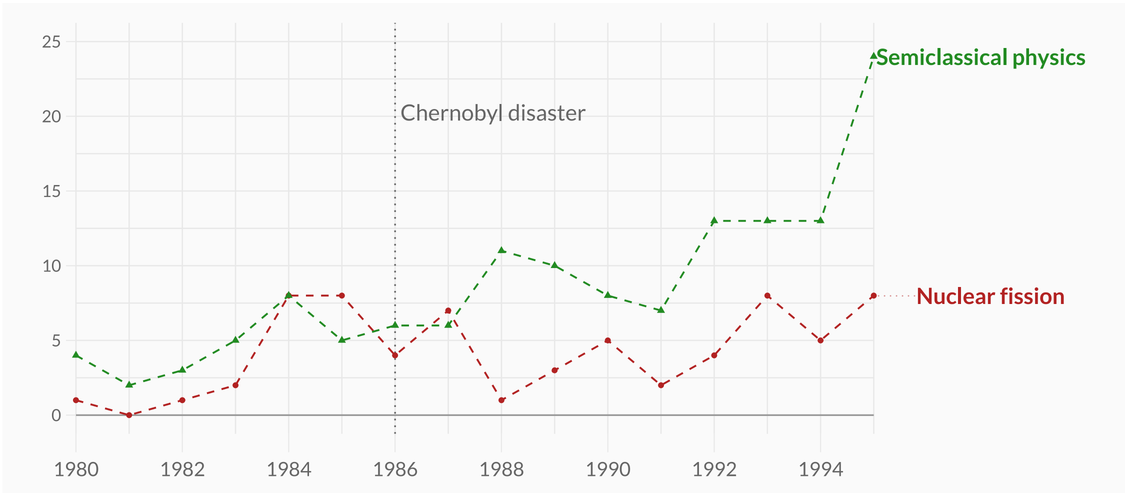
The table reports the number of papers produced in nuclear fission, in every country. It provides also the number of papers that were cited in patents, and the number of papers that - among the ones cited in patents - were produced by private research institutions.
Years 1980-1995 are considered.
Data source: Marx and Fuegi (2022)

Table 2.3. Links to other countries and fields

Period	Published Papers	Citing papers			
		Nr	All-Italian	Share of Italian authors	In-field
1980-1985	20	134	21 (16%)	29 %	16 (12%)
1986-1995	47	283	53 (19%)	42 %	69 (24%)

The table reports information on the publications of Italian nuclear scientists. It reports the amount of published papers in nuclear fission and the amount of papers that cite them. Among the citing papers, it reports how many were all-Italian (i.e., written only by authors affiliated with Italian universities), what was the average share of Italian authors per paper, and how many of the citing papers belonged to the field itself.
Source: Microsoft Academic Graph database

Figure 2.2. Time series of scientific publications



Yearly number of academic publications per scientific field. The red line represents nuclear fission. The green line is semiclassical physics.
Apparently, the two fields were following a parallel trend before the Chernobyl disaster.
Source: Microsoft Academic Graph database

fission using a Difference- in-Difference model, as

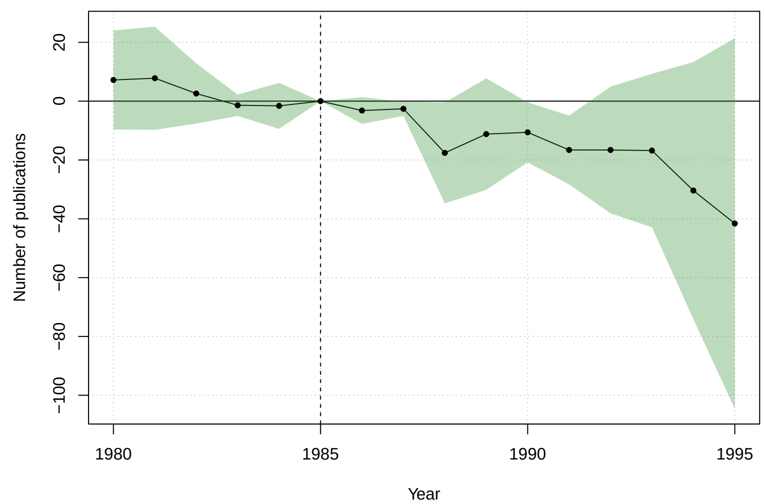
$$y_{ft} = \alpha + \beta Treat_f + \gamma Post_t + \delta Post_t \times Treat_f + \theta_t + \epsilon_{ft} \tag{2.1}$$

Where y_{ft} is the yearly number of publications in each field (or a function of it), $Treat_f$ is a dummy variable for the treated field (fission), $Post_t$ is a dummy variable being equal to 1 if the observation relates to year 1986 or later, $Post_t \times Treat_f$ is the interaction term in the DiD and θ_t controls for year-specific fixed effects.

Figure 2.3 shows estimates and 95% confidence intervals of an event study regression, in which the control group is made not only of semiclassical physics, but also of medical physics, engineering physics, theoretical physics and quantum mechanics. I chose this set of fields because all of them have - as semiclassical physics - little relation to the production of energy. As it can be seen in Figure 2.3, prior to 1986 nuclear fission was following the same trend of all other fields. After Chernobyl, the number of publications started to reduce (in particular, in periods 2, 4 and 5, i.e. years 1988, 1990 and 1991).

I estimate Equation 2.1, with two separate sets of controls. In the first specification I only use semiclassical physics as control group. In the second I use a synthetic control, created following Abadie, Diamond, and Hainmueller (2010), by using all fields considered when producing Figure 2.3. In both specifications I consider publications

Figure 2.3. Effects of Chernobyl disaster onto the number of publications



Event study plot. Fission is treated, semiclassical physics, medical physics, engineering physics, theoretical physics and quantum mechanics are used as controls. Period 0 is 1986.
Source: Microsoft Academic Graph database

Table 2.4. Effect of Chernobyl on aggregate publications

	Nr. publications		Log nr. publications		Asinh nr. publications	
	(1)	(2)	(3)	(4)	(5)	(6)
$Post \times Treat$	-5.233	-7.109	-0.2740	-0.5613	-0.2326	-0.6043
	(2.099)	(1.579)	(0.283)	(0.259)	(0.351)	(0.322)
R^2	0.84069	0.92046	0.85928	0.90551	0.85391	0.90157
Observations	32	32	32	32	32	32
Field F.E.	✓	✓	✓	✓	✓	✓
Year F.E.	✓	✓	✓	✓	✓	✓

Years 1980-1995 are considered.
Cols. (1), (3) and (5) report the estimates for δ in a DiD regression in which semi-classical physics is used as a control.
Cols. (2), (4) and (6) report the estimates for δ in a DiD regression with a synthetic control in the spirit of Abadie, Diamond, and Hainmueller (2010). It consists of a weighted mean of the following fields: semiclassical physics, medical physics, engineering physics, theoretical physics and quantum mechanics.
In (1) and (2) the dependent variable is the yearly number of publications in a scientific field, in (3) and (4) the dependent variable is the log number of publications, in (5) and (6) it is the inverse hyperbolic sine of the number of publications.
All regressions feature the diff-in-diff term δ, a post-treatment dummy $Post1986$, a dummy for the treated field, and year fixed effects.

in the years 1980-1995.

Table 2.4 reports the estimates. Every pair of columns is associated to a different outcome. In column (1) and (2) the outcome is just yearly number of publications. In column (1) only semiclassical physics is in the control group, while column (2) uses the synthetic control.

The pattern is repeated for the subsequent pairs , that differ only for the outcome variable. I perform two standard transformations for count data; in column (3) and (4) the yearly number of publications is log-transformed, while in column (5) and (6) I apply the inverse hyperbolic sine.

In columns (1) and (2) both simple and synthetic control show that after Chernobyl yearly publications significantly decreased. In particular, synthetic control estimates suggest that nuclear fission lost more than 7 articles per year. Turning absolute figures in percentage points, estimates in column (4) indicate that yearly publications decreased by more than 50%.

Appendix A contains also some robustness exercises. I performed the same exercise for other countries, such as the US, Japan and France. Estimates are reported in tables 2.12 - 2.13. It appears that some countries even increased their publications in fission (arguably because they decided to invest more on the safety of nuclear power plants). Italy appears the only that faced a significant reduction in the number of nuclear fission publications.

2.5 Impact of Chernobyl on individual careers

I identify 42 Italian authors that had at least a publication in nuclear fission between 1972 and 1986. Only 24 of them went on publishing in 1986 or later, and only 10 published papers in nuclear fission post-Chernobyl. Out of the 24 that produced publications after the shock, only 2 of them changed affiliation.

In order to understand to what extent those authors were damaged, I focus on two measures: number of publications and number of citations. Equation 2.2 presents a formal test for this hypothesis.

$$y_{it} = \alpha_i + \beta Treat_i + \gamma Post_t + \delta Post_t \times Treat_i + \theta_t + \epsilon_{ft} \tag{2.2}$$

Where y_{it} is the yearly number of publications (or citations) per author, α_i is an author-specific fixed effect, $Treat_i$ is a dummy variable that is equal to 1 if the author had published in fission pre-1986, $Post_t$ is a dummy variable being equal to 1 if the observation relates to year 1986 or later, $Post_t \times Treat_i$ is the interaction term in the DiD, and θ_t controls for year-specific fixed effects.

As a control group, I consider all authors who published at least once in semiclassical physics, medical physics, engineering physics, theoretical physics and quantum mechanics in the years 1972-1986. Figure 2.7 in Appendix B shows estimates and confidence interval for an event study regression. Pre-treatment trends are not clearly visible, although there is arguably a slight decrease in publications per author before 1986.

One can think that younger authors can relocate more easily in different fields, while more experienced authors have a harder time in changing the object of their studies. However, given that I introduce author-specific fixed effects, time-invariant features, such as age in 1986 (which can be a proxy for experience at the time of the shock) are already captured. In this design, estimates may be biased only by omitted time-varying variables that affect each author's productivity (such as health, for instance).

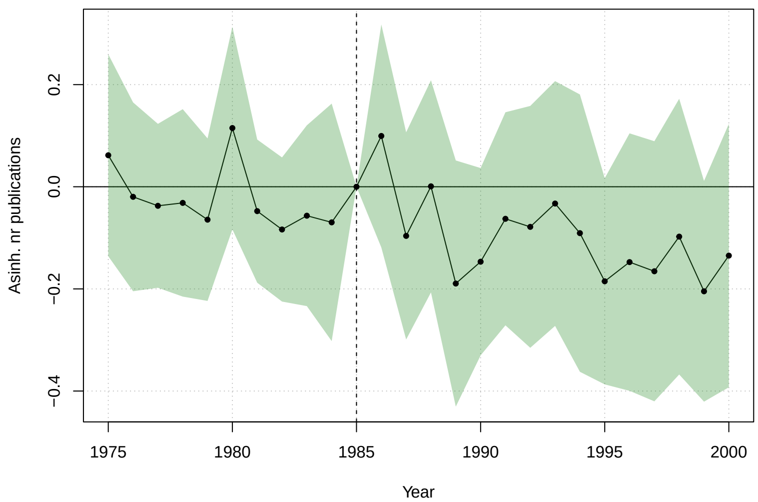
2.5.1 Publications per author

Table 2.5. Effect of Chernobyl on individual publications

	Nr. publications (1)	Log nr. publications (2)	Asinh nr. publications (3)
$Post \times Treat$	-0.4460	-0.0661	-0.0809
	(0.1643)	(0.0554)	(0.0713)
R^2	0.50255	0.56955	0.56973
Observations	50,440	50,440	50,440
Author F.E.	✓	✓	✓
Year F.E.	✓	✓	✓

Years 1980-2000 are considered.
Every observation is an author-year pair.
Authors who published in nuclear fission before 1986 are considered as treated.
The control group is made of authors who published before 1986 in semiclassical physics, medical physics, engineering physics, theoretical physics and quantum mechanics.
In (1) the dependent variable is the number of publications per year.
In (2) the dependent variable is the log number of publications per year.
In (3) the dependent variable is the inverse hyperbolic sine of the number of publications per year.
Every regression features year and author fixed effects, plus the DiD term $Post \times Treat$.
Standard errors are clustered by author.

Figure 2.4. Individual publications drop by 7% after Chernobyl



Event study plot. Effects of Chernobyl disaster onto the inverse hyperbolic sine of the number of publications of Italian scientists. Scientists who published at least once in 1972-1986 are part of the sample. Scientists who published at least once in fission (pre-Chernobyl) are considered as treated. Scientists who published at least once (pre-Chernobyl) in semiclassical physics, medical physics, engineering physics, theoretical physics and quantum mechanics are used as controls. Standard errors are clustered at the scientist level. Timespan considered for the event study: 1975-2000. Source: Microsoft Academic Graph database

Table 2.5 reports estimates for Equation 2.2, where the outcome variable is number of publications per year. Estimates in column (1) show that authors who were active in fission before 1986 faced a reduction in productivity of almost a paper per year.

In columns (2) and (3) dependent variables are the log and the inverse hyperbolic sine of yearly number of publications.The DiD estimates suggest that productivity was reduced of almost 10%. However, estimates are significant only at 10% level, with errors clustered at author level. Arguably, with a higher level of clustering (e.g. at field level), estimates would have been more precise. However, there is no unique way to map authors into fields (since authors contributed to several fields), therefore I decided to stick to this more conservative clustering level.

Table 2.6. Effect of Chernobyl on individual citations

	Nr. citations (1)	Log nr. citations (2)	Asinh nr. citations (3)
$Post \times Treat$	-19.85	-0.2348	-0.2626
	(3.457)	(0.1316)	(0.1547)
R^2	0.27507	0.53915	0.53835
Observations	50,440	50,440	50,440
Author F.E.	✓	✓	✓
Year F.E.	✓	✓	✓

Years 1980-2000 are considered.
Every observation is an author-year pair.
Authors who published in nuclear fission before 1986 are considered as treated.
The control group is made of authors who published before 1986 in semiclassical physics, medical physics, engineering physics, theoretical physics and quantum mechanics.
In (1) the dependent variable is the number of citations per year.
In (2) the dependent variable is the log number of citations per year.
In (3) the dependent variable is the inverse hyperbolic sine of the number of citations per year.
Every regression features year and author fixed effects, plus the DiD term $Post \times Treat$.
Standard errors are clustered by author.

2.5.2 Citations per author

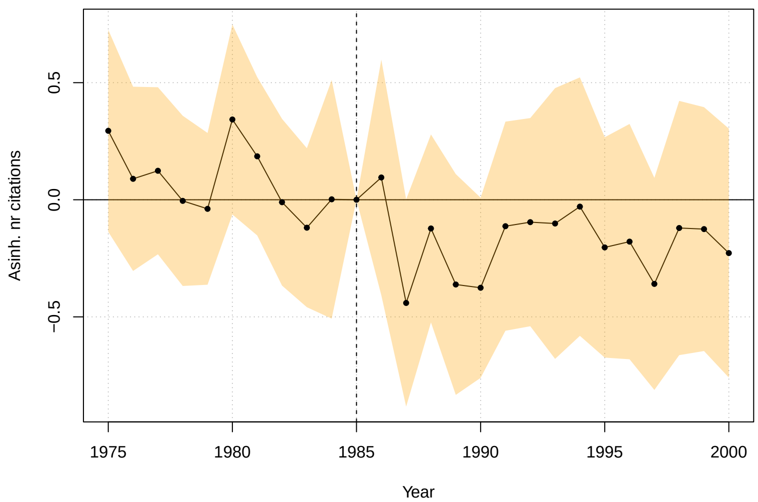
The second variable of interest is number of citations. Treated and control group are exactly the same as for publications.

Figure 2.5 shows estimates and confidence intervals for an event study regression. There is no evidence of pre-treatment trends, and the number of publications for fission scientists starts decreasing in the mid-90s.

Table 2.6 reports estimates for Equation 2.2, where the dependent variable is yearly number of citations. Authors who had published in nuclear fission pre Chernobyl lost more than 50 citations per year with respect to their peers (column 1), which accounted for around 25% of their citations (column 2). Not only nuclear fission scientists published fewer papers (10% reduction), but also they became less relevant, speaking to a smaller audience.

So far, I considered the whole career of nuclear fission scientists, taking into account papers published up to 2020. In Appendix A I perform some robustness exercises, considering shorter time periods. Still using 1980 as initial year, I use alternatively 2000 and 2010 as final years. In both cases all point estimates exhibit a negative sign. However, in regressions that use 2000 as a final year point estimates

Figure 2.5. Individual citations drop by 24% after Chernobyl



Event study plot. Effects of Chernobyl disaster onto the inverse hyperbolic sine of the number of citations of Italian scientists. Scientists who published at least once in 1972-1986 are part of the sample. Scientists who published at least once in fission (pre-Chernobyl) are considered as treated. Scientists who published at least once (pre-Chernobyl) in semiclassical physics, medical physics, engineering physics, theoretical physics and quantum mechanics are used as controls. Standard errors are clustered at the scientist level. Timespan considered for the event study: 1975-2000.
Source: Microsoft Academic Graph database

are too small to be significant.

On the other hand, estimates with the 2010 sample are almost identical to the full sample. This is reassuring, because in Figure 2.7 estimates for years from 2012 on are quite noisy. This implies that the reduction in citations per author is not driven by later years.

2.6 Conclusions

In this article I argued that Italian nuclear scientists have not been able to successfully relocate after the defunding of their field. They were not able to open the gates of new areas of research, and their careers slowed down. This piece of evidence informs the debate on science funding, showing that - in the absence of pre-designed programs for technological transfer - scientists are not able to smoothly relocate across fields. What is more, their contribution to areas of research they did not initially work on may be less valuable than the work in their original domain.

Further research should quantify the welfare implications of the defunding of a national research area, both at a country-level (i.e. how much research in nuclear fission mattered for Italy) and at a field level (how much Italian scientists mattered for nuclear fission worldwide).

However, this article has two major shortcomings: first, the number of nuclear fission authors is quite small (42 publishing in Italy before 1986). Second, I was not able to access granular budget data and compute any elasticity of academic production with respect to fundings.

Later versions of this article should address those shortcomings, as well as expand the analysis with machine learning techniques. For instance, I did not define any contiguity measure of research areas (i.e. mathematics is relatively closed to physics, while it is far apart from classic literature). Using the tools of network analysis, it would be possible to study the transition to new fields in a richer setting.

Another development of the article would focus on the innovative content of every article. Running text analysis algorithm (both on titles and abstracts), I would be able to estimate whether fission scientists who transitioned to new fields were able to produce more innovative content.

References

Abadie, Alberto, Alexis Diamond, and Jens Hainmueller (June 2010). "Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program". en. In: *Journal of the American Statistical Association* 105.490, pp. 493–505. ISSN: 0162-1459, 1537-274X. DOI: 10.1198/jasa.2009.ap08746. URL: http://www.tandfonline.com/doi/abs/10.1198/jasa.2009.ap08746 (visited on 11/10/2022).

Agarwal, Ruchir et al. (2021). "Why U.S. Immigration Barriers Matter for the Global Advancement of Science". en. In: *SSRN Electronic Journal*. ISSN: 1556-5068. DOI: 10.2139/ssrn.3762886. URL: https://www.ssrn.com/abstract=3762886 (visited on 11/09/2022).

Aghion, Philippe, Ufuk Akcigit, et al. (2022). *A Year Older, A Year Wiser (and Farther from Frontier): Invention Rents and Human Capital Depreciation*. Tech. rep. National Bureau of Economic Research.

Aghion, Philippe and Peter Howitt (1992). "A Model of Growth Through Creative Destruction". In: *Econometrica* 60.2. Publisher: Wiley, Econometric Society, pp. 323–351.

Akcigit, Ufuk, Douglas Hanley, and Nicolas Serrano-Velarde (2021). "Back to basics: Basic research spillovers, innovation policy, and growth". In: *The Review of Economic Studies* 88.1. Publisher: Oxford University Press, pp. 1–43.

Akcigit, Ufuk, Jeremy G Pearce, and Marta Prato (2020). "Tapping into Talent: Coupling Education and Innovation Policies for Economic Growth". en. In: p. 69.

Arts, Sam and Lee Fleming (Dec. 2018). "Paradise of Novelty—Or Loss of Human Capital? Exploring New Fields and Inventive Output". en. In: *Organization Science* 29.6, pp. 1074–1092. ISSN: 1047-7039, 1526-5455. DOI: 10.1287/orsc.2018.1216. URL: http://pubsonline.informs.org/doi/10.1287/orsc.2018.1216 (visited on 12/13/2022).

Azoulay, Pierre, Joshua S Graff Zivin, et al. (2019). "Public R&D investments and private-sector patenting: evidence from NIH funding rules". In: *The Review of economic studies* 86.1. Publisher: Oxford University Press, pp. 117–152.

Azoulay, Pierre, Joshua S Graff Zivin, and Jialan Wang (2010). "Superstar Extinction". en. In: *The Quarterly Journal of Economics* 125.2, pp. 549–589.

Barro, Robert J (2001). "Human Capital and Growth". en. In: *HUMAN CAPITAL* 91.2, p. 6.

Becker, Gary S. (Oct. 1962). "Investment in Human Capital: A Theoretical Analysis". en. In: *Journal of Political Economy* 70.5, Part 2, pp. 9–49. ISSN: 0022-3808, 1537-534X. DOI: 10.1086/258724. URL: https://www.journals.uchicago.edu/doi/10.1086/258724 (visited on 12/16/2022).

Becker, Sascha O. et al. (Nov. 2021). *Scholars at Risk: Academic Networks and High-Skilled Emigration from Nazi Germany*. CReAM Discussion Paper Series 2105. Centre for Research and Analysis of Migration (CReAM), Department of Economics, University College London.

Bhattacharya, Vivek (2021). "An empirical model of R&D procurement contests: An analysis of the DOD SBIR program". In: *Econometrica* 89.5. Publisher: Wiley Online Library, pp. 2189–2224.

Bini, Elisabetta and Igor Londero (2017). *Nuclear Italy. An international history of Italian nuclear policies during the Cold War*. EUT Edizioni Università di Trieste.

Borjas, George J and Kirk B Doran (2012). "The collapse of the Soviet Union and the productivity of American mathematicians". In: *The Quarterly Journal of Economics* 127.3. Publisher: MIT Press, pp. 1143–1203.

— (2015). "Cognitive Mobility: Labor Market Responses to Supply Shocks in the Space of Ideas". en. In.

Chen, Yiqun (Dec. 2021). "Team-Specific Human Capital and Team Performance: Evidence from Doctors". en. In: *American Economic Review* 111.12, pp. 3923–3962. ISSN: 0002-8282. DOI: 10.1257/aer.20201238. URL: https://pubs.aeaweb.org/doi/10.1257/aer.20201238 (visited on 11/08/2022).

Deming, David J and Kadeem Noray (Nov. 2020). "Earnings Dynamics, Changing Job Skills, and STEM Careers*". en. In: *The Quarterly Journal of Economics* 135.4, pp. 1965–2005. ISSN: 0033-5533, 1531-4650. DOI: 10.1093/qje/qjaa021. URL: https://academic.oup.com/qje/article/135/4/1965/5858010 (visited on 11/08/2022).

Ductor, Lorenzo et al. (2014). "Social networks and research output". In: *Review of Economics and Statistics* 96.5. Publisher: The MIT Press, pp. 936–948.

Ganguli, Ina (July 2015). "Immigration and Ideas: What Did Russian Scientists "Bring" to the United States?" en. In: *Journal of Labor Economics* 33.S1, S257–S288. ISSN: 0734-306X, 1537-5307. DOI: 10.1086/679741. URL: https://www.journals.uchicago.edu/doi/10.1086/679741 (visited on 12/09/2022).

Gathmann, Christina (2010). "How General Is Human Capital? A Task-Based Approach". en. In.

Gross, Daniel P and Bhaven N Sampat (2020a). "Inventing the Endless Frontier: The Effects of the World War II Research Effort on Post-war Innovation". en. In: p. 57.

— (2020b). *The World War II Crisis Innovation Model: What Was It, and Where Does It Apply?* Working Paper 27909. Series: Working Paper Series. National Bureau of Economic Research.

Hannah Ritchie, Max Roser and Pablo Rosado (2020). "Energy". In: *Our World in Data*.

Iaria, Alessandro, Carlo Schwarz, and Fabian Waldinger (2018). "Frontier knowledge and scientific production: evidence from the collapse of international science". In: *The Quarterly Journal of Economics* 133.2. Publisher: Oxford University Press, pp. 927–991.

Jäger, Simon (2022). "How Substitutable Are Workers? Evidence from Worker Deaths". en. In: p. 52.

Jia, Ruixue et al. (Apr. 2022). *The Impact of U.S.-China Tensions on U.S. Science*. en. Tech. rep. w29941. Cambridge, MA: National Bureau of Economic Research, w29941. DOI: 10.3386/w29941. URL: http://www.nber.org/papers/w29941.pdf (visited on 11/09/2022).

Kantor, Shawn and Alexander Whalley (2022). "Moonshot: Public R&D and Growth". en. In.

Lazear, Edward P (2009). "Firm-Specific Human Capital: A Skill-Weights Approach". en. In: *journal of political economy*.

Marx, Matt (Oct. 2011). "The Firm Strikes Back: Non-compete Agreements and the Mobility of Technical Professionals". en. In: *American Sociological Review* 76.5, pp. 695–712. ISSN: 0003-1224, 1939-8271. DOI: 10.1177/0003122411414822. URL: http://journals.sagepub.com/doi/10.1177/0003122411414822 (visited on 12/13/2022).

Marx, Matt (Sept. 2022). "Employee Non-compete Agreements, Gender, and Entrepreneurship". en. In: *Organization Science* 33.5, pp. 1756–1772. ISSN: 1047-7039, 1526-5455. DOI: 10.1287/orsc.2021.1506. URL: http://pubsonline.informs.org/doi/10.1287/orsc.2021.1506 (visited on 12/15/2022).

Marx, Matt, Deborah Strumsky, and Lee Fleming (June 2009). "Mobility, Skills, and the Michigan Non-Compete Experiment". en. In: *Management Science* 55.6, pp. 875–889. ISSN: 0025-1909, 1526-5501. DOI: 10.1287/mnsc.1080.0985. URL: http://pubsonline.informs.org/doi/10.1287/mnsc.1080.0985 (visited on 11/09/2022).

Moretti, Enrico, Claudia Steinwender, and John Van Reenen (2019). *The intellectual spoils of war? Defense R&R, productivity and international spillovers*. Tech. rep. National Bureau of Economic Research.

Moser, Petra, Alessandra Voena, and Fabian Waldinger (2014). "German Jewish émigrés and US invention". In: *American Economic Review* 104.10, pp. 3222–55.

Myers, Kyle (Oct. 2020). "The Elasticity of Science". en. In: *American Economic Journal: Applied Economics* 12.4, pp. 103–134. ISSN: 1945-7782, 1945-7790. DOI: 10.1257/app.20180518. URL: https://pubs.aeaweb.org/doi/10.1257/app.20180518 (visited on 11/09/2022).

Myers, Kyle R. and Lauren Lanahan (July 2022). "Estimating Spillovers from Publicly Funded R&D: Evidence from the US Department of Energy". en. In: *American Economic Review* 112.7, pp. 2393–2423. ISSN: 0002-8282. DOI: 10.1257/aer.20210678. URL: https://pubs.aeaweb.org/doi/10.1257/aer.20210678 (visited on 11/09/2022).

Neal, Derek (Oct. 1995). "Industry-Specific Human Capital: Evidence from Displaced Workers". en. In: *Journal of Labor Economics* 13.4, pp. 653–677. ISSN: 0734-306X, 1537-5307. DOI: 10.1086/298388. URL: https://www.journals.uchicago.edu/doi/10.1086/298388 (visited on 12/16/2022).

Nelson, Richard R. (June 1959). "The Simple Economics of Basic Scientific Research". en. In: *Journal of Political Economy* 67.3, pp. 297–306. ISSN: 0022-3808, 1537-534X. DOI: 10.1086/258177. URL: https://www.journals.uchicago.edu/doi/10.1086/258177 (visited on 11/08/2022).

Prato, Marta (2022). "The Global Race for Talent: Brain Drain, Knowledge Transfer and Growth". en. In: p. 91.

Romer, Paul M (1990). "Endogenous technological change". In: *Journal of political Economy* 98.5, Part 2. Publisher: The University of Chicago Press, S71–S102.

Waldinger, Fabian (2010). "Quality matters: The expulsion of professors and the consequences for PhD student outcomes in Nazi Germany". In: *Journal of political economy* 118.4. Publisher: University of Chicago Press Chicago, IL, pp. 787–831.

— (2012). "Peer effects in science: Evidence from the dismissal of scientists in Nazi Germany". In: *The review of economic studies* 79.2. Publisher: Oxford University Press, pp. 838–861.

— (2016). "Bombs, brains, and science: The role of human and physical capital for the creation of scientific knowledge". In: *Review of Economics and Statistics* 98.5. Publisher: The MIT Press, pp. 811–831.

Appendix A: Tables

Table 2.7. Poisson DiD with semiclassical physics

	Number of publications	
	(1)	(2)
$Post \times Treat$	-0.5593	-0.5593
	(0.343)	(0.343)
Observations	32	32
Year F.E.		✓

The dependent variable is the yearly number of publications in a scientific field. Years 1980-1995 are considered. Fission is considered the treated field. Semiclassical physics is the control group. (1) features only the diff-in-diff term, a post-treatment dummy *Post* and a dummy for fission. (2) features also year-specific fixed effects.

Table 2.8. Effect of Chernobyl on individual publications (up to 2010)

	Nr. publications (1)	Log nr. publications (2)	Asinh nr. publications (3)
Post × *Treat*	-0.7097	-0.1016	-0.1233
	(0.1779)	(0.0585)	(0.0752)
R^2	0.45283	0.59779	0.59966
Observations	69,840	69,840	69,840
Author F.E.	✓	✓	✓
Year F.E.	✓	✓	✓

Years 1980-2010 are considered.
Every observation is an author-year pair.
Authors who published in nuclear fission before 1986 are considered as treated.
The control group is made of authors who published before 1986 in semiclassical physics, medical physics, engineering physics, theoretical physics and quantum mechanics.
In (1) the dependent variable is the number of publications per year.
In (2) the dependent variable is the log number of publications per year.
In (3) the dependent variable is the inverse hyperbolic sine of the number of publications per year.
Every regression features year and author fixed effects, plus the DiD term *Post* × *Treat*.
Standard errors are clustered by author.

Table 2.9. Effect of Chernobyl on individual publications (up to 2020)

	Nr. publications (1)	Log nr. publications (2)	Asinh nr. publications (3)
Post × *Treat*	-0.9367	-0.1019	-0.1213
	(0.1838)	(0.0544)	(0.0698)
R^2	0.33537	0.57224	0.57705
Observations	89,240	89,240	89,240
Author F.E.	✓	✓	✓
Year F.E.	✓	✓	✓

Years 1980-2020 are considered.
Every observation is an author-year pair.
Authors who published in nuclear fission before 1986 are considered as treated.
The control group is made of authors who published before 1986 in semiclassical physics, medical physics, engineering physics, theoretical physics and quantum mechanics.
In (1) the dependent variable is the number of publications per year.
In (2) the dependent variable is the log number of publications per year.
In (3) the dependent variable is the inverse hyperbolic sine of the number of publications per year.
Every regression features year and author fixed effects, plus the DiD term *Post* × *Treat*.
Standard errors are clustered by author.

Table 2.10. Effect of Chernobyl on individual citations (up to 2010)

	Nr. citations (1)	Log nr. citations (2)	Asinh nr. citations (3)
Post × *Treat*	-32.83	-0.2854	-0.3160
	(3.938)	(0.1323)	(0.1552)
R^2	0.27074	0.56782	0.56885
Observations	69,840	69,840	69,840
Author F.E.	✓	✓	✓
Year F.E.	✓	✓	✓

Years 1980-2010 are considered.
Every observation is an author-year pair.
Authors who published in nuclear fission before 1986 are considered as treated.
The control group is made of authors who published before 1986 in semiclassical physics, medical physics, engineering physics, theoretical physics and quantum mechanics.
In (1) the dependent variable is the number of citations per year.
In (2) the dependent variable is the log number of citations per year.
In (3) the dependent variable is the inverse hyperbolic sine of the number of citations per year.
Every regression features year and author fixed effects, plus the DiD term $Post \times Treat$.
Standard errors are clustered by author.

Table 2.11. Effect of Chernobyl on individual citations (up to 2020)

	Nr. citations (1)	Log nr. citations (2)	Asinh nr. citations (3)
Post × *Treat*	-53.19	-0.2526	-0.2761
	(5.400)	(0.1153)	(0.1355)
R^2	0.19609	0.55016	0.55262
Observations	89,240	89,240	89,240
Author F.E.	✓	✓	✓
Year F.E.	✓	✓	✓

Years 1980-2020 are considered.
Every observation is an author-year pair.
Authors who published in nuclear fission before 1986 are considered as treated.
The control group is made of authors who published before 1986 in semiclassical physics, medical physics, engineering physics, theoretical physics and quantum mechanics.
In (1) the dependent variable is the number of citations per year.
In (2) the dependent variable is the log number of citations per year.
In (3) the dependent variable is the inverse hyperbolic sine of the number of citations per year.
Every regression features year and author fixed effects, plus the DiD term $Post \times Treat$.
Standard errors are clustered by author.

Table 2.12. US: Aggregate DiD

	Dependent variable:					
	Nr. pub		Log nr. pub		Asinh nr. pub	
	(1)	(2)	(3)	(4)	(5)	(6)
Post × *Treat*	86.033***	66.979**	0.322**	0.330**	0.322**	0.331**
	(25.414)	(22.897)	(0.124)	(0.111)	(0.125)	(0.112)
Treat	46.667**	−5.075	0.506***	−0.055	0.512***	−0.041
	(20.091)	(18.102)	(0.098)	(0.088)	(0.099)	(0.088)
Post	103.483**	118.636***	0.850***	0.683***	0.859***	0.692***
	(37.047)	(33.378)	(0.181)	(0.162)	(0.182)	(0.163)
Constant	47.667*	110.735***	3.959***	4.702***	4.634***	5.367***
	(26.578)	(23.946)	(0.130)	(0.116)	(0.131)	(0.117)
Observations	32	32	32	32	32	32
R^2	0.904	0.876	0.946	0.902	0.946	0.903
Adjusted R^2	0.787	0.725	0.880	0.782	0.881	0.786
Note:				*p<0.1; **p<0.05; ***p<0.01		

Years 1980-1995 are considered.
In (1) and (2) the dependent variable is the number of publications in a scientific field (in a given year).
In (3) and (4) the dependent variable is the log number of publications in a scientific field (in a given year).
In (5) and (6) the dependent variable is the inverse hyperbolic sine of the number of publications in a scientific field (in a given year).
Nuclear fission is the treated field.
In (1), (3) and (5) the control group is semiclassical physics.
In (2), (4) and (6) the control group is a synthetic control, in the spirit of Abadie, Diamond, and Hainmueller (2010). It consists of a weighted mean of the following fields: semiclassical physics, medical physics, engineering physics, theoretical physics and quantum mechanics.
Every regression features year and field fixed effects, a post-treatment dummy *Post*, a dummy for fission and the DiD term.

Table 2.13. JP: Aggregate DiD

	Dependent variable:					
	Nr. pub		Log nr. pub		Asinh nr. pub	
	(1)	(2)	(3)	(4)	(5)	(6)
Post × *Treat*	21.833**	19.650**	0.292	0.445*	0.256	0.231
	(8.886)	(8.132)	(0.229)	(0.221)	(0.242)	(0.219)
Treat	14.667*	7.559	1.076***	0.246	1.183***	0.443**
	(7.025)	(6.429)	(0.181)	(0.174)	(0.191)	(0.173)
Post	26.083*	30.925**	1.587***	1.363***	1.743***	1.716***
	(12.954)	(11.855)	(0.335)	(0.322)	(0.353)	(0.319)
Constant	−2.333	4.735	1.240***	2.094***	1.702***	2.472***
	(9.294)	(8.505)	(0.240)	(0.231)	(0.253)	(0.229)
Observations	32	32	32	32	32	32
R^2	0.855	0.849	0.937	0.875	0.938	0.900
Adjusted R^2	0.680	0.665	0.860	0.724	0.864	0.779

Note: *p<0.1; **p<0.05; ***p<0.01

Years 1980-1995 are considered.
In (1) and (2) the dependent variable is the number of publications in a scientific field (in a given year).
In (3) and (4) the dependent variable is the log number of publications in a scientific field (in a given year).
In (5) and (6) the dependent variable is the inverse hyperbolic sine of the number of publications in a scientific field (in a given year).
Nuclear fission is the treated field.
In (1), (3) and (5) the control group is semiclassical physics.
In (2), (4) and (6) the control group is a synthetic control, in the spirit of Abadie, Diamond, and Hainmueller (2010). It consists of a weighted mean of the following fields: semiclassical physics, medical physics, engineering physics, theoretical physics and quantum mechanics.
Every regression features year and field fixed effects, a post-treatment dummy *Post*, a dummy for fission and the DiD term.

Table 2.14. FR: Aggregate DiD

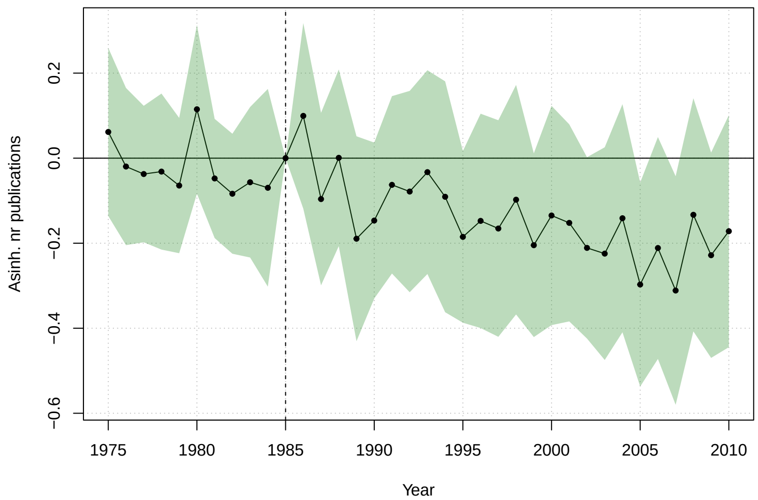
	Dependent variable:					
	Nr. pub		Log nr. pub		Asinh nr. pub	
	(1)	(2)	(3)	(4)	(5)	(6)
$Post \times Treat$	-2.700	-5.098^{*}	-0.038	-0.152	-0.010	-0.147
	(2.367)	(2.499)	(0.215)	(0.191)	(0.246)	(0.220)
$Treat$	-1.000	-1.878	-0.189	0.059	-0.233	0.053
	(1.871)	(1.976)	(0.170)	(0.151)	(0.195)	(0.174)
$Post$	13.850^{***}	17.382^{***}	1.141^{***}	1.357^{***}	1.248^{***}	1.509^{***}
	(3.450)	(3.643)	(0.313)	(0.278)	(0.359)	(0.320)
Constant	5.500^{**}	6.153^{**}	1.872^{***}	1.519^{***}	2.410^{***}	2.001^{***}
	(2.475)	(2.614)	(0.224)	(0.200)	(0.258)	(0.230)
Observations	32	32	32	32	32	32
R^2	0.838	0.865	0.828	0.864	0.820	0.859
Adjusted R^2	0.642	0.700	0.618	0.700	0.601	0.689

Note: *p<0.1; **p<0.05; ***p<0.01

Years 1980-1995 are considered.
In (1) and (2) the dependent variable is the number of publications in a scientific field (in a given year).
In (3) and (4) the dependent variable is the log number of publications in a scientific field (in a given year).
In (5) and (6) the dependent variable is the inverse hyperbolic sine of the number of publications in a scientific field (in a given year).
Nuclear fission is the treated field.
In (1), (3) and (5) the control group is semiclassical physics.
In (2), (4) and (6) the control group is a synthetic control, in the spirit of Abadie, Diamond, and Hainmueller (2010). It consists of a weighted mean of the following fields: semiclassical physics, medical physics, engineering physics, theoretical physics and quantum mechanics.
Every regression features year and field fixed effects, a post-treatment dummy $Post1986$, a dummy for fission and the DiD term δ.

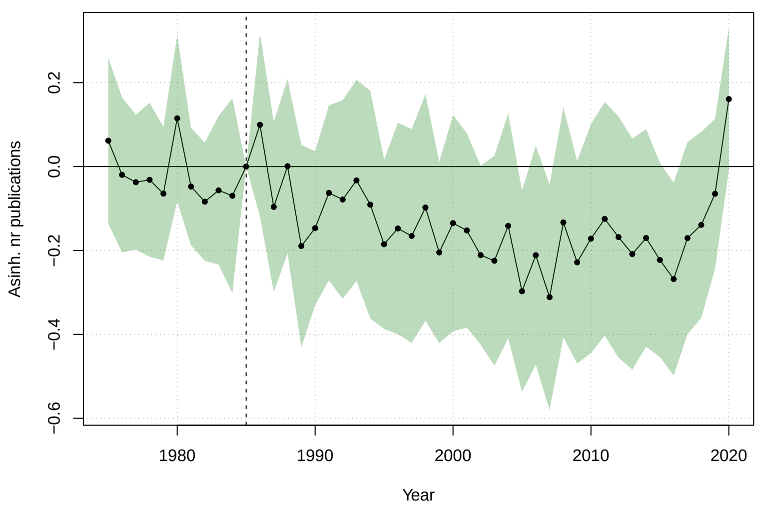
Appendix B: Figures

Figure 2.6. Effect of Chernobyl on individual publications (up to 2010)



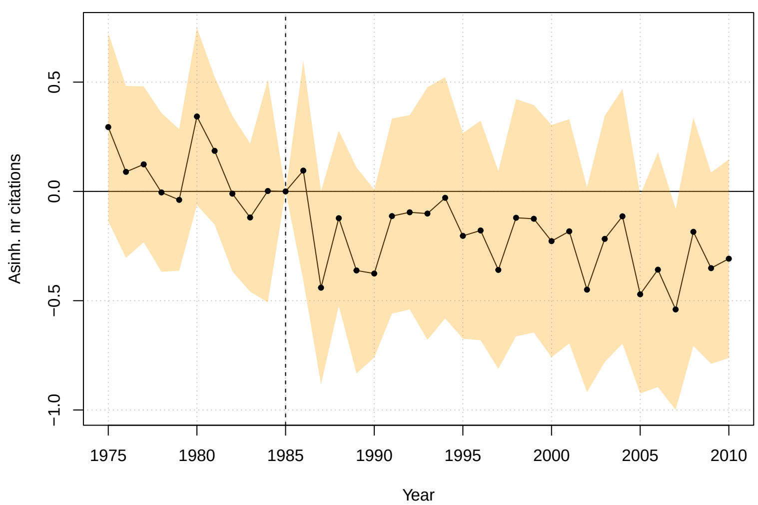
Event study plot. Effects of Chernobyl disaster onto the inverse hyperbolic sine of the number of publications of Italian scientists. Scientists who published at least once in 1972-1986 are part of the sample. Scientists who published at least once in fission (pre-Chernobyl) are considered as treated. Scientists who published at least once (pre-Chernobyl) in semiclassical physics, medical physics, engineering physics, theoretical physics and quantum mechanics are used as controls. Standard errors are clustered at the scientist level. Timespan considered for the event study: 1975-2010.
Source: Microsoft Academic Graph database

Figure 2.7. Effect of Chernobyl on individual publications (up to 2020)



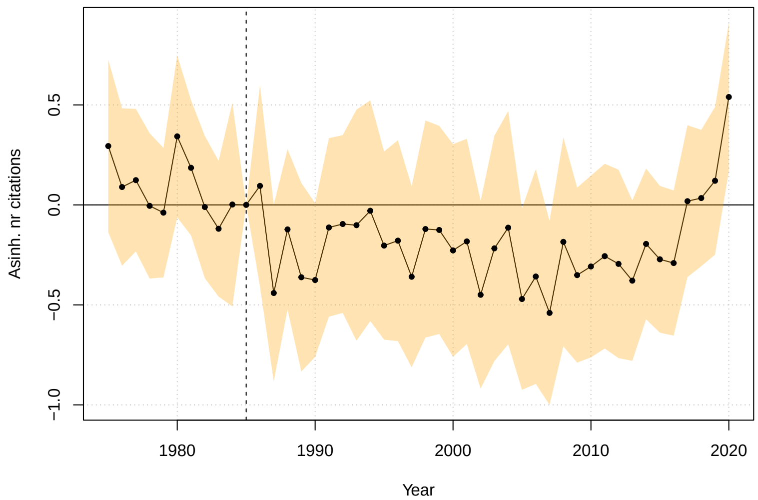
Event study plot. Effects of Chernobyl disaster onto the inverse hyperbolic sine of the number of publications of Italian scientists. Scientists who published at least once in 1972-1986 are part of the sample. Scientists who published at least once in fission (pre-Chernobyl) are considered as treated. Scientists who published at least once (pre-Chernobyl) in semiclassical physics, medical physics, engineering physics, theoretical physics and quantum mechanics are used as controls. Standard errors are clustered at the scientist level. Timespan considered for the event study: 1975-2020.
Source: Microsoft Academic Graph database

Figure 2.8. Effect of Chernobyl on individual citations (up to 2010)



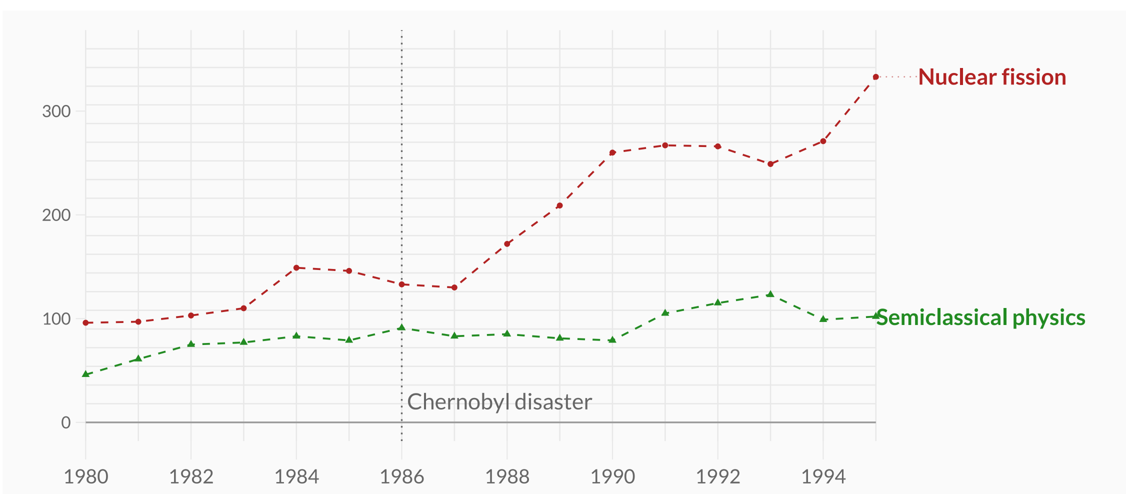
Event study plot. Effects of Chernobyl disaster onto the inverse hyperbolic sine of the number of citations of Italian scientists. Scientists who published at least once in 1972-1986 are part of the sample. Scientists who published at least once in fission (pre-Chernobyl) are considered as treated. Scientists who published at least once (pre-Chernobyl) in semiclassical physics, medical physics, engineering physics, theoretical physics and quantum mechanics are used as controls. Standard errors are clustered at the scientist level. Timespan considered for the event study: 1975-2010.
Source: Microsoft Academic Graph database

Figure 2.9. Effect of Chernobyl on individual citations (up to 2020)



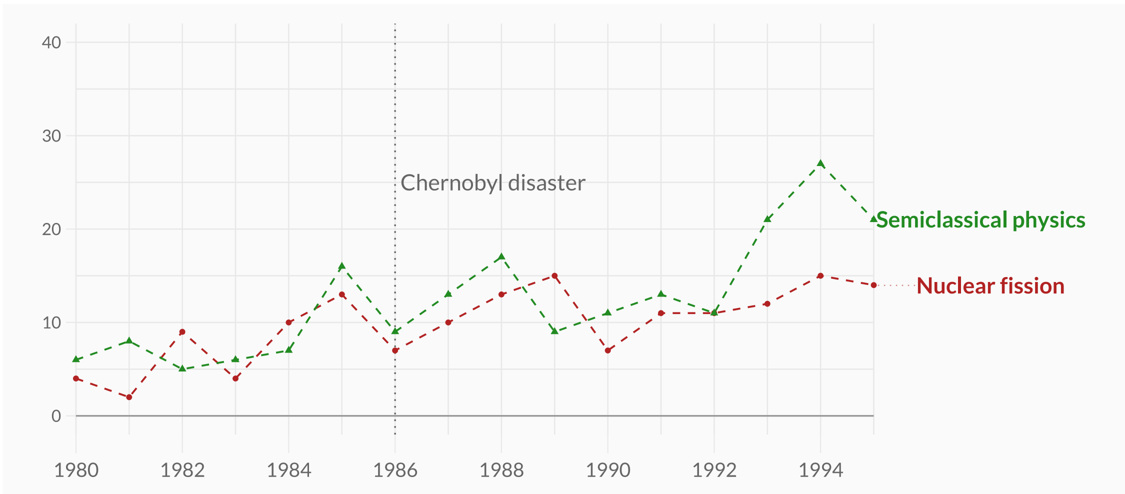
Event study plot. Effects of Chernobyl disaster onto the inverse hyperbolic sine of the number of citations of Italian scientists. Scientists who published at least once in 1972-1986 are part of the sample. Scientists who published at least once in fission (pre-Chernobyl) are considered as treated. Scientists who published at least once (pre-Chernobyl) in semiclassical physics, medical physics, engineering physics, theoretical physics and quantum mechanics are used as controls. Standard errors are clustered at the scientist level. Timespan considered for the event study: 1975-2020.
Source: Microsoft Academic Graph database

Figure 2.10. US: Time series of scientific publications



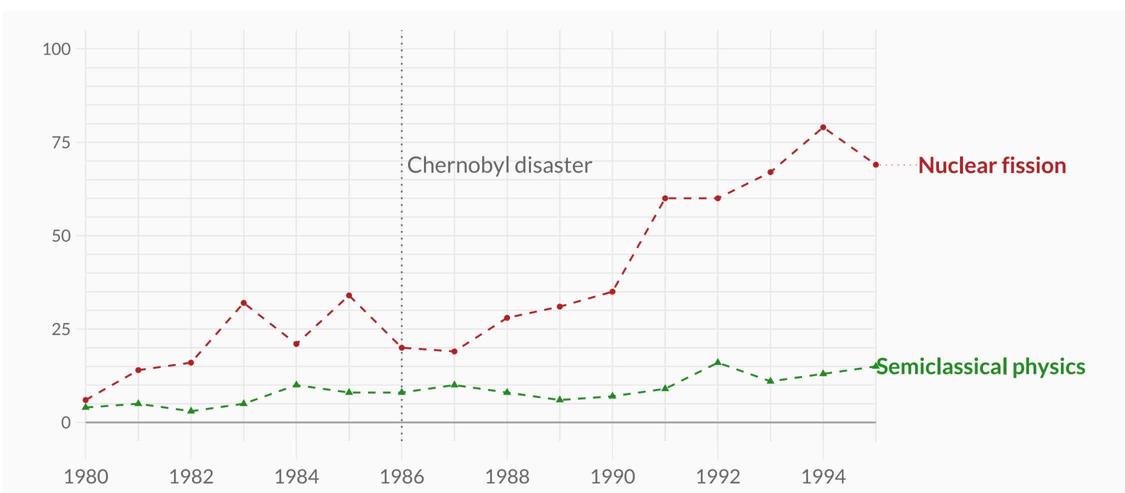
Yearly number of academic publications per scientific field by US authors. The red line is nuclear fission. The green line is semiclassical physics.
Source: Microsoft Academic Graph database

Figure 2.11. FR: Time series of scientific publications



Yearly number of academic publications per scientific field by French authors. The red line is nuclear fission. The green line is semiclassical physics.
Source: Microsoft Academic Graph database

Figure 2.12. JP: Time series of scientific publications



Yearly number of academic publications per scientific field by Japanese authors. The red line is nuclear fission. The green line is semiclassical physics.
Source: Microsoft Academic Graph database

Appendix C: Non-anticipation of the shock

The event of the Chernobyl disaster (and the subsequent referendum) is here interpreted as a sudden shock to the fundings for nuclear research. In order for this claim to be true, it has to be proved that Italian authorities were not already planning to reduce the contributions to nuclear fission and increase the amount of resources spent in fusion and renewables. Two documents support this claim:

1. *Relazione sul programma di attività e sui risultati conseguiti nel quinquennio 1980-1984*. It is a report presented to the Italian parliament documenting what ENEA has done in the years 1980-84, and describing the program for the following 5 years. The report for the 1980-84 years describes all the activities and the projects, including also detailed tables on the allocation of resources (p.59) and a long list of suppliers. In the part that explains the objectives for the 1985-89 period, most of the attention is devoted to nuclear fission. When listing the objectives of the plan (p.233), the extenders of the document put as the priority the realisation of new nuclear plants in Italy and the completion of the CIRENE reactor.

   The comment on the priorities is "the program is now at the apex of its effort, which is forecasted to last until the first nuclear plants of the National Electric Plan will start functioning. The current collaboration system features a whole bunch of big, small and medium enterprises, mostly belonging to the mechanical, electric and electronic sector."

2. *Attività svolta dall'ENEA nel biennio 1985-1986: Relazione del Presidente*. This document is a biennial report written by the ENEA president. It covers the activities carried on by the institution in the years 1985-86. With respect to the previous one, it contains more technical pieces of information. It covers the first two years of the 1985-89 plan, and presents updates on the realisation of the most important projects. Moreover, the report has been completed after the Chernobyl incident (but before the referendum).

   In its introductory pages, the president clearly states that "The irreplaceable role of this source [*nuclear*] has been firmly reaffirmed in every occasion, especially after the Chernobyl accident. [...] At the Tokyo G7 summit, held between the 4th and the 6th of May 1986, to which the Italian government took part, the participants agreed on the importance of the nuclear development in the industrialised countries, in order not to put excessive pressure on the costs of the fuel, which is crucial for economic growth in developing countries.(p.22)"

Then the document proceeds in analysing ENEA's major projects, saying that the CIRENE reactor had reached a completion percentage of "more than 90% at the end of 1985", while the PEC plant was "above 65% at the end of 1985". Both projects were suddenly interrupted in 1987.

Appendix D: Coding of the variables

I list here some technical notes on the way I defined variables.

Fission fields Every entry in the MAG database is associated to one or more fields of studies. I define an indicator "fission indicator" equal to 1 if any paper is associated to one of the following fields : "nuclear fission", "nuclear reactor", "nuclear power plant", "nuclear reactor safety systems", "nuclear reactor core", "nuclear material", "special nuclear material", "nuclear fission product", "nuclear reactor physics", "nuclear fuel cycle", "nuclear reactor coolant", "ford nuclear reactor", "economics of nuclear power plants", "convention on the physical protection of nuclear material", "weapons grade nuclear material"

Private institutions In the MAG database, universities and research centres are often associated to their website. I use this information in order to establish whether a research institution is public or not. In particular, I define an indicator being equal to 1 if any institutions has its website terminating with ".com" or having the string ".co." (typical in some countries such as the United Kingdom, where companies have often websites terminating in ".co.uk").

Chapter 3

Procuring innovation: evidence from the SBIR program

Abstract

Economists have been long debating about the effectiveness of policies aimed at inducing directed technical change. Policymakers can use several tools, such as military-driven public investment, research grants, or public procurement. I examine a unique amalgamation of such tools in a US context, where small businesses winning federal research grants (SBIR) can also secure procurement contracts from the Department of Defense. Employing a Difference-in-Differences framework and three comprehensive datasets - SBIR grants, DoD procurement contracts (USAspending.gov), and Patstat - I find that while procurement contract winners do not patent more than other SBIR participants, they get 1 million dollars more in government contracts every year.

3.1 Introduction

Fostering innovation is one of the primary objectives of economists and policymakers. Since Arrow (1962) researchers have studied the role economic actors have in directing technical change (Acemoglu et al., 2012; Slavtchev and Wiederhold, 2016). A central actor is arguably the state, which can set technological priorities and allocate resources to achieve them. However, how to best achieve them is still at the heart of the debate. Some economists have shown that big technological projects, such as the Manhattan project (Gross and Sampat, 2020), the Apollo program (Kantor and Whalley, 2022), or military spending (Moretti, Steinwender, and Van Reenen, 2019) can generate spillovers and crowd in private innovation. Others have shown the efficacy of smaller and more targeted tools, such as research grants (Sabrina T. Howell, 2017; Santoleri et al., 2022; Myers and Lanahan, 2022) or public procurement (Belenzon and Cioaca, 2022; Rassenfosse, Jaffe, and Raiteri, 2019). This paper studies a combination of grants and procurement, in a context in which funding priorities are set by the Department of Defense (DoD).

I focus on the SBIR program, a US federal program that awards research and development grants to small businesses. Originally created in 1982, this program aims to help small businesses conduct research and development. It is federally funded, but it is run separately by governmental agencies (such as the Department of Defense, the Department of Energy, and the National Institute of Health), that award grants according to their needs. Firms can apply to the program, and if they are selected, they receive a research grant to develop a proof of concept. If the proof of concept is successful, the firm can apply for a second research grant to develop a prototype. SBIR funds end there. Then, some agencies - among which the DoD - can award a procurement contract to help the firm commercialize the product.

One of the most famous examples of the program is iRobot, a firm that was awarded a research grant in 1998 by the Department of Defense to develop a robot (called PackBot) that could be used to detect and disarm bombs. The prototype was successful, and the firm was awarded a procurement contract by the DoD itself. At the same time, iRobot used the technology developed for the PackBot to create a commercial product, Roomba, a vacuum cleaner robot. Roomba was a commercial success, and the firm went public in 2005.

In this paper, I try to provide systematic evidence of cases like iRobot. First, I try to understand whether demand from the public sector directs innovation. I find that procurement contract winners do not patent more than other SBIR participants, neither in terms of quantity nor quality. Second, I investigate whether public procurement - on top of research grants - can impact firm performance. I find that procurement contracts help firms strenghten their ties with the DoD, getting around 0.6 more contracts (for a value of $1,000,000$$) per year. However, it is not yet clear whether those firms increase their market value, or they just become preferred suppliers of the DoD.

I combine three datasets: the first one contains all SBIR grants awarded from 1983 to 2023, the second one (USAspending.gov) contains all DoD procurement contracts from 2000 to 2021, and the third one (Patstat) contains all patents granted by the USPTO from 1976 to 2023. I test my hypotheses in a Difference-in-Differences framework, comparing procurement winners (i.e. firms that have been awarded a Phase III contract) to SBIR participants that have received only Phase I and Phase II contracts. My design does not allow me to get neat identification, as there are no public rankings of the firms that apply to the program.

This paper contributes to two strands of literature. First, there is a growing interest in the effect of research grants on innovation, with a particular focus on the SBIR program. Since the early 2000s, many economists (Lerner, 2000; Audretsch, 2003; Gans and Stern, 2003; Link and Scott, 2010; Keller and Block, 2013) have described the program as an effective policy instrument for the US government. However, only recently the availability of micro-data has allowed researchers to provide a quantitative assessment of the program.

Sabrina T. Howell (2017) cleanly shows that landing early stage SBIR financing (i.e. Phase I) makes firms more likely to access venture capital financing, patent and increase revenues. She also shows that this effect is stronger for liquidity constrained firms. Myers and Lanahan (2022) document that SBIR funding crowds in private investment, as for every patent produced by grant recipients, there are three more produced by other firms.

Recent work has tried to understand if the program was effective in directing technical change. S. Howell et al. (2021) show that an "open innovation" approach, that was introduced in the Air Force in 2018, has crowded in applicants and increased the quality of the projects. Firms that were awarded a grant under the new approach become more likely to be awarded a procurement contract. Bhattacharya (2021) tries to give a welfare assessment of the program. He focuses on procurement contracts of a specific DoD branch (the Navy) and shows that there is a substantial trade off between the aggregate social welfare (i.e. more innovation) and the DoD's objective (i.e. procurement of goods that are useful for military purposes).

I contribute to this literature by arguing that procurement contracts - on top of research grants - apparently do not direct innovation.

A second strand of literature studies the effect of procurement contracts on firm performance. Belenzon and Cioaca (2022) argue that changes in the content of DoD procurement contracts can affect the direction of innovation. They show that when awarded federal contracts, firms produce more publications, and tend not to protect their knowledge with patents. Hvide and Meling (2022) focus on a narrow set of procurement contracts, regarding road construction in Norway. They show that startups winning a procurement auction increase their sales and employment by 20% and become more profitable, compared to narrow losers. These effects seem to persist for several years. In Italy, construction firms that win a procurement contract are 70% more likely than losers to survive after 3 years, but they do not increase their productivity (Cappelletti and Giuffrida, 2021). Furthermore, their earnings become much more dependent on public sector contracts, generating a potentially vicious cycle. Two papers focus on firm behavior after receiving a SBIR grant. Sabrina T Howell and Brown (2022) show that SBIR funds are also used to raise compensations for incumbent workers. Lanahan, Joshi, and Johnson (2021) complement this evidence, showing that SBIR recipients create less jobs than losers, and even within the group of winners, receiving more money does not translate into creating more jobs. I complement this literature by focusing on a different industry (military procurement), which is more technology-intensive than construction, and more strictly linked to government priorities. As in Cappelletti and Giuffrida (2021), I find that procurement contracts help firms strengthen their ties with the public sector, but I do not find evidence of direct impact on firm performance.

The remainder of the paper proceeds as follows. Section 3.2 describes the SBIR program and the data I use. Section 3.3 presents the preliminary results. Section 3.4 concludes.

3.2 Context

3.2.1 The SBIR program at the Department of Defense

SBIR is a US federal program that aims to help small businesses conduct research and development. It was created in 1982 by the Reagan administration, but all US governments since then have supported it with increased funding.[1] The program is federally funded, but it is run separately by 12 governmental agencies (such as the Department of Defense, the Department of Energy or the National Institute of Health).[2] It is open to all small businesses that are majority-owned by US citizens and consists of two phases. In Phase I, each agency lists their own priorities (i.e. the problems they want to solve), and firms can apply to the program with their own proposals. There is no need for prototypes. Firms just have to show that their idea

[1]See National Research Council (2014), National Research Council (2020) or similar reports for a detailed history of the program.

[2]Full list is available at https://www.sbir.gov/agencies-landing.

is feasible.

If a firm is selected, it receives a research grant to develop a proof of concept. The grant can be up to $200,000, and it lasts for 6 months. If the proof of concept is successful, the firm can apply for a second research grant ("Phase II") to develop a prototype. The grant can be up to $1,500,000, and it lasts for 2 years. After the second phase, SBIR funds end.

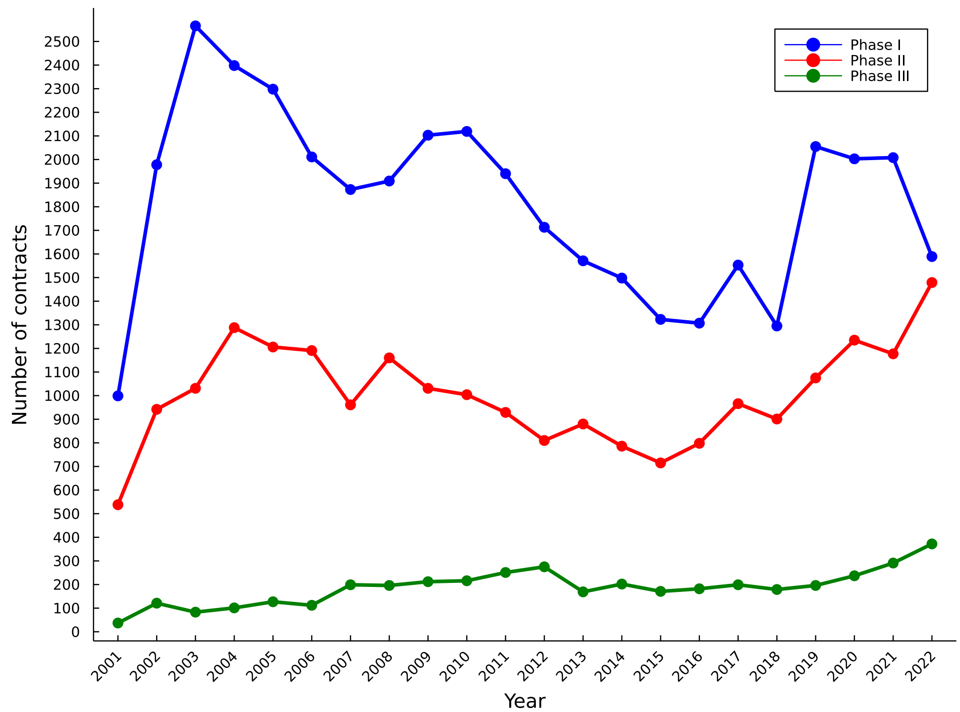
Those features are common to all SBIR grants, regardless of the agency that awards them. However, there is vast heterogeneity in the way the program is run by each agency. Indeed, some (such as the National Institute of Health or National Science Fundation) are more similar to traditional grant-awarding agencies. Their work is mostly to establish scientific priorities, and they award grants to firms that can help them achieve those priorities. On the other hand, agencies such as the Department of Defense or NASA are more concerned about their own needs. They use the program to individuate firms that can help them solve their problems. Formally speaking, the Department of Defense does not award grants, but signs contracts with firms, even for Phase I and Phase II.

Furthermore, in recent years some agencies have introduced a third phase, that is financed by their own budget.[3] It aims to help Phase II winners commercialize their product. However, there is even starker heterogeneity in the way this phase is run. Grant-awarding agencies (such as the Department of Energy) are more focused on finding external customers for the product, while agencies that award contracts (such as the Department of Defense) are interested in buying the product themselves. In particular, after winning a Phase III contract at the DoD, firms are entitled to sell their specific product (i.e. the one they developed using Phase I and II funds) to the DoD without having to face competition from other firms.

Throughout this paper, I will focus on Phase III (also referred to as "procurement") contracts awarded by the Department of Defense. The graph in Figure 3.1 shows the number of SBIR contracts awarded by the Department of Defense from 2000 to 2022. Phase I and Phase II contracts represent the largest share of the contracts awarded by the DoD, averaging more than 1,000 contracts per year. However, the number of Phase III contracts has increased in the recent years, as almost 300 contracts were awarded in 2022.

[3] There is no official introduction year, but this type of contracts have been used more and more frequently from the 2000s on. See https://www.sbir.gov/about for more information.

Figure 3.1. Department of Defense, number of contracts



3.2.2 Data

I use three different datasets, one on SBIR grants, one on DoD procurement contracts, and one on patents. I combine them to create a panel of firms that have been awarded at least a Phase II contract, and observe their activities, on innovation (measured by patents) and government contracts (measured by non-SBIR procurement contracts).

SBIR grants The first dataset contains all SBIR grants awarded by all agencies from 1983 to 2023. It is retrieved from the SBIR website (https://www.sbir.gov/sbirsearch/award/all). Each contract has a unique identifier, and it contains information on the awarded firm, the agency and branch that awarded the contract, the year, the phase and the amount of money awarded. It also contains the title and the abstract of the proposal, which would allow me to define the technological content of the contract, for further analysis. The dataset contains 197,664 observations (each one corresponding to a grant or contract), 89,134 of whose are contracts awarded by the Department of Defense.

Filtering only post-2000 observations, I obtain 66,609 unique contracts, 45,232 unique Phase I contracts, and unique 23,441 Phase II contracts. Most of the times the contract id changes from Phase I to Phase II. Sometimes it does not, and it creates discrepancies in the numbers above. The total number of firms is 9298 (identified by Dun and Bradstreet number). In the dataset it is quite common to find firms that have been awarded multiple contracts. Among firms that have been awarded at least a Phase I contract, 4658 (equal to 52.27%) have been awarded just one. Similarly, among firms that have been awarded at least a Phase II contract, 2925 (equal to 52.84%) have been awarded just one.

DoD procurement contracts The second dataset contains all DoD procurement contracts awarded from 2000 to 2021. It is retrieved from USAspending.gov, a publicly available database that registers all federal contracts awarded since 2000, broken down by agency. Each entry corresponds to a contract, and it contains several (288) features on the awarded firm, the awarding agencies (and branches), the competition rules that were applied, and the type of contract. I use this dataset for two purposes. First, I filter only SBIR-related contracts, exploiting the fact that not only it contains Phase I and Phase II (already present in the SBIR dataset), but also Phase III contracts. Second, I use it to retrieve non-SBIR contracts awarded to SBIR firms, that are further used to construct the firm panel.

I downloaded all DoD contracts since 2000, and I filtered them by contract type. I kept only contracts that were labeled as "research" contracts, corresponding to any of the three phases of the SBIR program. This procedure results in a dataset of 66,340 contracts, of which 40,206 are Phase I, 22,401 are Phase II, and 4219 are Phase III. There are two possible identifiers for the firm that was awarded the contract: the Duns number and the UEI number. Being consistent with the SBIR dataset, I use the Duns number, and I find 9819 unique firms. An alternative identifier is the UEI number, which identifies a slightly larger number of firms (10,555).

As it was already clear from the SBIR dataset, it is quite common to find firms that have been awarded multiple Phase I or Phase II contracts. 7934 firms have been awarded at least a Phase I contract. Out of those, 4374 (55.13%) have been awarded just one. Similarly, 2738 firms (53.67%) out of the 4431 Phase II winners, have won the contract just once. On the other hand, the vast majority of Phase III winners have been awarded just one contract. Indeed, the dataset contains 1,926 Phase III winners, and 1367 of them (70.98%) have been awarded just one contract.

Patents The third dataset is PATSTAT, which tracks all wordwide patent applications. It is retrieved from the European Patent Office, and it contains several pieces of information on the patent, such as inventor, year of application, title, abstract, follow-on citations, etc. Each patent is assigned to its inventor, that may be an individual or a corporation. However, in the case of firms, there is no identifier such as Duns or UEI. I perform a fuzzy matching algorithm, selecting all patent assignees that have a name that is similar to a SBIR grant winner. I choose a threshold of 0.9, with Levenshtein distance. Practically speaking, it means that I select all assignees whose name is at least 90% similar to a SBIR grant winner. Results are robust to different thresholds (0.8, 0.95).

Matching the datasets In order to create a dataset that can be used in regression analysis, I need to join the three datasets described above. First, I match the SBIR dataset with the DoD dataset, using the contract identifier. I keep only Phase I and Phase II contracts, because Phase III contracts are not stored in the SBIR dataset. The resulting dataset contains 48,266 unique contracts, 31,791 unique Phase I contracts, and unique 16,885 Phase II contracts. The number of unique firms is 7271 (identified by Duns number). A small number of firms have multiple Duns, so that the unique number of firms, identified by name-Duns pairs, is 7089. I use information on those firms to search for their patents in Patstat, as described above.

Firm panel Throughout the regression exercises I will use a panel of firms that have been awarded at least a Phase II contract. I will do it for two reasons. First, it is not possible to directly link Phase III contracts to Phase I and Phase II contracts, because there is no common identifier. It would be possible to impute Phase III contracts to the closest Phase II contract, in terms of technological content, but it would be a very rough approximation, worsened by the fact that the description of Phase III contracts is often very vague in the USAspending.gov dataset. Second, although contracts are awarded for a specific technology, funds are awarded to firms, that have a certain freedom in deciding how to use them. Therefore, a firm-level analysis is more appropriate to properly evaluate the program. Notice also that most of Phase III winners have been awarded just one contract.

The panel is built as follows: I select all 4134 firms that have been awarded at least a Phase II contract, and I observe their activities in the following years. I observe the number of Phase I and Phase II contracts awarded to the firm, the number of patents granted to the firm, and the number of citations received by the firm's patents. I also observe when and whether the firm has been awarded a Phase III contract (which is my treatment variable, in econometric terms), and the number of non-SBIR contracts awarded to the firm.

3.3 Results

3.3.1 Empirical strategy

I now empirically test whether getting access to public procurement affects firm outcomes. I focus on two set of outcomes: innovative activities (measured by patents) and further procurement contracts (measured by non-SBIR contracts). In order to do so, I consider all firms that received at least a Phase II contract, and I compare procurement winners (i.e. firms that have been awarded a Phase III contract) with non-procurement winners. I do it in a Difference-in-Differences framework, where the treatment is the award of a Phase III contract. Notice that some (30%) firms have received multiple Phase III contracts, so that they are treated multiple times. I consider as treatment only the first Phase III contract awarded to each firm. Notice also that treatment is staggered throughout time, as firms are awarded Phase III contracts in different years.

More important, the results that I present here do not rely on any identification strategy. Although some policies have been introduced in the recent years, there has been no abrupt change in the way Phase III contracts are awarded. Furthermore, as discussed in Rathje (2019), the decision to award a Phase III is rather subjective, and depends on decisions made by the program managers. Figure 3.11 in Appendix A shows how the conditional probabilities to get to Phase III have changed over time, and across DoD branches.

My baseline specification is the following:

$$y_{i,t} = \sum_{\tau=-5}^{5} \beta_\tau \text{YearsSinceProcurement}_{i,t+\tau} + \gamma_i + \delta_t + \epsilon_{i,t} \tag{3.1}$$

where $y_{i,t}$ is the number of patents (or non-SBIR contracts) granted to firm i in year t (in alternative specifications I use also the inverse hyperbolic sine, the log or just an indicator), $\text{YearsSinceProcurement}_{i,t}$ is a variable that counts how many years have passed since firm i was awarded a procurement contract (ranging from -5 to 5), γ_i and δ_t are - respectively - firm and year fixed effects. I estimate the equation using OLS, and I cluster the standard errors at the firm level. As described in section 3.2.2, my panel is made of 4134 firms that have been awarded at least a Phase II contract in the years 2000-2021.

3.3.2 Innovative activities

On the one hand, getting access to public procurement could provide firms with some source of guaranteed demand (Belenzon and Cioaca, 2022; Arora, Belenzon, and Sheer, 2021; Hvide and Meling, 2022), that could ease their financial constraints and help them develop their product, and so produce more innovation. On the other hand, getting access to public procurement could soften the incentive to innovate, as firms could rely on secured profits without having to further sharpen their product. Here I consider a standard - although imperfect - measure of innovation, which is the number of patents granted to the firm. I consider two dimensions of patenting activity: the number of patents granted to the firm, and the number of citations received by the firm's patents. While the former is a measure of quantity, the latter should also capture - at least partially - the quality of the patents.

Conditional on patenting, firms in my dataset file around 0.2 successful patent applications (and 3.8 citations) per year. The average number of patents per firm is 3.9 (73.3 citations).

Figure 3.2. Procurement does not affect patenting activity

(a) Asinh of patents (b) Asinh of citation-weighted-patents

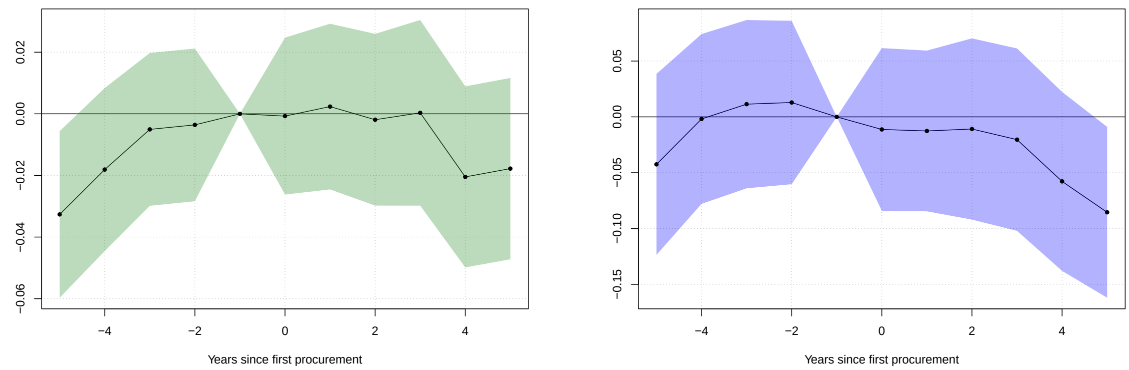


Figure 3.2 shows the event study plot, where I estimate Equation 3.1 for the inverse hyperbolic sine of the number of patents (panel a) and citations (panel b). It is not possible to reject the null hypothesis that landing a Phase III contract has no effect on patenting. Some slight form of anticipation seems to be present, as the number of patents slightly increases in the years before the award of the contract. It may signal that as firms approach the timing of the award, their technology is already mature for a patent application. However, the effect is not statistically significant, and it is even more attenuated when considering citations.

Results shown in Figure 3.2 reinforce the findings of both S. Howell et al. (2021) and Belenzon and Cioaca (2022). S. Howell et al. (2021) show that - in the context of traditional SBIR grants in the Air Force - Phase II winners are not more likely to patent than just Phase I winners. The same appears to hold true even comparing Phase III to Phase II winners. More generally, Belenzon and Cioaca (2022) argue that procurement contracts do not have a direct impact on patenting. However, they consider the whole universe of DoD procurement contracts, encompassing also many contracts that are not research-related. In this case, narrowing down the analysis to research-related contracts does not seem to change the results.

Table 3.1 and 3.2 in Appendix A shows the Diff-in-Diff results for different specifications, for patents and citation-weighted patents respectively.

3.3.3 Non-SBIR contracts

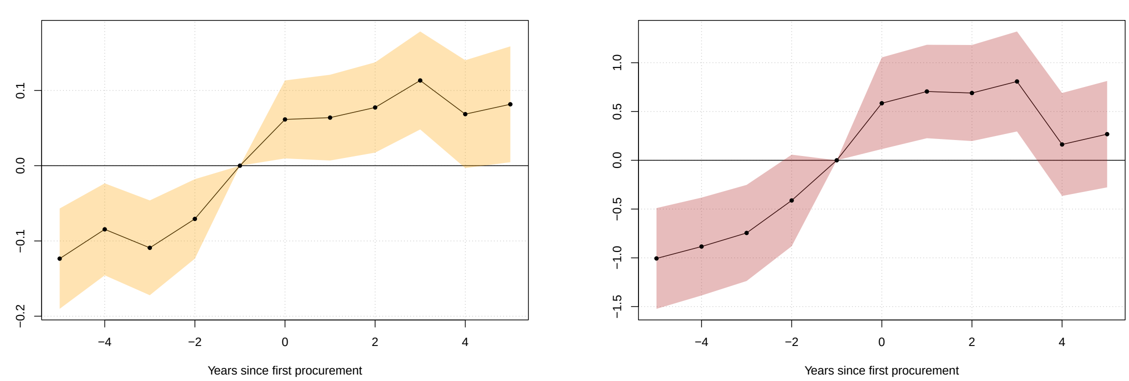
The second outcome I consider is the number of non-SBIR contracts awarded to the firm. As it was explained in Section 3.2, Phase III at the DoD is meant to be a way to open the gates of the DoD procurement system to SBIR firms. It gives the awarded firm the legal right to sell its product to the DoD, without having to face competition from other firms. However, this right is limited to the product that was developed throughout the previous phases of the SBIR program (National Research Council, 2020). It would be reasonable to expect that firms that have been awarded a Phase III contract are more likely to be awarded further contracts. Results in Figure 3.3 confirm this intuition.

Both the number of contracts (panel a) and the dollar value of contracts (panel b) increase after the award of a Phase III contract. Conditional on being awarded at least a contract in the whole timeframe, firms in my dataset are awarded around 2.2 contracts per year, for a total value of $840,000. The average number of contracts per firm is 38.8, for a total value of $14,976,000.

Figure 3.3. Non-SBIR contracts increase after Phase III

(a) Asinh of contracts number

(b) Asinh of contracts dollar value



After getting access to public procurement, firms are awarded around 0.6 more contracts per year, for a total value of $1,000,000. These results reinforce the findings of S. Howell et al. (2021), who show that Phase II winners are more likely to be awarded further procurement contracts. In this case, it appears that - among Phase II winners - those that are awarded a Phase III contract are even more likely to strengthen their ties with the DoD. At the same time, they are also consistent with Hvide and Meling (2022), who analyze public procurement auctions in Norway, and show that firms that are awarded a public procurement contract are more likely to receive further contracts from the same agency.

I then break down the number of contracts into competed and non-competed. Indeed, Phase III winners are allowed to sell the product they developed without having to face competition from other firms. It would be then reasonable to expect that the increase in procurement contracts is mostly due to non-competed contracts, rather than competed ones. Figures 3.4 and 3.5 confirm this intuition. The number and the dollar value of competed contracts (figure 3.4) increases only slightly after the award of a Phase III contract, and the effect seems to fade away after 5 years. On the other hand, the number and the dollar value of non-competed contracts (figure 3.5) increases more steadily, and the effect is still visible after 5 years. In the latter case, after 5 years firms are awarded around 0.5 more contracts per year, for a yearly value of $730,000.

Figure 3.4. Competed non-SBIR contracts slightly increase after Phase III

(a) Asinh of contracts number

(b) Asinh of contracts value

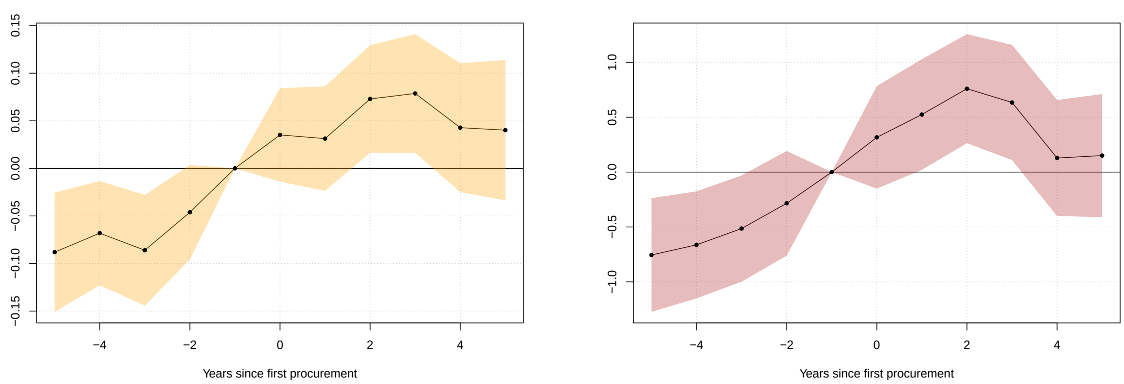
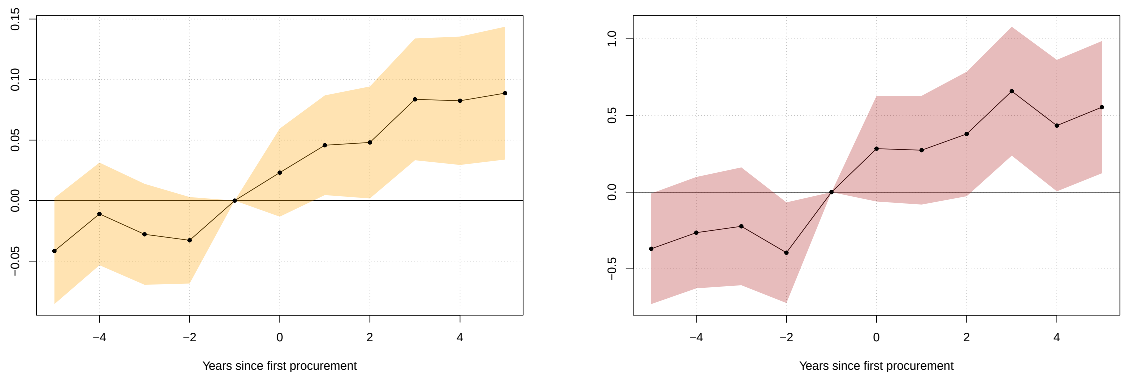


Figure 3.5. Noncompeted non-SBIR contracts increase more after Phase III

(a) Asinh of contracts number (b) Asinh of contracts value



It appears, then, that the increase in non-SBIR contracts is mostly driven by the increase in non-competed contracts. Put it differently, Phase III is effective in opening the gates of the DoD procurement system to SBIR firms, but mostly for a specific product. Based on this evidence, it is not possible to assess formally whether the program is welfare improving or not. On the one hand, the SBIR program has the broad purpose of helping small businesses grow and innovate. On the other hand, every agency is free to manage it as it prefers. The DoD is explicitly interested in addressing its own needs. Their goal is to develop technologies that can be used for military purposes, strenghten the ties with the firms that developed them, and add them to the pool of potential contractors. The results that I showed so far suggest that Phase III winners do not become more innovative (at least in terms of patenting). However, they seemingly become more valuable to the DoD, that is more likely to award them further contracts, especially for the product that was developed in the SBIR program. The economic question that arises is whether and how becoming part of the DoD procurement system can benefit small businesses.

3.3.4 Robustness

Recent evolutions in the debate about the use of Two Way Fixed Effects estimators (TWFE) have shown that when there is heterogeneity in the treatment effect, the standard TWFE estimator is biased (De Chaisemartin and D'Haultfœuille, 2020; Callaway and Sant'Anna, 2021; Borusyak and Hull, 2020). In the case of the SBIR program, the treatment (i.e. the award of a Phase III contract) is not awarded a single time, but it is staggered throughout time. In principle, there is no reason to suppose that - as the program developed - firms funded in different years received a similar treatment effect. Lanahan and Feldman (2015) and Lanahan and Feldman (2018) show that the program might have heterogeneous effects at state level, mainly due to the presence of statal funds (i.e. the State match program) that come on top of SBIR funding. While there is no State match program within the DoD, cohort-specific treatment effects might still be present. Indeed, the program might have gotten more effective over time, or firms might have learned how to better exploit the program. In order to take into account cohort-specific treatment effects, I use the estimator proposed by Sun and Abraham (2021). The estimator is more effective when the sample is divided into a small group of treated units (around 20% of the sample) and a large group of control units, that are never treated. In my sample I have 901 treated units, out of 4134 firms. There are 20 cohorts, from 2001 to 2020. In Appendix B I reproduce Figures 3.2 - 3.5 of the main text using the Sun and Abraham (2021) estimator. Results are qualitatively and quantitatively similar to the ones reported in the main text.

However, there is another potential problem with the estimates reported in this paper. It is not possible to observe why some firms have been awarded a Phase III contract, while others have not. Therefore, ex ante, the treatment is not exogenous. While the Sun and Abraham (2021) estimator takes into account cohort-specific treatment effects, it does not address the problem of endogenous selection into treatment, which is still a potential concern in this case. Borusyak and Hull (2020) propose a propensity-score method to assess - at least partially - the problem of non-random exposure to exogenous shocks. However, it would not fit well in my case, for two separate reasons. First, I have not been able to identify an exogenous aggregate shock (e.g. in budget allocation or in Defense priorities) that could have affected the award of Phase III contracts. Second, their method heavily relies on covariates that are not currently available in my database.

3.4 Conclusion

The effectiveness of public procurement as a catalyst for innovation has been a subject of ongoing discourse among economists (Slavtchev and Wiederhold, 2016). This study delves into a specific subset of procurement beneficiaries – firms already granted research awards – and their potential to engage in military contracts. Interestingly, while winning a procurement contract does not correlate with heightened patenting activity, these firms become integrated into the DoD procurement ecosystem, enjoying increased likelihood of securing subsequent contracts (approximately 0.6 additional contracts per year, with a cumulative value of $1,000,000). The majority (80%) of these contracts are non-competitive, directly awarded by DoD officials, facilitating new contractor retention. From the DoD's perspective, the program is successful in attracting and maintaining new contractors. However, assessing the program's overall welfare impact remains intricate. SBIR participants are primarily small, young businesses. One of the most desirable outcomes for such firms is to be acquired by larger firms, which can provide the resources necessary to grow and innovate. It is crucial to consider the nature of the acquiring firm. If winning a procurement contract increases the value of SBIR participants only for large DoD contractors, which eventually acquire them, the program's intended benefits could be diluted. Instead of fostering competition, it may reduce it. Conversely, acquisitions by non-DoD contractors might offer more favorable welfare implications, facilitating technology transfer to broader sectors. Future research directions involve scrutinizing acquisition frequency among SBIR participants and the nature of the entities involved in such transactions.

References

Acemoglu, Daron et al. (Feb. 2012). "The Environment and Directed Technical Change". en. In: *American Economic Review* 102.1, pp. 131–166. ISSN: 0002-8282. DOI: 10.1257/aer.102.1.131. URL: https://pubs.aeaweb.org/doi/10.1257/aer.102.1.131 (visited on 11/08/2022).

Arora, Ashish, Sharon Belenzon, and Lia Sheer (Mar. 2021). "Knowledge Spillovers and Corporate Investment in Scientific Research". en. In: *American Economic Review* 111.3, pp. 871–898. ISSN: 0002-8282. DOI: 10.1257/aer.20171742. URL: https://pubs.aeaweb.org/doi/10.1257/aer.20171742 (visited on 11/16/2022).

Arrow, Kenneth J (1962). "Economic Welfare and the Allocation of Resources for Invention". en. In: p. 19.

Audretsch, David B (2003). "Standing on the Shoulders of Midgets: The U.S. Small Business Innovation Research Program (SBIR)". en. In.

Belenzon, Sharon and Larisa C Cioaca (2022). "Buyers of First Resort: Guaranteed Public Demand and Corporate Scientific Research". en. In: p. 85.

Bhattacharya, Vivek (2021). "An empirical model of R&D procurement contests: An analysis of the DOD SBIR program". In: *Econometrica* 89.5. Publisher: Wiley Online Library, pp. 2189–2224.

Borusyak, Kirill and Peter Hull (Sept. 2020). *Non-Random Exposure to Exogenous Shocks: Theory and Applications*. en. Tech. rep. w27845. Cambridge, MA: National Bureau of Economic Research, w27845. DOI: 10.3386/w27845. URL: http://www.nber.org/papers/w27845.pdf (visited on 11/09/2022).

Callaway, Brantly and Pedro H.C. Sant'Anna (Dec. 2021). "Difference-in-Differences with multiple time periods". en. In: *Journal of Econometrics* 225.2, pp. 200–230. ISSN: 03044076. DOI: 10.1016/j.jeconom.2020.12.001. URL: https://linkinghub.elsevier.com/retrieve/pii/S0304407620303948 (visited on 08/23/2023).

Cappelletti, Matilde and Leonardo M Giuffrida (2021). "Procuring survival". en. In.

De Chaisemartin, Clément and Xavier D'Haultfœuille (Sept. 2020). "Two-Way Fixed Effects Estimators with Heterogeneous Treatment Effects". en. In: *American Economic Review* 110.9, pp. 2964–2996. ISSN: 0002-8282. DOI: 10.1257/aer.20181169. URL: https://pubs.aeaweb.org/doi/10.1257/aer.20181169 (visited on 08/23/2023).

Gans, Joshua and Scott Stern (Aug. 2003). "When does funding research by smaller firms bear fruit?: Evidence from the SBIR program*". en. In: *Economics of Innovation and New Technology* 12.4, pp. 361–384. ISSN: 1043-8599, 1476-8364. DOI: 10.1080/10438590290004555. URL: http://www.tandfonline.com/doi/abs/10.1080/10438590290004555 (visited on 02/01/2023).

Gross, Daniel P and Bhaven N Sampat (2020). "Inventing the Endless Frontier: The Effects of the World War II Research Effort on Post-war Innovation". en. In: p. 57.

Howell, Sabrina et al. (Apr. 2021). *Opening up Military Innovation: Causal Effects of Reforms to U.S. Defense Research*. en. Tech. rep. w28700. Cambridge, MA: National Bureau of Economic Research, w28700. DOI: 10.3386/w28700. URL: http://www.nber.org/papers/w28700.pdf (visited on 02/01/2023).

Howell, Sabrina T and J David Brown (2022). "Do Cash Windfalls Affect Wages? Evidence from R&D Grants to Small Firms". en. In: *The Review of Financial Studies* 00.0.

Howell, Sabrina T. (Apr. 2017). "Financing Innovation: Evidence from R&D Grants". en. In: *American Economic Review* 107.4, pp. 1136–1164. ISSN: 0002-8282. DOI: 10.1257/aer.20150808. URL: https://pubs.aeaweb.org/doi/10.1257/aer.20150808 (visited on 11/15/2022).

Hvide, Hans K and Tom G Meling (Dec. 2022). "Do Temporary Demand Shocks Have Long-Term Effects for Startups?" en. In: *The Review of Financial Studies* 36.1. Ed. by Tarun Ramadorai, pp. 317–350. ISSN: 0893-9454, 1465-7368. DOI: 10.1093/rfs/hhac028. URL: https://academic.oup.com/rfs/article/36/1/317/6588700 (visited on 03/02/2023).

Kantor, Shawn and Alexander Whalley (2022). "Moonshot: Public R&D and Growth". en. In.

Keller, Matthew R. and Fred Block (Oct. 2013). "Explaining the transformation in the US innovation system: the impact of a small government program". en. In: *Socio-Economic Review* 11.4, pp. 629–656. ISSN: 1475-147X, 1475-1461. DOI: 10.1093/ser/mws021. URL: https://academic.oup.com/ser/article-lookup/doi/10.1093/ser/mws021 (visited on 02/06/2023).

Lanahan, Lauren and Maryann P. Feldman (Sept. 2015). "Multilevel innovation policy mix: A closer look at state policies that augment the federal SBIR program". en. In: *Research Policy* 44.7, pp. 1387–1402. ISSN: 00487333. DOI: 10.1016/j.respol.2015.04.002. URL: https://linkinghub.elsevier.com/retrieve/pii/S0048733315000621 (visited on 02/01/2023).

— (Oct. 2018). "Approximating Exogenous Variation in R&D: Evidence from the Kentucky and North Carolina SBIR State Match Programs". en. In: *The Review of Economics and Statistics* 100.4, pp. 740–752. ISSN: 0034-6535, 1530-9142. DOI: 10.1162/rest_a_00681. URL: https://direct.mit.edu/rest/article/100/4/740-752/58494 (visited on 11/16/2022).

Lanahan, Lauren, Amol M. Joshi, and Evan Johnson (Sept. 2021). "Do public R&D subsidies produce jobs? Evidence from the SBIR/STTR program". en. In: *Research Policy* 50.7, p. 104286. ISSN: 00487333. DOI: 10.1016/j.respol.2021.104286. URL: https://linkinghub.elsevier.com/retrieve/pii/S0048733321000883 (visited on 02/01/2023).

Lerner, Josh (May 2000). "The Government as Venture Capitalist: The Long-Run Impact of the SBIR Program". en. In: *The Journal of Private Equity* 3.2, pp. 55–78. ISSN: 1096-5572, 2168-8508. DOI: 10.3905/jpe.2000.319960. URL: http://jpe.pm-research.com/lookup/doi/10.3905/jpe.2000.319960 (visited on 02/01/2023).

Link, Albert N. and John T. Scott (June 2010). "Government as entrepreneur: Evaluating the commercialization success of SBIR projects". en. In: *Research Policy* 39.5, pp. 589–601. ISSN: 00487333. DOI: 10.1016/j.respol.2010.02.006. URL: https://linkinghub.elsevier.com/retrieve/pii/S0048733310000545 (visited on 02/01/2023).

Moretti, Enrico, Claudia Steinwender, and John Van Reenen (2019). *The intellectual spoils of war? Defense R&R, productivity and international spillovers*. Tech. rep. National Bureau of Economic Research.

Myers, Kyle R. and Lauren Lanahan (July 2022). "Estimating Spillovers from Publicly Funded R&D: Evidence from the US Department of Energy". en. In: *American Economic Review* 112.7, pp. 2393–2423. ISSN: 0002-8282. DOI: 10.1257/aer.20210678. URL: https://pubs.aeaweb.org/doi/10.1257/aer.20210678 (visited on 11/09/2022).

National Research Council (Sept. 2014). *SBIR at the Department of Defense*. en. Pages: 18821. Washington, D.C.: National Academies Press. ISBN: 978-0-309-

30656-0. DOI: 10.17226/18821. URL: http://www.nap.edu/catalog/18821 (visited on 02/15/2023).

National Research Council (Apr. 2020). *Review of the SBIR and STTR Programs at the Department of Energy*. en. Pages: 25674. Washington, D.C.: National Academies Press. ISBN: 978-0-309-67159-0. DOI: 10.17226/25674. URL: https://www.nap.edu/catalog/25674 (visited on 02/01/2023).

Rassenfosse, Gaetan de, Adam Jaffe, and Emilio Raiteri (Aug. 2019). "The procurement of innovation by the U.S. government". en. In: *PLOS ONE* 14.8. Ed. by Cassidy Rose Sugimoto, e0218927. ISSN: 1932-6203. DOI: 10.1371/journal.pone.0218927. URL: https://dx.plos.org/10.1371/journal.pone.0218927 (visited on 02/10/2023).

Rathje, Jason (Aug. 2019). *Essays on Public Funding and For-Profit Innovation*. PhD Dissertation. Stanford University.

Santoleri, Pietro et al. (July 2022). "The Causal Effects of R&D Grants: Evidence from a Regression Discontinuity". en. In: *The Review of Economics and Statistics*, pp. 1–42. ISSN: 0034-6535, 1530-9142. DOI: 10.1162/rest_a_01233. URL: https://direct.mit.edu/rest/article/doi/10.1162/rest_a_01233/112422/The-Causal-Effects-of-R-amp-D-Grants-Evidence-from (visited on 08/25/2023).

Slavtchev, Viktor and Simon Wiederhold (Apr. 2016). "Does the Technological Content of Government Demand Matter for Private R&D? Evidence from US States". en. In: *American Economic Journal: Macroeconomics* 8.2, pp. 45–84. ISSN: 1945-7707, 1945-7715. DOI: 10.1257/mac.20130069. URL: https://pubs.aeaweb.org/doi/10.1257/mac.20130069 (visited on 01/30/2023).

Sun, Liyang and Sarah Abraham (Dec. 2021). "Estimating dynamic treatment effects in event studies with heterogeneous treatment effects". en. In: *Journal of Econometrics* 225.2, pp. 175–199. ISSN: 03044076. DOI: 10.1016/j.jeconom.2020.09.006. URL: https://linkinghub.elsevier.com/retrieve/pii/S030440762030378X (visited on 08/23/2023).

Appendix A: Appendix A: Tables and figures

Table 3.1. Patents

	Patents	Asinh Patents	Log Patents	Pat. Indicator
	(1)	(2)	(3)	(4)
Post × *Treat*	0.0177	0.0050	0.0746	0.0016
	(0.0181)	(0.0080)	(0.0834)	(0.0062)
R^2	0.25564	0.25160	0.49454	0.20778
Observations	76,849	76,849	3,118	76,849
Firm F. E.	✓	✓	✓	✓
Year F. E.	✓	✓	✓	✓

Table 3.2. Citation-Weighted Patents

	Pat. Citations	Asinh Pat. Citations	Log Pat. Citations
	(1)	(2)	(3)
Post × *Treat*	-0.7231	-0.0278	-0.1082
	(0.4816)	(0.0218)	(0.1878)
R^2	0.11069	0.19127	0.72820
Observations	76,849	76,849	2,424
Firm F. E.	✓	✓	✓
Year F. E.	✓	✓	✓

Table 3.3. Non-SBIR Contracts

	N. Contr. (1)	Asinh Contr. (2)	Log Contr. (3)	Contr. Indicator (4)
Post × *Treat*	0.7499	0.1498	-0.0324	0.0826
	(0.4109)	(0.0230)	(0.0428)	(0.0109)
R^2	0.53689	0.65236	0.74121	0.44261
Observations	76,849	76,849	12,585	76,849
Firm F. E.	✓	✓	✓	✓
Year F. E.	✓	✓	✓	✓

Table 3.4. Non-SBIR Contracts Dollar Value

	Contract Value (1)	Asinh Contract Value (2)	Log Contract Value (3)
Post × *Treat*	758,060.6	1.119	0.0312
	(200,860.2)	(0.1524)	(0.0829)
R^2	0.59099	0.47856	0.58954
Observations	76,849	76,849	12,235
Firm F. E.	✓	✓	✓
Year F. E.	✓	✓	✓

Table 3.5. Cumulative Patents

	Cum. Patents (1)	Asinh Cum. Patents (2)	Log Cum. Patents (3)
Post × *Treat*	0.3818	0.1166	0.1305
	(0.0886)	(0.0169)	(0.0326)
R^2	0.64339	0.74014	0.84709
Observations	76,849	76,849	16,324
Firm F. E.	✓	✓	✓
Year F. E.	✓	✓	✓

Table 3.6. Cumulative Citation-Weighted Patents

	Cum. Pat. Cit. (1)	Asinh Cum. Pat. Cit. (2)	Log Cum. Pat. Cit. (3)
Post × *Treat*	5.525	0.2018	0.1022
	(1.652)	(0.0381)	(0.0342)
R^2	0.74147	0.77377	0.96173
Observations	76,849	76,849	13,976
Firm F. E.	✓	✓	✓
Year F. E.	✓	✓	✓

Table 3.7. Phase 1 Grants

	N. Grants	Asinh N. Grants	Log N. Grants	Grant Indicator
	(1)	(2)	(3)	(4)
Post × *Treat*	-0.0257	-0.0335	0.0516	-0.0435
	(0.0387)	(0.0157)	(0.0218)	(0.0107)
R^2	0.55115	0.40265	0.57373	0.27400
Observations	76,849	76,849	11,566	76,849
Firm F. E.	✓	✓	✓	✓
Year F. E.	✓	✓	✓	✓

Table 3.8. Phase 2 Grants

	N. Grants	Asinh N. Grants	Log N. Grants	Grant Indicator
	(1)	(2)	(3)	(4)
Post × *Treat*	0.0184	-2.57×10^{-6}	0.0431	-0.0145
	(0.0226)	(0.0130)	(0.0213)	(0.0104)
R^2	0.37772	0.28796	0.51435	0.18974
Observations	76,849	76,849	9,124	76,849
Firm F. E.	✓	✓	✓	✓
Year F. E.	✓	✓	✓	✓

Table 3.9. Competed Non-SBIR Contracts

	N. Contr. (1)	Asinh Contr. (2)	Log Contr. (3)	Contr. Indicator (4)
Post × *Treat*	0.4570	0.1038	-0.0554	0.0622
	(0.3347)	(0.0209)	(0.0492)	(0.0105)
R^2	0.54796	0.62765	0.73207	0.40941
Observations	76,849	76,849	10,184	76,849
Firm F. E.	✓	✓	✓	✓
Year F. E.	✓	✓	✓	✓

Table 3.10. Competed Non-SBIR Contracts Dollar Value

	Contract Value (1)	Asinh Contract Value (2)	Log Contract Value (3)
Post × *Treat*	443,395.2	0.8390	0.0259
	(151,217.0)	(0.1481)	(0.0987)
R^2	0.55387	0.44035	0.56762
Observations	76,849	76,849	9,833
Firm F. E.	✓	✓	✓
Year F. E.	✓	✓	✓

Table 3.11. Noncompeted Non-SBIR Contracts

	N. Contr. (1)	Asinh Contr. (2)	Log Contr. (3)	Contr. Indicator (4)
Post × *Treat*	0.2930	0.0808	0.0027	0.0509
	(0.1212)	(0.0161)	(0.0614)	(0.0091)
R^2	0.34564	0.57646	0.69081	0.42283
Observations	76,849	76,849	5,756	76,849
Firm F. E.	✓	✓	✓	✓
Year F. E.	✓	✓	✓	✓

Table 3.12. Noncompeted Non-SBIR Contracts Dollar Value

	Contract Value (1)	Asinh Contract Value (2)	Log Contract Value (3)
Post × *Treat*	314,665.4	0.6584	0.0894
	(122,872.5)	(0.1171)	(0.1196)
R^2	0.30352	0.44717	0.64155
Observations	76,849	76,849	5,608
Firm F. E.	✓	✓	✓
Year F. E.	✓	✓	✓

Figure 3.6. Cumulative patents increase constantly

(a) Asinh of cumulative patents

(b) Asinh of cumulative citation-weighted-patents

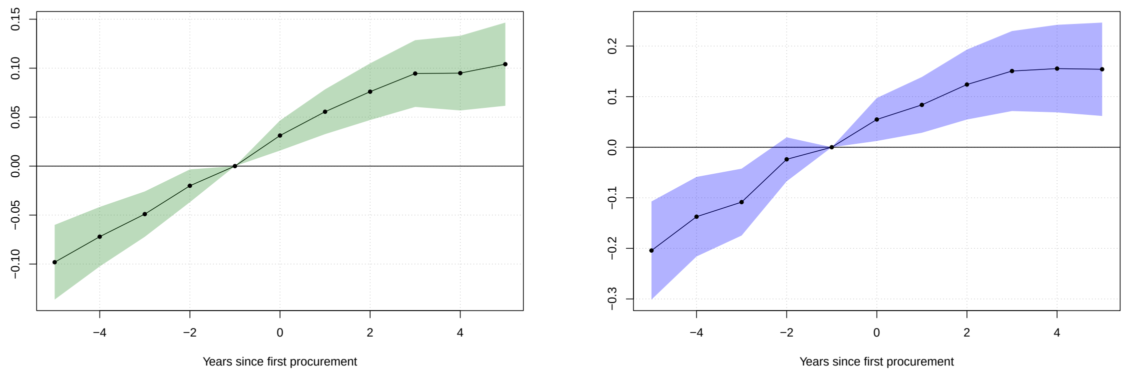


Figure 3.7. No more Phase I after Phase III

(a) Asinh of phase I grants number

(b) Grants indicator

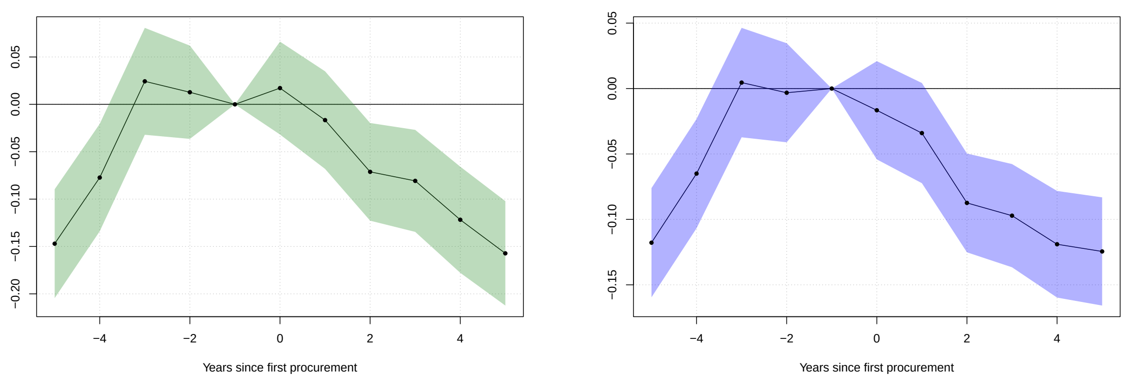
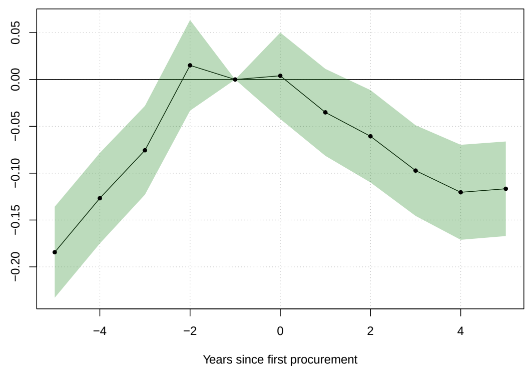


Figure 3.8. Phase II grants decrease after Phase III

(a) Asinh of phase II grants number



(b) Grants indicator

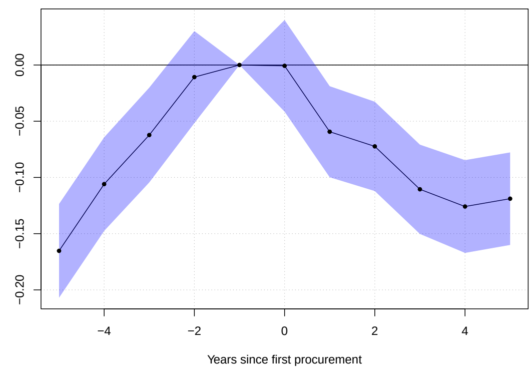


Figure 3.9. Effect of procurement on non sbir contracts indicator

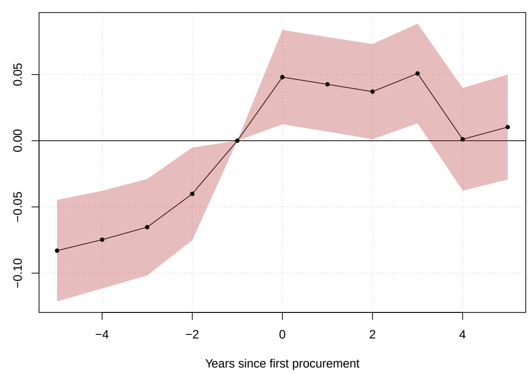


Figure 3.10. Competed and noncompeted procurement contracts indicators

(a) Competed contracts indicator (b) Noncompeted contracts indicator

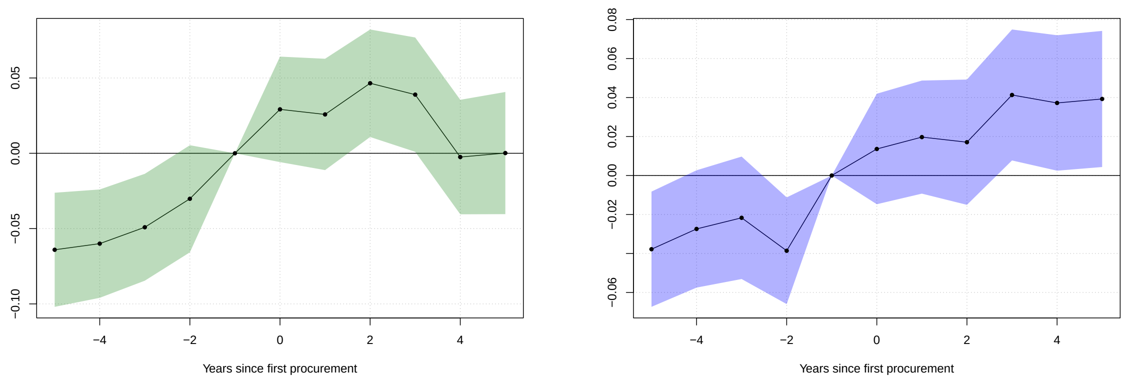
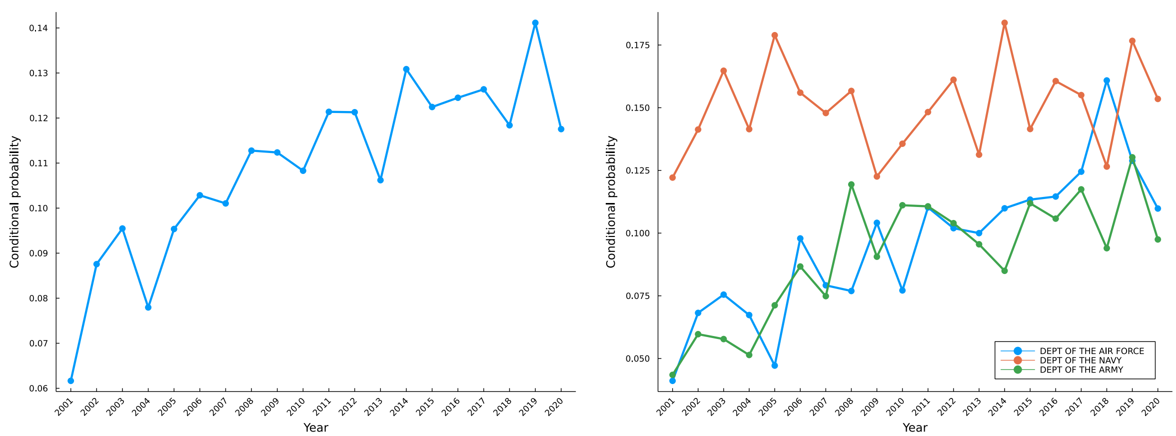


Figure 3.11. Conditional probabilities of reaching Phase III

(a) All Agencies (b) Top 3 Agencies



Appendix B: Appendix B: Robustness checks

Figure 3.12. Procurement does not affect patenting activity

(a) Asinh of patents (b) Asinh of citation-weighted-patents

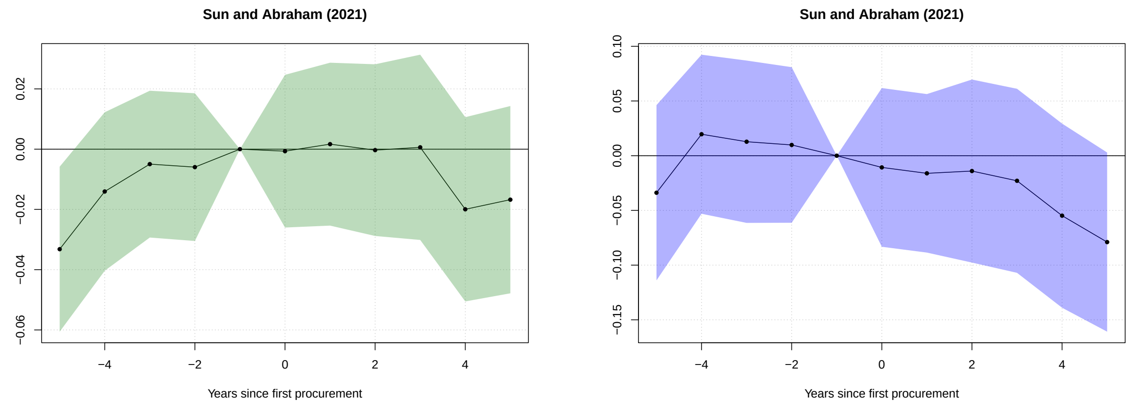


Figure 3.13. Non-SBIR contracts increase after Phase III

(a) Asinh of contracts number (b) Asinh of contracts dollar value

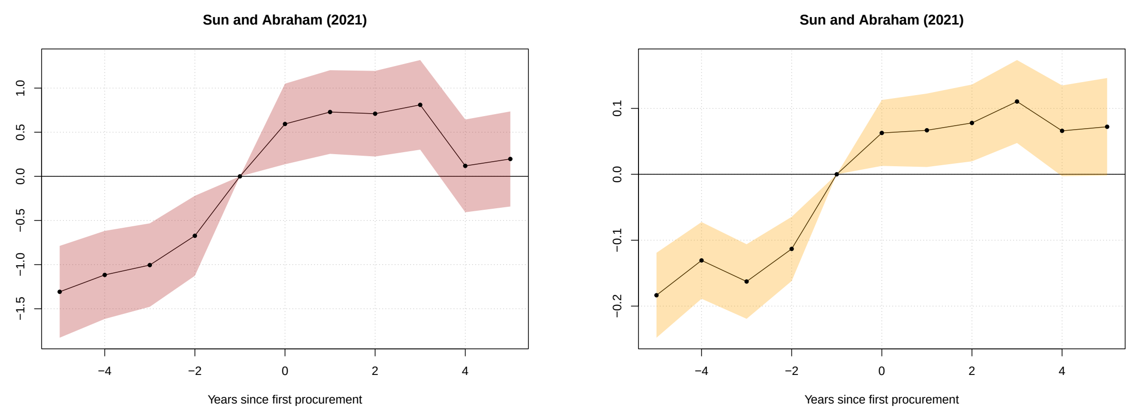
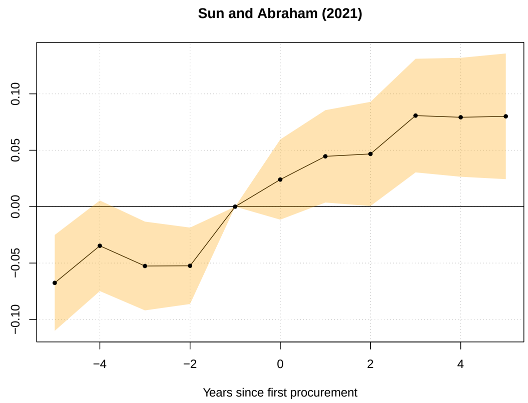


Figure 3.14. Noncompeted non-SBIR contracts increase more after Ph. III

(a) Asinh of contracts number



(b) Asinh of contracts value

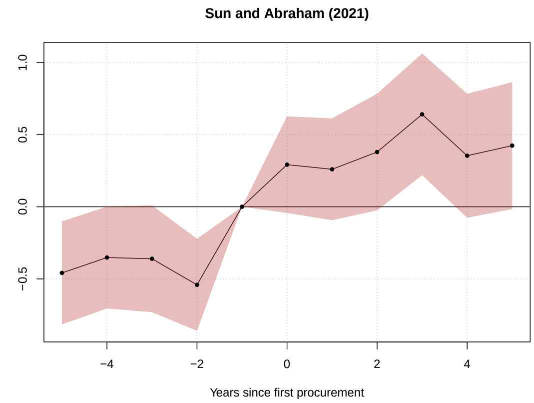
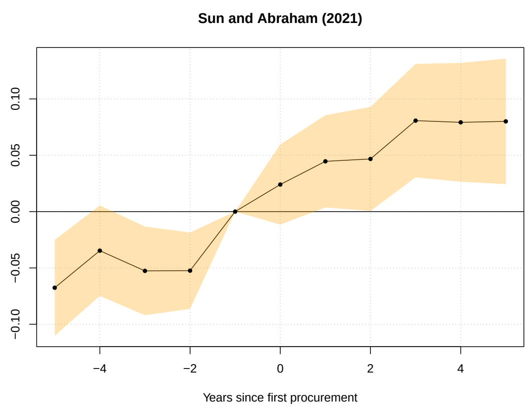


Figure 3.15. Noncompeted non-SBIR contracts increase more after Ph. III

(a) Asinh of contracts number



(b) Asinh of contracts value