Commercialization Infrastructure, External Finance, and Academic Entrepreneurship*

Xuelai Li† Xizhao Wang‡

July 2026

Abstract

We study how university commercialization infrastructure and external finance shape academic scientists' entry into entrepreneurship. We construct a novel panel linking a quasi-universe of U.S. academic scientists across fields to business registrations. Exploiting staggered technology transfer office (TTO) adoption across universities in the wake of the Bayh-Dole Act, we find that TTO establishment increases business formation by academic scientists. The effect is concentrated in counties with stronger banking markets and venture capital presence, and is larger among lower-income scientists. Nevertheless, scientists at universities that adopt TTOs experience a decline in publication output, potentially due to the reallocation of effort from academic research to commercialization. These findings show that commercialization infrastructure and external finance act as complements in translating academic research into entrepreneurial activity.

*We thank Juanita González-Uribe, Ryan Hill, Raj Iyer, Benjamin Jones, Ramana Nanda, and Dimitris Papanikolaou for their valuable comments and suggestions. All errors are our own.

†Imperial Business School, Imperial College London, x.li20@imperial.ac.uk.

‡Kellogg School of Management, Northwestern University, xizhao.wang@kellogg.northwestern.edu.

1 Introduction

University-based scientific research is widely recognized as a foundational input to commercial innovation and economic growth (Babina et al., 2023; Jaffe, 1989; Lerner et al., 2024; Nelson, 1986). Academic entrepreneurship represents an important pathway for academic commercialization, because university inventions are frequently embryonic; without dedicated development, these discoveries may never realize their commercial potential (Jensen and Thursby, 2001; Marx and Hsu, 2022). Moreover, this channel has profound implications for scientific career trajectories and economic dynamism. Thus, identifying the determinants of academic entrepreneurship is central to understanding how science translates into economic activity.

Despite the importance of academic entrepreneurship, measuring it systematically has been difficult. Entrepreneurial activity by individual scientists is hard to observe at scale, and the institutional and financial environments that shape these decisions usually evolve gradually rather than sharp changes. As a result, we know relatively less about which university scientists enter entrepreneurship and how commercialization infrastructure interacts with external finance in shaping this entry.

This paper addresses this gap by constructing novel individual-level data on a quasi-universe of U.S. academic scientists and their entrepreneurial activity. The data combine the historical source with nationwide business registration records and scientific publications. This linked dataset allows us to document the prevalence, career timing, field distribution, and publication activities of academic scientist-entrepreneurs, as well as the local external-finance environments in which they operate..

With the novel data, we begin by documenting a set of descriptive facts about academic scientist-entrepreneurs. Entrepreneurial entry is relatively rare, but it varies substantially across individuals, institutions, and fields. Scientists most commonly form their first business in mid-career, with entry peaking around ages in the late 30s to early 40s. Entry rates are higher among scientists with higher income and advanced educational attainment. Finally, there is pronounced variation across fields, with engineering and applied domains exhibiting substantially higher entry rates than life sciences and

social sciences in our sample period.

We then study the establishment of technology transfer offices (TTOs) as a broader institutional shift in universities' commercialization infrastructure. We ask how the emergence of TTOs encouraged academic entrepreneurship. We focus on the period when TTOs proliferated across U.S. universities, creating a setting in which commercialization infrastructure was still emerging rather than already fully institutionalized. This setting is important because existing datasets often observe scientists only after universities had established commercialization infrastructures.

Our empirical design exploits variation in the timing of TTO adoption across universities. While TTO formation was not randomly assigned, adoption was often determined at the state-system level for public universities, limiting the scope for individual scientists to drive timing. Consistent with this institutional feature, we obtain similar results when restricting the analysis to public universities. Moreover, the response is stronger among scientists with lower pre-adoption income, who were plausibly less likely to have the bargaining power or institutional influence needed to shape TTO establishment. These patterns weaken the concern that leading entrepreneurial scientists both pushed for adoption and subsequently increased firm formation.

We document three pieces of empirical results. First, TTO adoption increases business formation among academic scientists. Younger scientists and scientists with advanced degrees respond more strongly. The effects are stronger at universities with weaker pre-existing industry links but greater government-funded research, where scientists may have more commercializable ideas but fewer established commercialization channels. The results are robust across different specifications.

Second, TTOs and external finance are complements in facilitating academic entrepreneurship. The effect of TTO establishment on business formation is concentrated in counties with greater bank branch availability and an existing VC presence. Moreover, the effects are stronger among lower-income scientists, suggesting that the translation of these opportunities into entrepreneurial entry depends on access to capital. This pattern highlights the important role of TTOs as intermediaries connecting inno-

vative university projects with financing opportunities.

Lastly, TTO establishment is associated with a decline in scientists' publication output, suggesting a potential reallocation of effort from academic research toward commercialization. Importantly, this decline is not driven by scientists who eventually start businesses. Scientists who do not enter entrepreneurship experience a reduction in publications. Since observed business formation captures only realized entrepreneurial entry, the underlying reallocation of effort toward commercialization is broader.

We add to the literature on university innovation and commercialization. Prior work studies university patenting, licensing, and the organization of technology transfer (Jensen and Thursby, 2001; Mowery et al., 2001; Thursby and Thursby, 2002; Siegel et al., 2003). Related work examines academic entrepreneurship and the effects of commercialization incentives on scientific output (Azoulay et al., 2009; Toole and Czarnitzki, 2010; Marx and Hsu, 2022; Lerner et al., 2024). Closest to our paper, Stuart and Ding (2006) studies entrepreneurial entry among academic life scientists, and Hvide and Jones (2018) examines how ownership rights over university inventions shape university innovation and entrepreneurship. We differ by constructing a broad U.S. scientist-level panel, measuring business formation, and combining staggered TTO establishment with variation in external capital supply to study which scientists are drawn into entrepreneurship.

We contribute to the literature on financing innovation and technology entrepreneurship (Lerner and Nanda, 2020). This literature highlights the importance of external finance, through both debt and equity markets, for innovation by small firms (Black and Strahan, 2002; Chava et al., 2013; Hombert and Matray, 2017; Kerr and Nanda, 2009; Robb and Robinson, 2014). We show that commercialization infrastructure and external finance are complement in shaping the university academic entrepreneurship, as it can be embedded into the local finance network (Hochberg et al., 2007; Fehder et al., 2025).

We also contribute to the literature on selection into entrepreneurship (Bernstein

et al., 2022; Gottlieb et al., 2022; Hacamo and Kleiner, 2022; Hombert et al., 2020; Nanda and Sørensen, 2010; Schoar, 2010; Wallskog, 2025). Related work studies selection into VC-backed entrepreneurship, investor selection among startups, new-firm financing at entry, and the geography of high-growth entrepreneurship (Ewens et al., 2024; Kaplan et al., 2009; Guzman and Stern, 2020; Chen and Ewens, 2025; Jang and Kaplan, 2025). We bring this perspective to university scientists by showing that TTO establishment increases their transition into entrepreneurship, especially in counties with stronger venture-capital and banking markets. Our findings suggest that selection into science-based entrepreneurship is shaped jointly by university commercialization infrastructure and local external capital supply.

The rest of this paper is organized as follows. Section 2 provides an overview of the historical context of TTOs. Section 3 describes the data sources and presents descriptive statistics on the scientists included in the analysis. Section 4 provides an identification strategy. Section 5 examines the reduced-form relationship between the TTOs' establishment and business formation. Section 6 concludes.

2 Institutional Background

U.S. universities rely substantially on federal funding to support their research activities (Babina et al., 2023). The contemporary landscape of academic entrepreneurship in the U.S. is closely tied to the Bayh-Dole Act of 1980. Before Bayh-Dole, ownership and control rights over inventions arising from federally funded research were often fragmented across federal agencies, and universities faced limited and uncertain rights to patent, license, or otherwise commercialize these inventions (Mowery et al., 2001; Jensen and Thursby, 2001).

The Bayh-Dole Act changed this environment by allowing universities and other nonprofit institutions to retain rights to inventions developed with federal funding (Henderson et al., 1998; Sampat, 2006). In the years following Bayh-Dole, many universities expanded their commercialization activities by establishing TTOs. Their growth

marked a shift from more decentralized and informal commercialization arrangements toward a more centralized and professionalized model, in which licensing officers, attorneys, and business managers served as intermediaries between academic scientists and commercial markets (Siegel et al., 2003; Thursby and Thursby, 2002).

Whether such commercialization infrastructure can activate entrepreneurial activity among scientists, and which scientists are drawn into entrepreneurship, is an important question for innovation policy. Yet answering this question is difficult because institutions, scientists, and entrepreneurial outcomes co-evolve, making it hard to isolate the effect of any single policy lever. We exploit the historical establishment of TTOs at the U.S. universities as a unique setting for studying this question.

The early TTO function is essentially a broker that benefits universities by providing funds for research and benefits industry by providing research findings but not yet equity holders (Feldman et al., 2002). By serving as a user broker, TTOs align laboratory outputs with market demand, ensuring the structural stability of the transfer mechanism (Herdendorf and Linsteadt, 1982). Most TTOs focus on managing information clearinghouses for available patents and designing flexible intellectual property policies to facilitate industrial access. Some TTOs oversee programs such as computerized risk-capital networks and university-affiliated business incubators (Hisrich and Smilor, 1988).

The establishment of a TTO changed not only universities' emphasis on commercialization, but also the commercialization infrastructure through which scientists could pursue patenting, licensing, and entrepreneurial activity. During this period, TTOs were adopted in a staggered way across universities, creating temporal and cross-sectional variation that can be exploited for identification. Importantly, the scientists working at these universities during this era can now be linked to their entrepreneurial activities and academic publications through our digitized historical records. This enables the first comprehensive, individual-level analysis of how commercialization infrastructure shapes scientists' selection into entrepreneurship.

3 Data and Descriptive Statistics

3.1 Data Sources

Scientists Data The primary data source is the National Register of Scientific and Technical Personnel (NRSTP), obtained from the National Archives Access to Archival Databases. Originally established by the National Science Foundation (NSF) to identify specialized professionals for national emergencies, the NRSTP evolved into a comprehensive statistical resource for science policy, covering diverse disciplines in the physical, biological, and social sciences. While the register contains eight waves spanning 1954 to 1970,[1] This paper utilizes the 1964–1968 waves because they include the city-level geographic identifiers necessary for business registration, patent, and paper linkage, and also employer information for TTO linkage. The dataset is highly representative of the high-skilled workforce; the NSF reported that the 1964 wave captured over 90% of all U.S. science doctorates.

Data processing involved extracting information from digitized numerical codes and mapping them to original codebooks. The mapping between codes and descriptions was manually cleaned to ensure accuracy. In instances where original scans were faint, transformer-based language models were employed to reconstruct text based on standard scientific nomenclature. The final panel includes individual-level variables such as education, employment specialty, institutional affiliation, and self-reported income (Li, 2026).[2]

Figure 1a presents the original NRSTP codebook, which was utilized to digitize and decode the raw data entries. As illustrated in Figure 1b, the analytical sample is restricted to university employees to maintain a specific focus on university scientists and TTO.

[1]The waves include 1954, 1958, 1960, 1962, 1964, 1966, 1968, and 1970. Started in 1973, the Survey of Doctorate Recipients (SDR) became the primary tool for tracking Ph.D.-level scientists.

[2]Similar historical bibliographies, such as American Men of Science, do not include income information for scientists.

U.S. Business Registration Data While self-employment is occasionally equated with entrepreneurship, this study aligns with established literature by utilizing incorporated business owners as the primary metric for entrepreneurs. This distinction is particularly relevant when focusing on ventures that commercialize innovative ideas, as incorporation serves as a more precise indicator of such activity (Levine and Rubinstein, 2017).

Data regarding U.S. business registrations, which are maintained at the state level by each Secretary of State, were obtained through OpenCorporates. This organization aggregates these disparate records into a centralized download package, and the version utilized for this research reflects data as of January 2025. This study incorporates data from all U.S. jurisdictions. However, it should be noted that state registry records do not include details on bankruptcies or other forms of litigation; such information must be retrieved through specialized litigation searches.

The comprehensive business registry from OpenCorporates encompasses 76 million businesses across the United States, providing critical variables such as incorporation and dissolution dates, registration addresses, and jurisdictional status.[3]

TTO Data The timing of TTO formation comes from the Association of University Technology Managers (AUTM), which reports information on U.S. university technology transfer offices, including their year of establishment.[4] We combine these records with the scientists data from NRSTP by linking institutions based on university names. Because institution names are not standardized across the two sources, the merge required manual checks. In particular, we reviewed alternative assignee names that correspond to the same university and collapsed them into a common institutional identifier.[5]

[3] Although Census data, such as the Longitudinal Business Database (LBD), contain information on business registrations, the LBD begins only in 1976, yet our sample starts from 1970. OpenCorporates offers the most consistent publicly available source of historical U.S. business registration records.

[4] The version utilized for this research is from 2020 AUTM Licensing Activity Survey.

[5] In some cases, AUTM reports TTO information at the university-system level rather than for individual campuses. For instance, campuses such as UCLA, UCSD, UC Merced, UC Davis, UC Berkeley, UC Irvine, UC Riverside, UCSF, UC Santa Barbara, and UC Santa Cruz are all recorded under the University of California System in the AUTM data.

Publication Data Scientific productivity is measured using SciSciNet (Lin et al., 2023), a large-scale open data resource that aggregates publication and citation data from the Microsoft Academic Graph. For each paper, SciSciNet provides comprehensive information such as bibliographic information, author identifiers, affiliated institutions, publication year, field classifications, and citation links.

3.2 Variable Construction

3.2.1 Dependent Variables

Business Registrations Figure 2 shows the matching procedure used to identify scientist entrepreneurship. Scientists' names and locations from NRSTP are linked to business registrations from OpenCorporates based on officer or director names and registered addresses. The matching procedure follows the name-and-address linkage approach in Engelberg et al. (2026), but applies it to a distinct historical population of U.S. academic scientists and constructs a new scientist-level dataset.

Business addresses within the OpenCorporates dataset were regularized to extract ZIP codes, which were subsequently mapped to their respective counties. For the primary matching period (1945–1990), approximately 57% of the 14,495,168 firms in the dataset contained registered address information. Among these firms, 84% initially included a ZIP code. To improve data density, the OpenStreetMap API was employed to retrieve ZIP codes for the remaining 16% of non-standard addresses. This process successfully geocoded an additional 155,820 addresses, increasing the overall ZIP code coverage to 93%.

To distinguish between individual and corporate entities within the officer records, the spaCy library was utilized for natural language processing. The classification process involved identifying entities labeled as "PERSON" by the NLP model to determine if an officer's name represented a human individual. This analysis revealed that 88.08% of the officer names are classified as human names rather than company names, providing a clearer understanding of the composition of the entities recorded in the dataset.

The OpenCorporates data were subsequently mapped to the NRSTP datasets using

a combination of individual names and county FIPS codes. This matching procedure was restricted to businesses established between 1945 and 1990 to align with the demographic profile of the scientists, most of whom were born in the 1930s. Limiting the timeframe accounts for the decreased likelihood of business formation at advanced ages and mitigates the risk of false-positive matches between different individuals sharing the same name.

Publications The matching is restricted to English-language journal publications from SciSciNet and utilizes institutional affiliation FIPS codes to verify identities. To minimize false positives, records appearing after 2005 are excluded given the cohort's age.

3.2.2 Independent Variable

TTO Establishment Our primary sample comprises scientists identified in the NRSTP survey between 1964 and 1968. Because the survey ceased tracking specific employer transitions after 1970, we assign each scientist's institutional affiliation based on their recorded university during the 1964–1968 window and merge this with the timing of TTO establishments.

A limitation of this approach is the potential for unobserved inter-university mobility; in cases where scientists switched institutions after 1968, we would fail to update their treatment status. However, since our business formation data is also geographically anchored to the scientist's 1960s residence, any migratory entrepreneurship would result in the scientist dropping out of the detectable outcome pool. This results in a downward bias of the TTO's impact, as we only observe businesses started by scientists who remained in their original geographic area. If a TTO inspires a scientist to move and start a company elsewhere, our model records that as a "zero" for business formation, even though the TTO may have been the catalyst.

Age Since the NRSTP dataset lacks the explicit birth date information, birth years are estimated based on the recorded Year of Highest Degree and Level of Highest Degree. This methodology assumes that scientists typically reach educational milestones

at specific ages: 30 for a doctoral-level degree (e.g., PhD or MD), 25 for a Master's, and 22 for a Bachelor's. A proxy for the birth year is derived by subtracting these age constants from the graduation year. Figure 3 shows the ages of scientists at the time they started their businesses. The cohort has a mean age of 39.73 and a median age of 39, with the 10th percentile (P10) at 30.0 and the 90th percentile (P90) at 48.4.

Gross Income Annual gross income is self-reported in the NRSTP. It combines both base salary and other income. Since scientists in the sample worked at universities, they were typically paid on a nine-month basis, so using annual gross income rather than base salary is more appropriate.

Field To compare individual expertise with institutional strengths, we proxy for research field using the major of the highest degree, which is standardized to align with the university spending categories in the 1972 Higher Education Research and Development (HERD) survey.[6] Utilizing the codebook's taxonomic hierarchy, granular subfields—such as analytical chemistry, numerical methods, and electrical engineering—were aggregated into broader disciplinary classifications, including Physical Sciences, Math/CS, Engineering, etc. This mapping ensures consistency between educational backgrounds and institutional expenditure data.

R&D Funding The institutional research environment is further characterized by the shares of different types of funding in total R&D income, utilizing data from the HERD survey. Total R&D income is calculated as the aggregate of federal, industry, institutional, and state/local government funding as shown in Panel A of Table A1. For instance, the proportion of industry-sponsored funding serves as a proxy for the university's integration within the business ecosystem. High levels of industry R&D may indicate established pathways for commercialization, potentially reducing a scientist's reliance on a formal TTO for entrepreneurial activities.

[6]The data are available beginning in 1972, when the survey was conducted as the Survey of R&D Expenditures at Universities and Colleges; see https://ncses.nsf.gov/explore-data/microdata/higher-education-research-development

External Financing Bank branch data are drawn from the Federal Deposit Insurance Corporation (FDIC) and measure the number of active branches at the county level. As Figure A3 shows, the geographic distribution of bank branches is largely stable over the sample period. We therefore classify counties into high- and low-bank-branch groups based on their 1970 branch counts, and use this classification as a proxy for local credit supply facing scientists who may enter entrepreneurship (Black and Strahan, 2002; Célerier and Matray, 2019). Likewise, we obtain VC data from VentureXpert and measure local VC activity using the number of VC deals at the county level prior to 1970.

3.3 Descriptive Statistics

Our analysis draws on a panel of 109,584 scientists observed from 1970 to 1990. Of these, 50,171 are treated, meaning that their university establishes a TTO during the sample period, and 59,413 serve as never-treated controls, meaning that their university does not establish a TTO during our observation window.

Table 1 reports the distribution of scientists across top 20 employing institutions. Employment is highly concentrated in a small number of large public research universities and elite private universities. The University of California alone accounts for 8.25% of the sample. The presence of state universities is consistent with their larger scale and broader hiring capacity, while elite private institutions reflect high research intensity.[7]

Table 2 reports business formation rates across scientist characteristics, comparing the full sample with the subset of never-treated individuals. Several consistent patterns emerge across the two samples, with broadly similar magnitudes, suggesting that selection into treatment is unlikely to fully drive the observed differences.

The descriptive evidence suggests a life-cycle pattern where formation propensity diminishes with age. In the full sample, the youngest cohort (<45) exhibits the highest formation rate at 2.57%, falling to 1.93% for those over 60. Interestingly, while the

[7]Scientists affiliated with the Massachusetts Institute of Technology (MIT) are excluded from the sample, as MIT is always treated: its TTO was established prior to 1970, the start of the sample period.

never-treated sample follows a similar downward trend, the gap between the youngest and middle cohorts is much narrower (2.13% vs 2.23%), suggesting that the establishment of TTO may disproportionately stimulate formation among younger individuals. These findings are robustly supported by the logistic regressions in Figure 4, where the coefficients for "Age: 45-60" and "Age: Over 60" are both negative and statistically significant. This confirms that, relative to the omitted baseline of "Age: Below 45", advancing age serves as a significant structural barrier to formation activity.

Advanced degree attainment is a positive predictor of business formation. Data from the full sample shows that PhD or MD holders have a business formation rate of 2.42%, over non-PhD holders (2.03%). In the never-treated sample, this premium persists but is slightly attenuated (2.15% vs 2.02%), indicating that higher educational attainment may enhance an individual's responsiveness to the establishment of TTO.[8] The results in Figure 4 reinforce this, showing a positive and significant log-odds ratio for the "PhD or MD Degree" variable.

The descriptive statistics also show a significant relationship between income and business formation rates, likely reflecting the impact of budget constraints. High-income individuals in the full sample achieve a formation rate of 2.70%, compared to only 1.90% for the low-income group. This 80 basis point disparity remains substantial in the never-treated sample (2.47% vs 1.79%), suggesting that budget constraints are a primary driver of business formation regardless of treatment status. Figure 4 benchmarks this effect through the "Low Local Income" coefficient, which is negative and highly significant. This provides strong empirical evidence that individuals in lower-income scientists face higher barriers to entry.

Business formation activity varies considerably across academic fields. Engineering records the highest rates in both the full (3.72%) and never-treated (3.54%) samples. Across almost all categories, formation rates are lower in the never-treated sample, with the delta being particularly pronounced in Mathematical & Computer Sciences (a drop

[8]The sample includes many individuals who are reported as non-PhD holders in the original survey. Manual checks suggest that some later obtained PhD degrees, while others held PhD degrees but reported a master's degree as their highest degree. To preserve the original survey records, we do not amend the education variable using manually collected information.

from 2.36% to 1.98%).

Figure 4 includes a proxy for "Low Field R&D Spending," which yields a negative coefficient. Scientists in fields that receive higher R&D spending at their employer's university show significantly higher business formation propensity compared to those in lower-spending fields, suggesting that institutional research support is a vital precursor to business formation.

4 Identification Strategy

We exploit the staggered establishment of TTOs across universities as a natural experiment. Treatment is defined at the university level: scientist i is treated when her university establishes a TTO. The staggered timing of TTO adoption across universities provides the identifying variation. We estimate treatment effects using the CSDID estimator for staggered adoption designs (Callaway and Sant'Anna, 2021). Let G_i denote the first year in which scientist i is treated, with $G_i = 0$ for never-treated individuals. For each treatment cohort g and calendar year t, we estimate the cohort-time average treatment effect:

$$ATT(g,t) = E\big[Y_{it}(1) - Y_{it}(0) \mid G_i = g\big], \quad t \geq g. \tag{1}$$

This estimator is well suited to settings with staggered treatment timing and heterogeneous treatment effects, and avoids the well-known biases of conventional two-way fixed effects estimators in such environments.

Identification requires that, absent treatment, untreated potential outcomes would have evolved in parallel across treated and comparison groups. We assess this assumption using pre-treatment event-study coefficients. In the baseline specification, we use never-treated scientists as the control group; in robustness checks, we instead use not-yet-treated scientists. We estimate $ATT(g,t)$ using outcome regression with a universal base period and then aggregate cohort-time effects using the simple aggregation procedure to obtain an overall average treatment effect for each outcome. We also report

event-study relative to TTO establishment to examine dynamic effects.

Several potential threats to identification merit discussion. First, universities that establish TTOs earlier may differ systematically from those that do not. The Callaway and Sant'Anna (2021) framework accommodates treatment effect heterogeneity across cohorts and therefore does not rely on homogeneous treatment effects across adoption cohorts. Second, other policy changes – such as shifts in federal research funding or changes in patent law – may coincide with TTO establishment. However, TTO adoption dates vary widely across universities, making it less likely that any single contemporaneous shock explains the results. As an additional check, we also estimate dynamic treatment effects using the Sun and Abraham (2021) interaction-weighted event-study estimator, and obtain similar patterns.

We next examine heterogeneous treatment effects along three dimensions: scientists' characteristics, university characteristics, and external financing environment. For each subgroup, we run separate CSDID estimations with both treated and control scientists restricted to the relevant subgroup. We examine heterogeneity in both business formation and publication output to assess whether the benefits and costs of TTO establishment accrue to the same or to different groups.

5 Empirical Results

5.1 TTO Establishment and Academic Entrepreneurship

Understanding who becomes an entrepreneur in this setting is important because academic entrepreneurs are not randomly drawn from the scientific workforce: they differ in age, skills, opportunity costs, and so on. These selection margins matter. If TTOs primarily induce entry by scientists with commercially valuable ideas but limited access to capital or support, then they can expand the set of innovations that reach the market. By contrast, if entry is concentrated only among scientists who would have commercialized anyway, the role of TTOs is more limited.

We first show that the effect of TTO establishment on business formation in the

full sample. Table 3 shows that TTOs have positively significant effect on business formation of university scientists. The estimates are highly stable across alternative estimators and control groups. Using the Callaway and Sant'Anna (2021) estimator, the aggregate ATT is 0.0012 with never-treated controls and 0.0013 with not-yet-treated controls. The Sun and Abraham (2021) interaction-weighted estimator gives a similar estimate of 0.0011. All three estimates are statistically significant at the 1 percent level.[9]

Figure 5 presents the corresponding event-study estimates. Across specifications, the pre-treatment coefficients are close to zero, while the effect on business formation rises after TTO establishment. The dynamic evidence therefore supports the interpretation that TTOs increased entrepreneurial entry among academic scientists.

5.1.1 Robustness Checks

Specifications Our main results so far treat business formation as a yearly count and examine how TTO establishment changes the average number of new ventures per scientist-year. A complementary framing is to treat entrepreneurship as a one-time *transition*: how does TTO establishment affect the rate at which a scientist moves from non-entrepreneurship into entrepreneurship?

To answer this question, we construct two outcomes from the panel. The first is a cumulative-ever indicator, defined as an absorbing-state variable that equals one once scientist i has started any business by year t. We re-estimate the same staggered-DID estimators using this outcome.

The second outcome is the time to first business formation, right-censored at 1990 for scientists who never enter entrepreneurship. We fit a Cox proportional hazards model in counting-process form, $\lambda_i(t) = \lambda_0(t) \exp(\beta \cdot PostTTO_{it})$, with a time-varying treatment indicator $PostTTO_{it}$ that switches to one once the scientist's employer has

[9]One caveat is that these estimates may understate the effect of TTO establishment on younger cohorts. The full sample is based on a snapshot of scientists observed around the 1960s and tracks their subsequent activities. As a result, the analysis does not fully capture new scientists who entered universities after the baseline sample was constructed. This is relevant when younger scientists are more responsive to the commercialization opportunities created by TTOs. Therefore, the estimates in Table **??** should be interpreted as the effect on a stock of scientists, rather than the full effect on the evolving scientific workforce.

established a TTO, and remains zero for never-treated scientists.

The estimates using the cumulative-ever indicator and the hazard model are reported in Table A3, and the corresponding event-study dynamics are reported in Figure A1 and A2. All results support that TTO establishment significantly increases scientists' entry into entrepreneurship.

Public Universities As another additional check, we restrict the analysis to scientists at public universities. This restriction is useful because TTO adoption at public universities was often shaped by state-system decisions, reducing the scope for individual scientists to influence the timing of adoption. Table A4 shows that the results are similar. This provides additional evidence that the main result is not driven by reverse causality, whereby scientists' entrepreneurial activity induces TTO establishment.

5.1.2 Scientist Characteristics

Table 4 examines heterogeneity by scientist characteristics. Panel A shows that the response is strongest among younger scientists. The ATT is 0.0023–0.0024 for scientists younger than 45, declines to 0.0007–0.0008 for those aged 45–60, and is smaller and statistically insignificant for scientists older than 60. This finding aligns with the descriptive evidence that scientist-entrepreneurs tend to be younger than the average faculty member. This is consist with MIT alumni survey data that shows median age of first time entrepreneurs was around 40 for those graduated in the 1950s (Hsu et al., 2007).

Panel B shows that the effect is concentrated among scientists with advanced degrees. The ATT is 0.0014–0.0016 for PhD or MD scientists and statistically significant across control groups, compared with smaller and insignificant estimates for scientists without advanced degrees. This result is consistent with descriptive evidence showing that scientist-founded startups are disproportionately led by individuals with advanced training.

Panel C splits scientists by income relative to peers in the same county. The estimates are positive for both groups, with larger point estimates for scientists below the

within-county median. This pattern indicates that the entrepreneurial response might be larger to low-income scientists.

Overall, the heterogeneity by scientist characteristics suggests that TTO-induced entrepreneurship is concentrated among younger scientists and scientists with stronger formal training. The income results provide more suggestive evidence that lower-income scientists may also respond more strongly. Taken together, these patterns are consistent with TTOs lowering commercialization frictions for scientists who are well positioned to act on new entrepreneurial opportunities.

5.1.3 University Characteristics

In line with the existing literature showing that university characteristics are important determinants of commercialization (e.g., Lerner et al., 2024), we examine heterogeneity in the relationship between TTO establishment and business formation by university characteristics.

Table 5 Panel A splits universities by pre-existing federal funding share. The effect is concentrated among universities with above-median federal funding shares: the ATT is 0.0010 using never-treated controls and 0.0015 using not-yet-treated controls, while the estimates for low-federal-funding universities are close to zero and statistically insignificant. Consistent with the framework, in which TTOs lower the effort required to commercialize the existing stock of scientific ideas rather than create new ideas directly, the effect is concentrated among universities receiving above median federal funding. This pattern suggests the complementarity between federal research funding and commercialization infrastructure: public funding generates a stock of quality opportunities, while TTOs lower the effort scientists must devote to converting them into business formation.

Panel B splits universities by pre-existing industry funding share. The estimates are small and statistically insignificant among universities with high industry funding shares. For universities with low industry funding shares, the point estimates are larger, although the effect is statistically significant only when using not-yet-treated

controls. This provides suggestive evidence that TTOs may matter more where universities had weaker pre-existing links to industry.

Panel C splits universities by own-field R&D spending. The evidence is less robust across control groups. Using never-treated controls, the effect is concentrated among universities with below-median own-field R&D spending. Using not-yet-treated controls, however, the estimates are positive, statistically significant, and similar in magnitude for both high and low own-field R&D universities. Thus, the own-field R&D split does not provide clear evidence of a stable heterogeneity pattern. One interpretation is that own-field R&D captures two offsetting forces in the framework. Stronger research environments tend to have lower pre-existing effort thresholds, so the TTO-induced reduction is smaller; but they also raise the commercial value of ideas, and hence the payoff to reallocating effort. These offsetting forces can explain why the heterogeneity by own-field R&D spending is less sharp than the heterogeneity by prestige or industry funding.

Overall, the university-level heterogeneity suggests that TTO effects are stronger in settings where universities combine substantial federally funded research with weaker pre-existing industry ties. The evidence is consistent with TTOs being more consequential when they introduce new commercialization capacity, rather than merely formalizing channels that were already present.

5.2 Complementarities Between TTOs and External Finance

A key argument in the literature is that the realization of commercial opportunities depends on access to external capital (Kerr and Nanda, 2015). The evidence in Section 5.1 shows that lower-income scientists are more responsive to TTO establishment, suggesting that financial constraints may be a barrier to entry. This motivates examining whether the effect of TTOs is stronger in counties with greater bank availability or VC presence. If TTOs and external finance are complements, then commercialization infrastructure should have larger effects.

Figure 6 plots county-level VC activity against the business formation rate of scien-

tists. Counties with greater VC activity exhibit higher rates of scientist business formation. This pattern is descriptive and does not admit a causal interpretation, but it motivates the analysis below: if external finance complements commercialization infrastructure, the effect of TTO establishment should be stronger in financially developed local markets.

Table 6 examines heterogeneity by county-level financial development. Panel A splits counties by bank branch availability in 1970. The effect of TTO establishment is concentrated in counties in the top quartile of bank branch availability: the ATT is 0.0013 using never-treated controls and 0.0015 using not-yet-treated controls. In counties with fewer bank branches, the estimates are smaller, around 0.0003, and statistically insignificant. Panel B shows a similar pattern for venture capital. In counties with any recorded VC activity, the ATT is 0.0015–0.0020 and statistically significant at the 1 percent level; in counties without VC activity, the point estimates are close to zero.

Table 7 further examines whether the role of local finance varies with scientist income. Panel A interacts within-county income with bank branch availability. Among lower-income scientists, the ATT is larger in counties with more banks branches. Among higher-income scientists, the estimates are still positive in counties with more banks branches but close to zero in counties with fewer bank branches. This pattern suggests that credit availability is associated with stronger TTO effects, with particularly robust responses among lower-income scientists.

Panel B shows a sharper pattern for venture capital. For lower-income scientists in counties with VC presence, the ATT is 0.0014, yet for those in counties without VC presence, the estimates are close to zero. Higher-income scientists also respond more in VC counties than in non-VC counties, although the estimates are smaller and insignificant. Because the baseline business formation rate is lower among low-income scientists, these estimates imply that the relative effect of external finance is substantially larger for this group.

The results in this section suggest that TTOs and local finance are complements. TTOs may lower the costs of commercialization, but the translation of those opportu-

nities into observed firm formation appears strongest in counties with greater access to bank credit or venture capital. External finance can amplify the effect of TTO establishment, especially for lower-income scientists. The results are also consistent with historical evidence that TTOs play an intermediary role in connecting innovative university projects with financing opportunities.

5.3 Publication Trade-offs and the Cost of Commercialization

Apart from business formation, TTOs may affect scientists' broader allocation of effort between academic research and commercial activity. Since publication is a common measurable productivity of scientists (Kim and Moser, 2025), we next examine publication output as an additional margin. The effect is theoretically ambiguous. Entrepreneurship may either crowd out academic research by reallocating scientists' effort away from publication-oriented work (Aghion et al., 2008), or crowd in research productivity by generating new ideas and (Fini et al., 2021). This section therefore examines whether TTO-induced business formation is associated with an increase or decline in scientists' publication output.

Table A6 reports the effect of TTO establishment on publication output in the full sample. The estimated ATT ranges from -0.0066 to -0.0074 publications per scientist-year, with all estimates statistically significant at the 1 percent level. Figure A4 presents the corresponding event-study estimates. Together, the Table A6 and Figure A4 indicate that TTO establishment is followed by a decline in average publication output among scientists at treated universities.

This decline should be interpreted as an institutional-level trade-off, rather than as the publication cost borne by scientists who start firms. The treatment varies at the university level, and the estimate captures the average response of scientists exposed to TTOs. Consistent with this interpretation, Appendix Table A7 shows no detectable decline in publication output around scientists' first observed business formation among TTO-treated scientists.

The evidence suggests a broader reallocation of effort within treated universities.

TTOs may draw a wider set of scientists into commercialization-related activities, such as invention disclosure, patenting, licensing discussions, and startup exploration, even when these activities do not culminate in observed firm formation. The cost of TTO establishment is therefore a diffuse reduction in measured academic output, rather than a direct publication penalty for the subset of scientists who become entrepreneurs.

6 Conclusion

This paper studies how commercialization infrastructure shapes the selection of university scientists into entrepreneurship. Using newly linked individual-level data on academic scientists, we document that TTO establishment increases entrepreneurial entry among university scientists. The response is stronger among younger scientists, scientists with advanced degrees, and scientists located at universities with weaker pre-existing industry links but greater government-funded research. These patterns suggest that TTOs are particularly important where commercializable ideas exist but commercialization channels are less developed.

The results also show that commercialization infrastructure does not operate in isolation. The effect of TTO establishment is concentrated in counties with stronger local banking markets and VC presence, and is larger among lower-income scientists. This evidence suggests that TTOs and external finance are complementary inputs into academic entrepreneurship. At the same time, TTO establishment is associated with lower publication output, indicating that commercialization may reallocate scientific effort away from publication-oriented research.

The evidence points to a broader policy lesson: effective academic commercialization requires coordination between university-level commercialization infrastructure and local financial markets. Policies that expand university commercialization capacity may have limited effects if they are not paired with access to entrepreneurial finance. Conversely, increasing the supply of capital may be insufficient when universities lack institutions that identify, disclose, and develop commercial opportunities.

References

Aghion, Philippe, Mathias Dewatripont, and Jeremy C. Stein, 2008, Academic freedom, private-sector focus, and the process of innovation, *RAND Journal of Economics* 39, 617–635.

Azoulay, Pierre, Waverly Ding, and Toby Stuart, 2009, The impact of academic patenting on the rate, quality and direction of (public) research output, *Journal of Industrial Economics* 57, 637–676.

Babina, Tania, Alex Xi He, Sabrina T. Howell, Elisabeth Ruth Perlman, and Joseph Staudt, 2023, Cutting the innovation engine: How federal funding shocks affect university patenting, entrepreneurship, and publications, *Quarterly Journal of Economics* 138, 895–954.

Bernstein, Shai, Emanuele Colonnelli, Xavier Giroud, and Benjamin Iverson, 2022, Who creates new firms when local opportunities arise?, *Journal of Financial Economics* 143, 107–130.

Black, Sandra E., and Philip E. Strahan, 2002, Entrepreneurship and bank credit availability, *Journal of Finance* 57, 2807–2833.

Callaway, Brantly, and Pedro HC Sant'Anna, 2021, Difference-in-differences with multiple time periods, *Journal of Econometrics* 225, 200–230.

Célerier, Claire, and Adrien Matray, 2019, Bank-branch supply, financial inclusion, and wealth accumulation, *The Review of Financial Studies* 32, 4767–4809.

Chava, Sudheer, Alexander Oettl, Ajay Subramanian, and Krishnamurthy V. Subramanian, 2013, Banking deregulation and innovation, *Journal of Financial Economics* 109, 759–774.

Chen, Jun, and Michael Ewens, 2025, Venture capital and startup agglomeration, *Journal of Finance* 80, 2153–2198.

Engelberg, Joseph, Jorge Guzman, Runjing Lu, and William Mullins, 2026, Partisan entrepreneurship, *Journal of Finance* .

Ewens, Michael, Ramana Nanda, and Christopher Stanton, 2024, Founder-ceo compensation and selection into venture capital-backed entrepreneurship, *Journal of Finance* 79, 3361–3405.

Fehder, Daniel C, Naomi Hausman, and Yael V Hochberg, 2025, Innovation and capital, *Journal of Financial Economics* 169, 104029.

Feldman, Maryann P., Irwin Feller, Janet Bercovitz, and Richard Burton, 2002, Equity and the technology transfer strategies of american research universities, *Management Science* 48, 105–121.

Fini, Riccardo, Markus Perkmann, and Jan-Michael Ross, 2021, Attention to exploration: The effect of academic entrepreneurship on the production of scientific knowledge, *Organization Science* 33, 688–715.

Gottlieb, Joshua D., Richard R. Townsend, and Ting Xu, 2022, Does career risk deter potential entrepreneurs?, *Review of Financial Studies* 35, 3973–4015.

Guzman, Jorge, and Scott Stern, 2020, The state of american entrepreneurship: New estimates of the quantity and quality of entrepreneurship for 32 us states, 1988–2014, *American Economic Journal: Economic Policy* 12, 212–243.

Hacamo, Isaac, and Kristoph Kleiner, 2022, Forced entrepreneurs, *Journal of Finance* 77, 49–83.

Henderson, Rebecca, Adam B. Jaffe, and Manuel Trajtenberg, 1998, Universities as a source of commercial technology: A detailed analysis of university patenting, 1965–1988, *The Review of Economics and Statistics* 80, 119–127.

Herdendorf, P. B., and G. F. Linsteadt, 1982, The transfer of Federal technology and the community college, *Journal of Technology Transfer* 6, 39–44.

Hisrich, Robert D., and Raymond W. Smilor, 1988, The university and business incubation: Technology transfer through entrepreneurial development, *Journal of Technology Transfer* 13, 14–19.

Hochberg, Yael V., Alexander Ljungqvist, and Yang Lu, 2007, Whom you know matters: Venture capital networks and investment performance, *Journal of Finance* 62, 251–301.

Hombert, Johan, and Adrien Matray, 2017, The real effects of lending relationships on innovative firms and inventor mobility, *Review of Financial Studies* 30, 2413–2445.

Hombert, Johan, Antoinette Schoar, David Sraer, and David Thesmar, 2020, Can unemployment insurance spur entrepreneurial activity? evidence from france, *Journal of Finance* 75, 1247–1285.

Hsu, David H., Edward B. Roberts, and Charles E. Eesley, 2007, Entrepreneurs from technology-based universities: Evidence from mit, *Research Policy* 36, 768–788.

Hvide, Hans K, and Benjamin F Jones, 2018, University innovation and the professor's privilege, *American Economic Review* 108, 1860–1898.

Jaffe, Adam B., 1989, Real effects of academic research, *American Economic Review* 79, 957–970.

Jang, Young Soo, and Steven N Kaplan, 2025, Venture capital start-up selection, *NBER Working Paper* .

Jensen, Richard, and Marie Thursby, 2001, Proofs and prototypes for sale: The licensing of university inventions, *American Economic Review* 91, 240–259.

Kaplan, Steven N, Berk A Sensoy, and Per Strömberg, 2009, Should investors bet on the jockey or the horse? evidence from the evolution of firms from early business plans to public companies, *Journal of Finance* 64, 75–115.

Kerr, William R., and Ramana Nanda, 2009, Democratizing entry: Banking deregulations, financing constraints, and entrepreneurship, *Journal of Financial Economics* 94, 124–149.

Kerr, William R., and Ramana Nanda, 2015, Financing innovation, *Annual Review of Financial Economics* 7, 445–462.

Kim, Scott, and Petra Moser, 2025, Women in science. lessons from the baby boom, *Econometrica* 93, 1521–1560.

Lerner, Josh, Henry Manley, Carolyn Stein, and Heidi Williams, 2024, The wandering scholars: Understanding the heterogeneity of university commercialization, *NBER Working Paper* .

Lerner, Josh, and Ramana Nanda, 2020, Venture capital's role in financing innovation: What we know and how much we still need to learn, *Journal of Economic Perspectives* 34, 237–261.

Levine, Ross, and Yona Rubinstein, 2017, Smart And Illicit: Who Becomes An Entrepreneur And Do They Earn More?, *Quarterly Journal of Economics* 132, 963–1018.

Li, Xuelai, 2026, Venture capital and scientists' selection into entrepreneurship, *SSRN Working Paper* .

Lin, Zihang, Yian Yin, Lu Liu, and Dashun Wang, 2023, SciSciNet: A Large-Scale Open Data Lake For The Science of Science Research, *Scientific Data* 10, 1–22.

Marx, Matt, and David H Hsu, 2022, Revisiting the entrepreneurial commercialization of academic science: Evidence from "twin" discoveries, *Management Science* 68, 1330–1352.

Mowery, David C, Richard R Nelson, Bhaven N Sampat, and Arvids A Ziedonis, 2001, The growth of patenting and licensing by us universities: An assessment of the effects of the bayh–dole act of 1980, *Research Policy* 30, 99–119.

Nanda, Ramana, and Jesper B. Sørensen, 2010, Workplace peers and entrepreneurship, *Management Science* 56, 1116–1126.

Nelson, Richard R, 1986, Institutions supporting technical advance in industry, *American Economic Review* 76, 186–189.

Robb, Alicia M, and David T Robinson, 2014, The capital structure decisions of new firms, *Review of Financial Studies* 27, 153–179.

Sampat, Bhaven N, 2006, Patenting and us academic research in the 20th century: The world before and after bayh-dole, *Research Policy* 35, 772–789.

Schoar, Antoinette, 2010, The divide between subsistence and transformational entrepreneurship, *Innovation Policy And The Economy* 10, 57–81.

Siegel, Donald S, David Waldman, and Albert Link, 2003, Assessing the impact of organizational practices on the relative productivity of university technology transfer offices: An exploratory study, *Research Policy* 32, 27–48.

Stuart, Toby E, and Waverly W Ding, 2006, When do scientists become entrepreneurs? the social structural antecedents of commercial activity in the academic life sciences, *American Journal of Sociology* 112, 97–144.

Sun, Liyang, and Sarah Abraham, 2021, Estimating dynamic treatment effects in event studies with heterogeneous treatment effects, *Journal of Econometrics* 225, 175–199.

Thursby, Jerry G, and Marie C Thursby, 2002, Who is selling the ivory tower? sources of growth in university licensing, *Management Science* 48, 90–104.

Toole, Andrew A, and Dirk Czarnitzki, 2010, Commercializing science: Is there a university "brain drain" from academic entrepreneurship?, *Management Science* 56, 1599–1614.

Wallskog, Melanie, 2025, Entrepreneurial Spillovers Across Coworkers, *Journal of Financial Economics* 170, 104077.

Tables

Table 1: Top 15 Scientist Employers

Employer Name	Count	Share (%)
University of California	6,361	8.25
University of Illinois	2,609	3.38
University of Michigan	1,875	2.43
Purdue University	1,734	2.25
Cornell University	1,644	2.13
Michigan State University	1,516	1.97
State University of New York	1,402	1.82
Pennsylvania State University	1,394	1.81
Ohio State University	1,309	1.70
University of Texas	1,249	1.62
Harvard University	1,238	1.61
University of Washington	1,228	1.59
University of Maryland	1,124	1.46
Columbia University	1,121	1.45
Indiana University	1,072	1.39
University of Missouri	1,058	1.37
University of Chicago	1,031	1.34
Stanford University	979	1.27
University of Pennsylvania	960	1.24
Yale University	937	1.22

Notes: This table reports the 20 most common employer names in the scientist sample. In most state university systems, TTOs are established at the system-wide level, meaning the treatment applies to all campuses. However, five institutions – University of Wisconsin Madison, University of Illinois Urbana-Champaign, UNC Chapel Hill, University of Arkansas Fayetteville, and Washington University in St. Louis – implemented TTOs at the individual campus level. Because scientists' affiliations often do not specify a particular campus, we assume that if a single campus within a system establishes a TTO, the entire system is treated as having received the shock.

Table 2: Business Formation Rates by Scientist Characteristics

		Full Sample		Never Treated Sample	
Category	Group	Count	Formation Rate (%)	Count	Formation Rate (%)
Age	<45	20,967	2.57	9,454	2.13
	45-60	63,225	2.30	36,817	2.23
	>60	24,368	1.93	12,638	1.68
Degree	PhD/MD	67,378	2.42	35,470	2.15
	Non-PhD	42,206	2.03	23,943	2.02
Income	High	48,549	2.70	25,896	2.47
	Low	51,639	1.90	28,380	1.79
Field	Engineering	3,652	3.72	1,610	3.54
	Environmental sciences	5,027	2.71	2,574	2.33
	Life sciences	18,110	2.06	9,249	1.84
	Mathematical & computer sciences	12,063	2.36	7,035	1.98
	Physical sciences	30,112	2.18	15,369	2.08
	Psychology	8,623	2.77	5,468	2.49
	Social sciences	27,230	2.07	15,442	2.01

Notes: The full sample contains 109,584 scientists in the regression sample. The never-treated sample contains the subset of 59,413 scientists with no observed treatment during the panel. Business formation rate is the percentage of scientists who ever start a business in the sample period. Age is defined as scientists' age at the time of TTO establishment. Government indicates whether scientists receive government funding. Degree groups classify PhD, MD, and PhD+ as PhD/MD, with all remaining degree categories grouped as Non-PhD. Field refers to the major of the scientist's highest degree.

Table 3: Effects of TTO Establishment on Business Formation

	(1)	(2)	(3)
ATT	0.0012***	0.0013***	0.0011***
	(0.0003)	(0.0003)	(0.0003)
N Scientists (treated)	50,171	50,171	50,171
N Scientists (control)	59,413	108,104	59,413

Notes: This table reports difference-in-differences estimates of the effect of TTO establishment on business formation from 1970 to 1990 using the full sample. Columns (1)–(2) report the simple aggregated ATT from Callaway and Sant'Anna (2021) using never-treated and not-yet-treated controls, respectively. Column (3) reports the aggregated ATT from the Sun and Abraham (2021) interaction-weighted estimator. Standard errors in parentheses. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table 4: Heterogeneous Effects: Scientist Characteristics

	Never-Treated Control			Not-Yet-Treated Control		
Panel A: Age at TTO Establishment						
	<45	45-60	>60	<45	45-60	>60
	(1)	(2)	(3)	(4)	(5)	(6)
ATT	0.0023***	0.0007*	0.0005	0.0024***	0.0008**	0.0006
	(0.0006)	(0.0004)	(0.0006)	(0.0006)	(0.0004)	(0.0006)
N Scientists (treated)	11,513	26,408	11,730	11,513	26,408	11,730
N Scientists (control)	9,454	36,817	12,638	19,872	62,956	24,271
Panel B: Degree						
	PhD/MD	Non-PhD		PhD/MD	Non-PhD	
	(1)	(2)		(3)	(4)	
ATT	0.0014***	0.0006		0.0016***	0.0007	
	(0.0003)	(0.0004)		(0.0003)	(0.0005)	
N Scientists (treated)	31,908	18,263		31,908	18,263	
N Scientists (control)	35,470	23,943		66,414	41,690	
Panel C: Within-County Income						
	High Income	Low Income		High Income	Low Income	
	(1)	(2)		(3)	(4)	
ATT	0.0009**	0.0013***		0.0011**	0.0013***	
	(0.0004)	(0.0004)		(0.0004)	(0.0004)	
N Scientists (treated)	22,653	23,259		22,653	23,259	
N Scientists (control)	25,896	28,380		47,885	50,935	

Notes: This table reports heterogeneous difference-in-differences estimates of the effect of TTO establishment on business formation from 1970 to 1990 using the full sample. Both treated and control scientists are restricted to the relevant subgroup in each panel. Columns (1)–(3) use never-treated controls; columns (4)–(6) use not-yet-treated controls. For treated scientists, age at treatment equals TTO establishment year minus birth year. Never-treated scientists are assigned to age groups based on the modal age group of treated scientists with the same birth year, ensuring comparable birth cohorts within each subgroup. Appendix Table **??** reports results using birth cohort groups directly, which require no imputation. Income is defined as the scientist's gross income; the High/Low split is taken at the within-county median, so each scientist is classified relative to peers at the same county. All estimates report simple ATT from Callaway and Sant'Anna (2021). Standard errors in parentheses. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table 5: Heterogeneous Effects: University Characteristics

	Never-Treated Control		Not-Yet-Treated Control	
Panel A: Federal Funding Share				
	High Gov Share (1)	Low Gov Share (2)	High Gov Share (3)	Low Gov Share (4)
ATT	0.0010*	−0.0004	0.0015**	−0.0003
	(0.0006)	(0.0004)	(0.0006)	(0.0004)
N Scientists (treated)	19,102	25,845	19,102	25,845
N Scientists (control)	7,793	8,647	25,916	33,991
Panel B: Industry Funding Share				
	High Ind. Share (1)	Low Ind. Share (2)	High Ind. Share (3)	Low Ind. Share (4)
ATT	0.0002	0.0006	0.0003	0.0012**
	(0.0003)	(0.0005)	(0.0003)	(0.0005)
N Scientists (treated)	20,763	24,184	20,763	24,184
N Scientists (control)	8,545	7,895	28,807	31,100
Panel C: Own-Field R&D Spending				
	High Own-Field R&D (1)	Low Own-Field R&D (2)	High Own-Field R&D (3)	Low Own-Field R&D (4)
ATT	0.0005	0.0014***	0.0017***	0.0014***
	(0.0006)	(0.0005)	(0.0005)	(0.0005)
N Scientists (treated)	26,584	14,944	26,584	14,944
N Scientists (control)	4,983	15,805	30,399	30,490

Notes: This table reports heterogeneous difference-in-differences estimates of the effect of TTO establishment on business formation from 1970 to 1990 using the full sample. Both treated and control scientists are restricted to the relevant subgroup in each panel. Columns (1)–(2) use never-treated controls; columns (3)–(4) use not-yet-treated controls. Federal funding share is federal R&D expenditure divided by total R&D (federal + industry + institutional + state & local government), split at the median across universities. Industry funding share is industry R&D expenditure divided by total R&D (federal + industry + institutional + state & local government), split at the median across universities. Own-field R&D spending is the university's R&D expenditure in the scientist's own field of science, split at the within-field median across unique university–field combinations. All estimates report simple ATT from Callaway and Sant'Anna (2021). Standard errors in parentheses. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table 6: Heterogeneous Effects: External Finance

	Never-Treated Control		Not-Yet-Treated Control	
Panel A: Bank Branch Availability in 1970: Top 25% vs. Remaining Counties				
	More Banks (1)	Less Banks (2)	More Banks (3)	Less Banks (4)
ATT	0.0013***	0.0003	0.0015***	0.0003
	(0.0004)	(0.0003)	(0.0004)	(0.0003)
N Scientists (treated)	39,913	10,172	39,913	10,172
N Scientists (control)	38,912	20,045	77,390	30,178
Panel B: Venture Capital Presence				
	Has VC (1)	No VC (2)	Has VC (3)	No VC (4)
ATT	0.0015***	−0.0003	0.0020***	−0.0003
	(0.0004)	(0.0003)	(0.0004)	(0.0003)
N Scientists (treated)	23,517	26,654	23,517	26,654
N Scientists (control)	20,128	39,285	42,685	65,419

Notes: This table reports heterogeneous difference-in-differences estimates of the effect of TTO establishment on business formation from 1970 to 1990 using the full sample. In each panel, both treated and control scientists are restricted to the relevant subgroup. Bank branch availability is measured using the number of county bank branches in 1970. Top-bank counties are defined as counties above the 75th percentile of the branch-count distribution. VC presence is an indicator equal to one if the county has any recorded VC deals before 1970. Standard errors are reported in parentheses. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

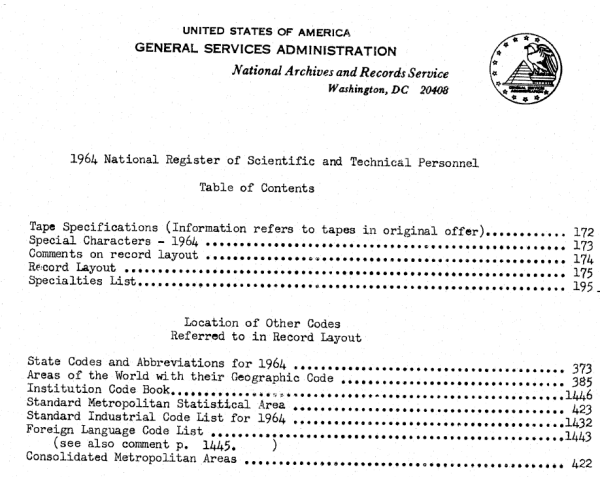
Table 7: Heterogeneous Effects: External Finance and Income

Panel A: Within-County Income × Bank Branch Availability				
	High Income		Low Income	
	More Banks (1)	Less Banks (2)	More Banks (3)	Less Banks (4)
ATT	0.0009*	0.0001	0.0013**	0.0006**
	(0.0005)	(0.0006)	(0.0005)	(0.0003)
Number of treated scientists	18,050	4,581	18,419	4,784
Number of control scientists	17,018	8,682	18,304	9,853
Panel B: Within-County Income × Venture Capital Presence				
	High Income		Low Income	
	Has VC (1)	No VC (2)	Has VC (3)	No VC (4)
ATT	0.0013	−0.0006	0.0014**	0.0001
	(0.0008)	(0.0005)	(0.0006)	(0.0006)
N Scientists (treated)	10,497	12,156	10,964	12,295
N Scientists (control)	8,850	17,046	9,429	18,951

Notes: This table reports heterogeneous difference-in-differences estimates of the effect of TTO establishment on business formation from 1970 to 1990 using the full sample and never-treated scientists as the control group. Results with not-yet-treated control groups are in Table A5. In each panel, both treated and control scientists are restricted to the relevant subgroup. Bank branch availability is measured using the number of county bank branches in 1970. Top-bank counties are defined as counties above the 75th percentile of the branch-count distribution. VC presence is an indicator equal to one if the county has any recorded VC deals before 1970. Income is split at the within-county median of first-observation gross income. All estimates report simple aggregated ATT from Callaway and Sant'Anna (2021). Standard errors are reported in parentheses.
$^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

7 Figures

Figure 1: NRSTP Documentation and Data Harmonization



UNITED STATES OF AMERICA
GENERAL SERVICES ADMINISTRATION
National Archives and Records Service
Washington, DC 20408

1964 National Register of Scientific and Technical Personnel

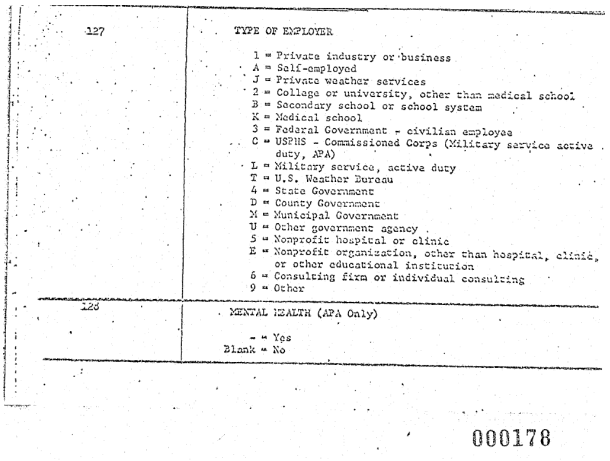
Table of Contents

Tape Specifications (Information refers to tapes in original offer)..... 172
Special Characters - 1964 173
Comments on record layout 174
Record Layout 175
Specialties List..... 195

Location of Other Codes
Referred to in Record Layout

State Codes and Abbreviations for 1964 373
Areas of the World with their Geographic Code 385
Institution Code Book.....1446
Standard Metropolitan Statistical Area 423
Standard Industrial Code List for 1964 1432
Foreign Language Code List 1443
(see also comment p. 1445.)
Consolidated Metropolitan Areas 422

(a) Original Code Book of NRSTP 1964



127 — TYPE OF EMPLOYER

1 = Private industry or business
A = Self-employed
J = Private weather services
2 = College or university, other than medical school
B = Secondary school or school system
K = Medical school
3 = Federal Government - civilian employee
C = USPHS - Commissioned Corps (Military service active duty, APA)
L = Military service, active duty
T = U.S. Weather Bureau
4 = State Government
D = County Government
M = Municipal Government
U = Other government agency
5 = Nonprofit hospital or clinic
E = Nonprofit organization, other than hospital, clinic, or other educational institution
6 = Consulting firm or individual consulting
9 = Other

128 — MENTAL HEALTH (APA Only)

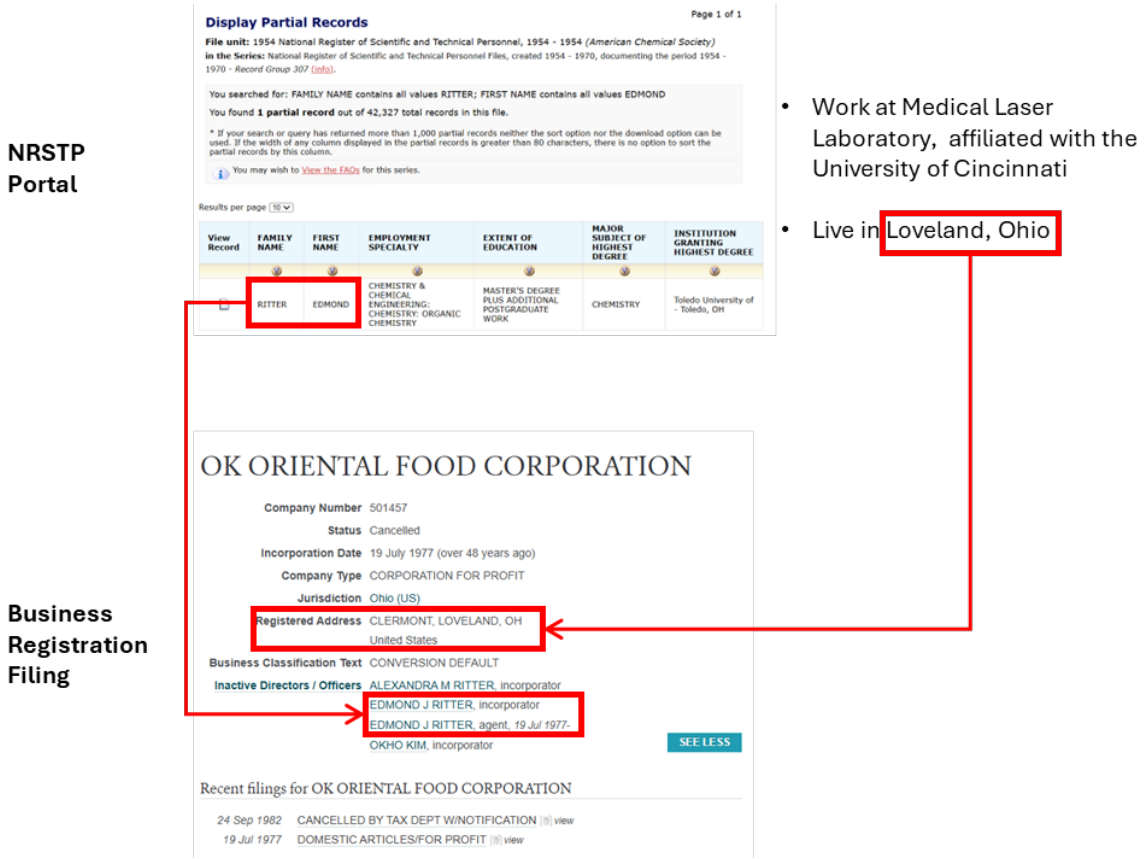
- = Yes
Blank = No

000178

(b) Classification of Employer Types

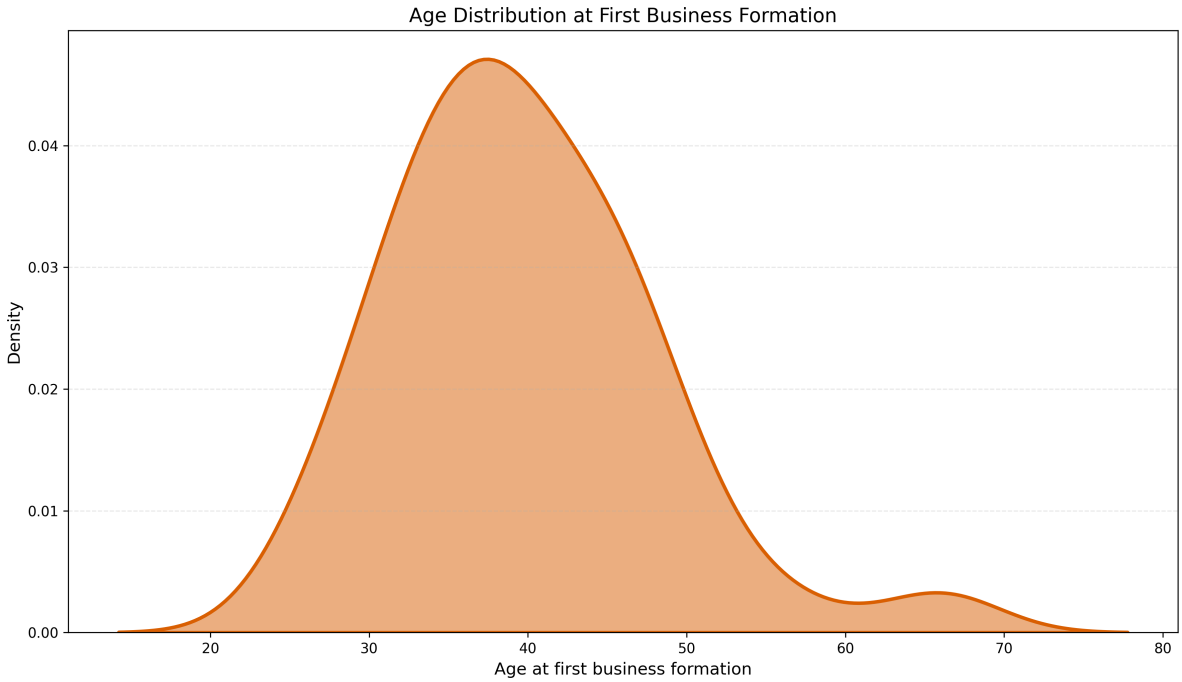
Notes: This figure illustrates the digitized codebook of the NRSTP and the classification of employer types used for data harmonization. Because the raw NRSTP data consist of coded entries, the codebook was digitized to facilitate the decoding and analysis of individual scientist information. In the employer classification for 1964, both colleges/universities and medical schools are harmonized as "university employees" to maintain consistency across different data waves.

Figure 2: Identifying Business Formation Activities



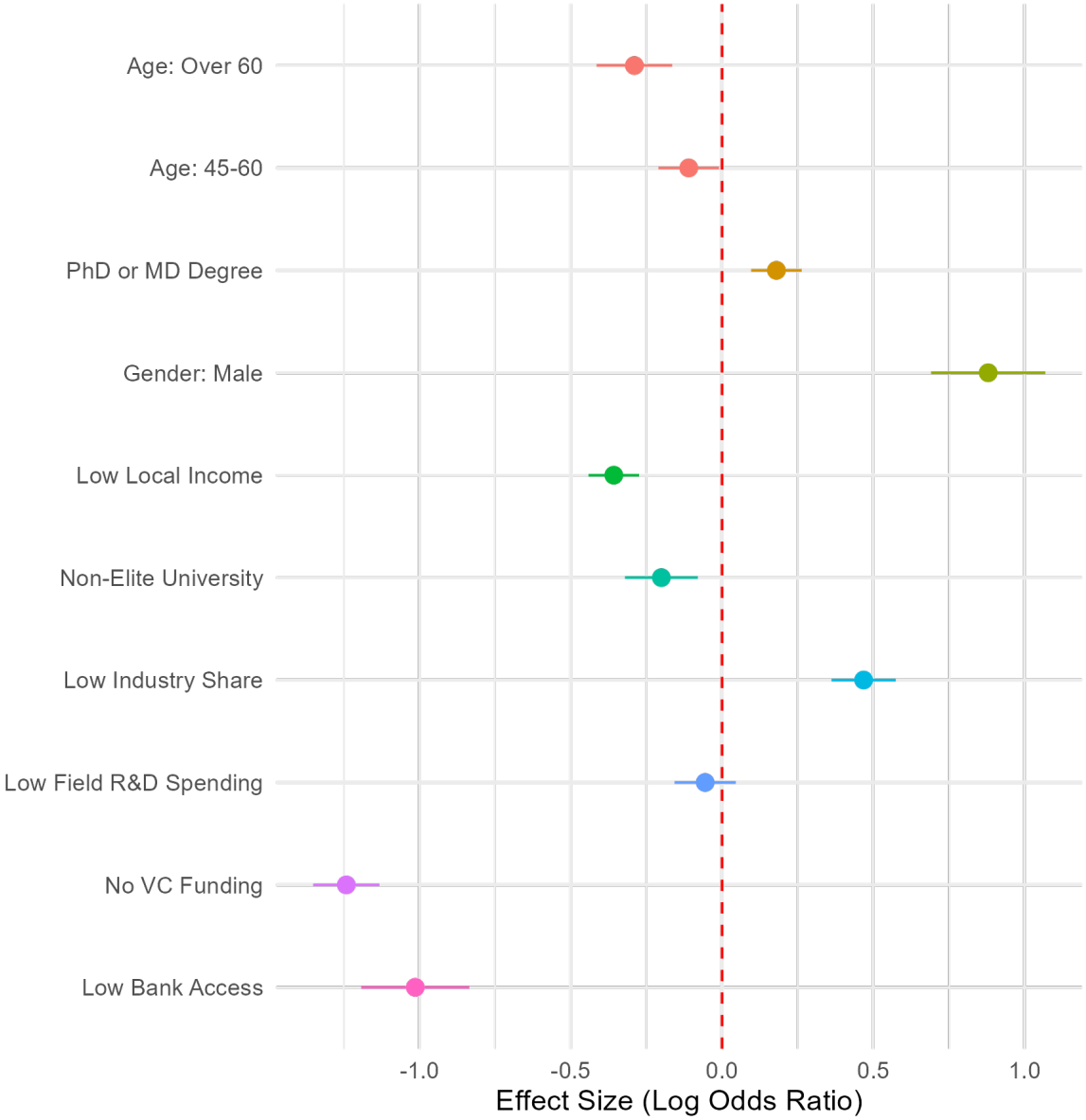
Notes: This figure illustrates the digitized codebook of the NRSTP and the classification of employer types used for data harmonization. As the raw NRSTP data consist of coded entries, the codebook was digitized to facilitate the decoding and analysis of information for each scientist. Both colleges/universities and medical schools are classified as university employees to ensure consistency across different data waves. The sample is restricted to university employees to focus specifically on university scientists and TTO.

Figure 3: Age Distribution of Entrepreneurs



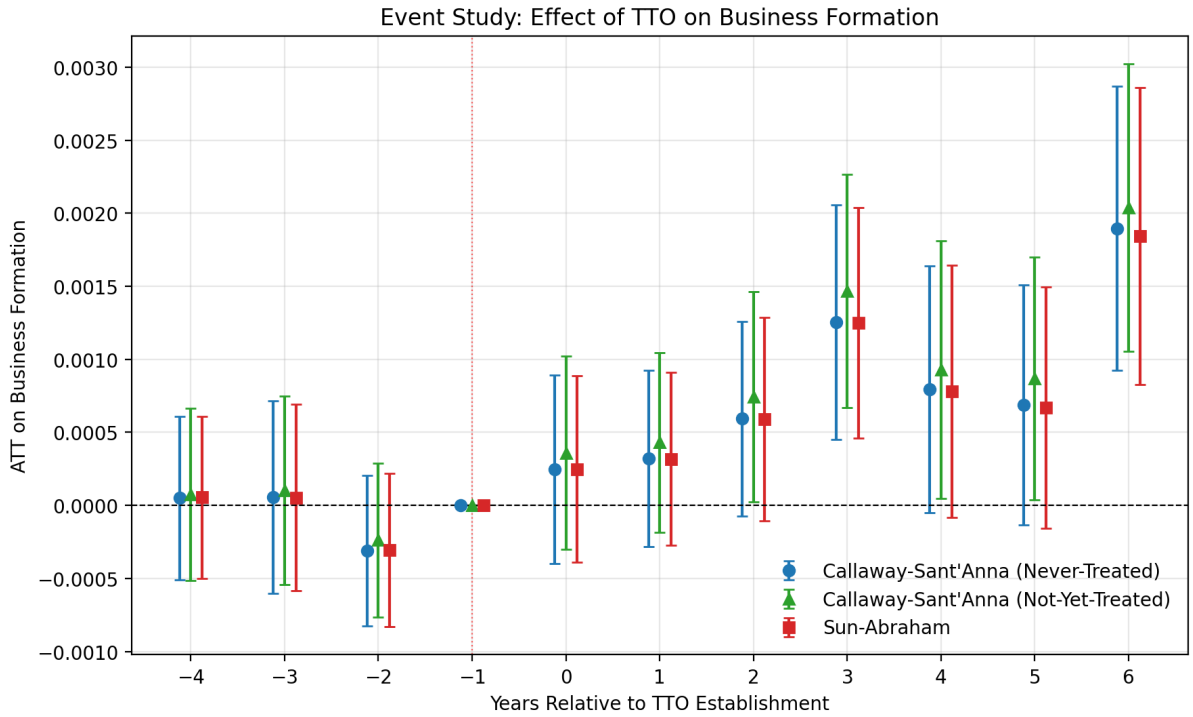
Notes: This figure plots the age at which scientists in the sample first engaged in business formation. Age is calculated as the difference between the first business start year and the individual's birth year. The density curve uses a Gaussian Kernel Density Estimate to smooth the distribution.

Figure 4: Correlates of Scientists' Business Formation



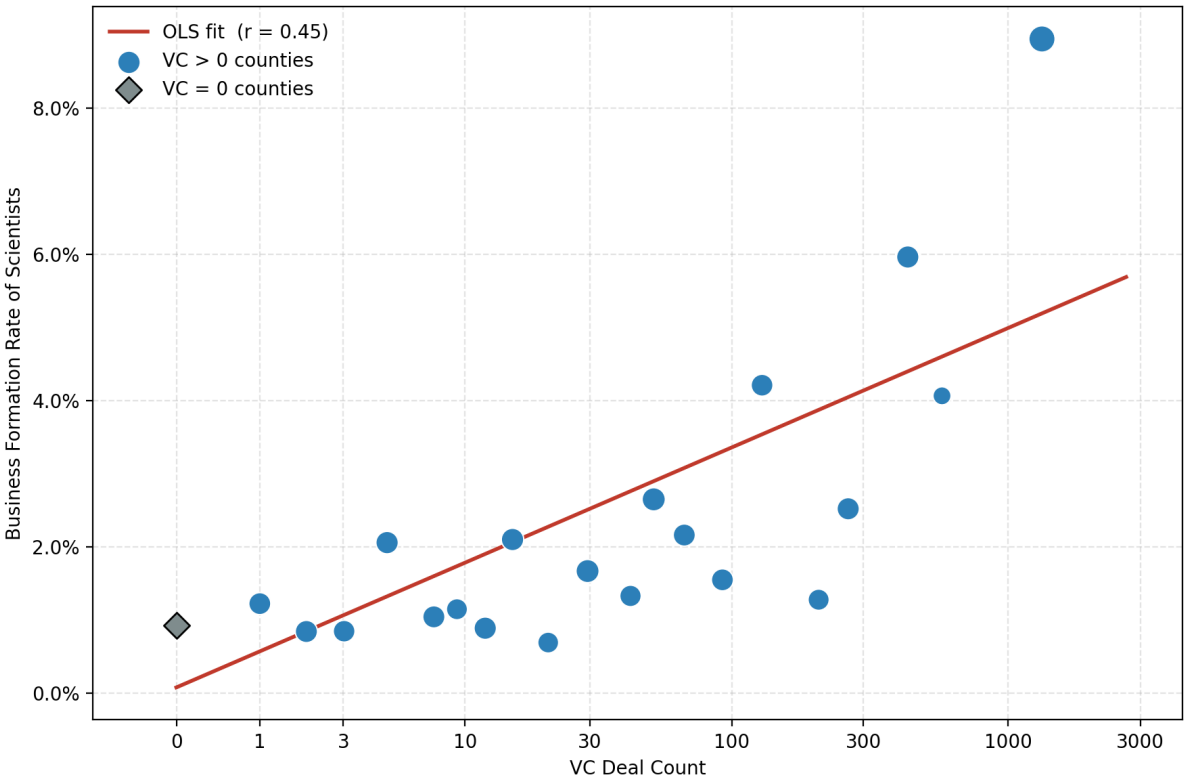
Notes: The figure reports coefficient estimates from separate logistic regressions in which the dependent variable is an indicator equal to one if the scientist starts a business. Each coefficient is estimated in a univariate specification using the subsample with non-missing values for the corresponding characteristic. Dots represent point estimates and whiskers denote 95% confidence intervals. Coefficients are reported in log-odds units. The omitted category in each case is the reference group not displayed on the figure; hence each estimate should be interpreted relative to that category.

Figure 5: Event Study: Effect of TTO Establishment on Business Formation



Notes: This figure plots dynamic treatment effect estimates from the Callaway and Sant'Anna (2021) and Sun and Abraham (2021) event study specifications on the full sample. The dependent variable is the number of businesses started by a scientist in a given year. Vertical bars represent 95% confidence intervals. The dotted vertical line marks the last pre-treatment period.

Figure 6: Local VC Activity and Scientist Business Formation



Notes: The figure is a binscatter of county-level venture capital activity against the business formation rate of scientists. The sample consists of U.S. counties observed over 1970–1990 with at least one scientist in the cross-section. The x-axis measures cumulative VC deal count in the county over the sample period; the axis scale is log(1 + VC deals), so the zero-VC point plots at a finite location. Counties with zero VC deals are displayed as a separate point. The remaining counties are grouped into 20 equal-weight bins sorted by deal count. The business formation rate is the share of scientists in a county who started a business. The fitted line is weighted OLS of the formation rate on log(1 + VC deals), with counties weighted by scientist coun. The weighted correlation between log(1 + VC deals) and the formation rate is 0.45.

Internet Appendix for

"Commercialization Infrastructure, External Finance, and Academic Entrepreneurship"

Xuelai Li and Xizhao Wang

July 2026

A Appendix Tables

Table A1: University Funding and Expenses (1972)

Variable	N	Mean	SD	P25	Median	P75	Min	Max
Panel A: Annual University Funding by Source (k USD)								
Federal	7,718	30,281.5	44,356.8	5,539	15,989	33,447	2	173,730
Industry	7,219	791.0	765.2	175	564	1,172	1	3,472
Institution funds	7,014	6,447.2	9,548.9	495	2,272	9,106	1	34,053
State/Local Gov.	7,269	2,847.9	3,719.6	294	1,239	4,384	1	15,635
Panel B: Annual University Expense by Field (k USD)								
Engineering	6,168	1,515.9	2,508.4	238	615	1,364	1	14,552
Environmental	5,583	581.5	1,145.1	58	164	402	1	5,119
Life sciences	7,229	14,008.9	16,504.1	1,200	5,196	26,458	1	56,021
Math / CS	5,244	556.8	919.6	18	158	486	1	3,386
Physical sciences	7,055	2,405.1	4,253.8	270	891	3,354	1	26,551
Psychology	4,953	471.5	987.1	19	70	262	1	4,844
Social sciences	6,325	1,617.5	3,157.6	65	427	2,005	1	15,445

Notes: Panel A uses the university-level HERD employer file. The variables *Federal*, *Industry*, *Institution funds*, and *State and local government* measure university R&D funding by source. Panel B uses the university-level HERD employer file. The reported variables measure university R&D expenditures by field: *Engineering*, *Environmental sciences*, *Life sciences*, *Mathematical and computer sciences*, *Physical sciences*, *Psychology*, and *Social sciences*.

Table A2: Variable Definitions and Construction

Variable Name	**Definition & Construction**
Business Registrations	Count of firms in the OpenCorporates dataset (1945–1990) matched to scientists via individual names and county FIPS codes. Data were regularized using the OpenStreetMap API for geocoding and filtered using spaCy NLP to identify "PERSON" entities, excluding corporate officer names.
Publications	Count of English-language journal publications matched via institutional affiliation FIPS codes. To account for the cohort's age and minimize false positives, records appearing after 2005 are excluded.
TTO Establishment	A binary or timing indicator for the establishment of a Technology Transfer Office at the university where the scientist was affiliated during the 1964–1968 NRSTP survey window.
Age	Estimated birth year derived by subtracting age constants from the Year of Highest Degree: 30 years for doctoral-level degrees, 25 for Master's, and 22 for Bachelor's.
Gross Income	Annual gross income self-reported in the NRSTP. This measure includes both base salary and other income.
Major	Standardized classification of the scientist's primary field of study. Granular subfields were aggregated into broad categories (e.g., Physical Sciences, Engineering, Math/CS) to align with HERD university spending taxonomy.
Industry Funding	The proportion of a university's total R&D income derived from industry-sponsored research, as recorded in the 1972 HERD survey.
Bank Access	Indicator equal to one if a scientist is located in a county with above-median FDIC bank branch counts in 1970, and zero otherwise. This variable proxies for greater local credit supply.
Has VC	Indicator equal to one if a scientist is located in a county that recorded at least one VC deal before 1979, and zero otherwise. This variable captures local exposure to VC activity prior to the ERISA reform.

Table A3: Survival-style DID: TTO Effect on First Business Formation

	Cumulative-Ever Indicator			Cox Hazard Ratio
	(1)	(2)	(3)	(4)
ATT / HR	0.0034***	0.0036***	0.0033***	1.5360***
	(0.0005)	(0.0005)	(0.0005)	(0.0501)
N Scientists (treated)	50,171	50,171	50,171	50,171
N Scientists (control)	59,413	109,584	59,413	59,413

Notes: Survival-style DID estimates of TTO establishment on first business formation, 1970–1990. Columns (1)–(3) report the simple aggregated ATT from staggered DID estimators with the outcome defined as the cumulative-ever indicator ($\mathbf{1}\{\text{any biz} \leq t\}$); Columns (1) and (2) use never-treated and not-yet-treated comparison groups respectively (Callaway and Sant'Anna 2021), and Sun-Abraham (Sun and Abraham 2021) uses the never-treated as the comparison. Column (4) reports the hazard ratio from a pooled Cox proportional hazards model with time-varying treatment indicator $\text{post_tto}_{it} = \mathbf{1}\{t \geq \text{gvar}_i\}$. Standard errors clustered at the scientist level in parentheses (SE on log-HR scale for column 4). $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table A4: Effects of TTO Establishment on Business Formation: Public Universities Only

	Never-Treated Control (1)	Not-Yet-Treated Control (2)	Sun-Abraham (3)
ATT	0.0016***	0.0018***	0.0016***
	(0.0004)	(0.0004)	(0.0003)
N Scientists (treated)	34,245	34,245	34,245
N Scientists (control)	15,940	49,684	15,940

Notes: This table reports difference-in-differences estimates of the effect of TTO establishment on business formation from 1970 to 1990 using the full sample restricted to scientists at *public* universities (HERD type-of-control code 1, supplemented by manual classification). Columns (1)–(2) report the simple aggregated ATT from Callaway and Sant'Anna (2021) using never-treated and not-yet-treated controls, respectively. Column (3) reports the aggregated ATT from the Sun and Abraham (2021) interaction-weighted estimator. Standard errors are clustered at the scientist level and reported in parentheses. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table A5: Exterinal Finance and Income with Not-Yet-Treated Control Group

Panel A: Within-County Income × Bank Branch Availability				
	High Income		Low Income	
	More Banks (1)	Less Banks (2)	More Banks (3)	Less Banks (4)
ATT	0.0013**	0.0001	0.0015***	0.0006**
	(0.0006)	(0.0006)	(0.0005)	(0.0003)
Number of treated scientists	18,050	4,581	18,419	4,784
Number of control scientists	34,418	13,251	36,047	14,613
Panel B: Within-County Income × Venture Capital Presence				
	High Income		Low Income	
	Has VC (1)	No VC (2)	Has VC (3)	No VC (4)
ATT	0.0018**	−0.0005	0.0017***	0.0001
	(0.0008)	(0.0005)	(0.0006)	(0.0006)
N Scientists (treated)	10,497	12,156	10,964	12,295
N Scientists (control)	18,923	28,962	19,930	31,005

Notes: This table reports heterogeneous difference-in-differences estimates of the effect of TTO establishment on business formation from 1970 to 1990 using the full sample and not-yet-treated scientists as the control group. In each panel, both treated and control scientists are restricted to the relevant subgroup. Bank branch availability is measured using the number of county bank branches in 1970. Top-bank counties are defined as counties above the 75th percentile of the branch-count distribution. VC presence is an indicator equal to one if the county has any recorded VC deals before 1970. Income is split at the within-county median of first-observation gross income. All estimates report simple aggregated ATT from Callaway and Sant'Anna (2021). Standard errors are reported in parentheses. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table A6: Effect of TTO Establishment on Publication Output

	(1)	(2)	(3)
ATT	-0.0074^{***}	-0.0066^{***}	-0.0072^{***}
	(0.0017)	(0.0017)	(0.0017)
N Scientists (treated)	50,171	50,171	50,171
N Scientists (control)	59,413	108,104	59,413

Notes: This table reports difference-in-differences estimates of the effect of TTO establishment on publication output from 1970 to 1990 using the full sample. Columns (1)–(2) report the simple aggregated ATT from Callaway and Sant'Anna (2021) using never-treated and not-yet-treated controls, respectively. Column (3) reports the aggregated ATT from the Sun and Abraham (2021) interaction-weighted estimator. Standard errors in parentheses. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

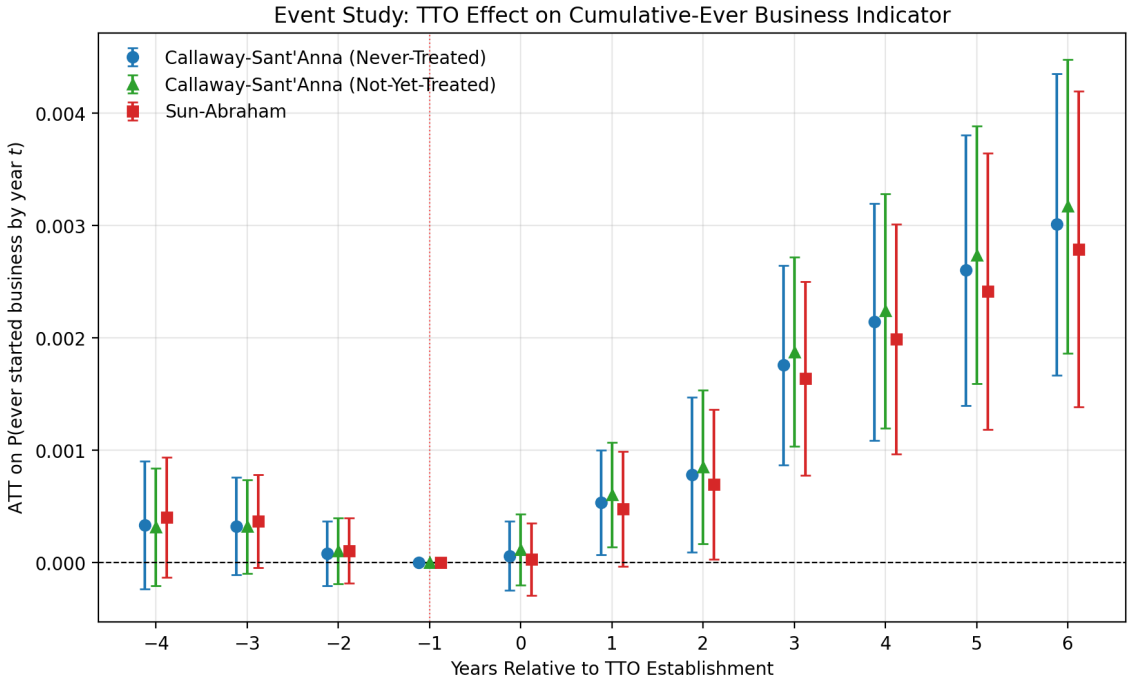
Table A7: Effect of Business Formation on Publication Output: TTO-Treated Scientists

	(1)	(2)	(3)
ATT	0.0089	0.0088	0.0089
	(0.0144)	(0.0143)	(0.0145)
N Scientists (treated)	1,241	1,241	1,241
N Scientists (control)	48,930	50,143	48,930

Notes: This table reports staggered difference-in-differences estimates of the effect of starting a business on publication output. The treatment variable is the year a scientist first records a new business ($bizcount > 0$); the control group consists of scientists who never start a business in Column (1), who have not yet started a business in Column(2), or never-biz scientists used as the comparison group in the Sun-Abraham interaction-weighted estimator in Column (3). The sample is restricted to TTO-treated scientists, i.e., scientists whose employer established a Technology Transfer Office during the panel period (1970–1990). Columns (1)–(2) report the simple aggregated ATT from Callaway and Sant'Anna (2021) using never-treated and not-yet-treated controls, respectively. Column (3) reports the aggregated ATT from the Sun and Abraham (2021) interaction-weighted estimator. Standard errors in parentheses. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

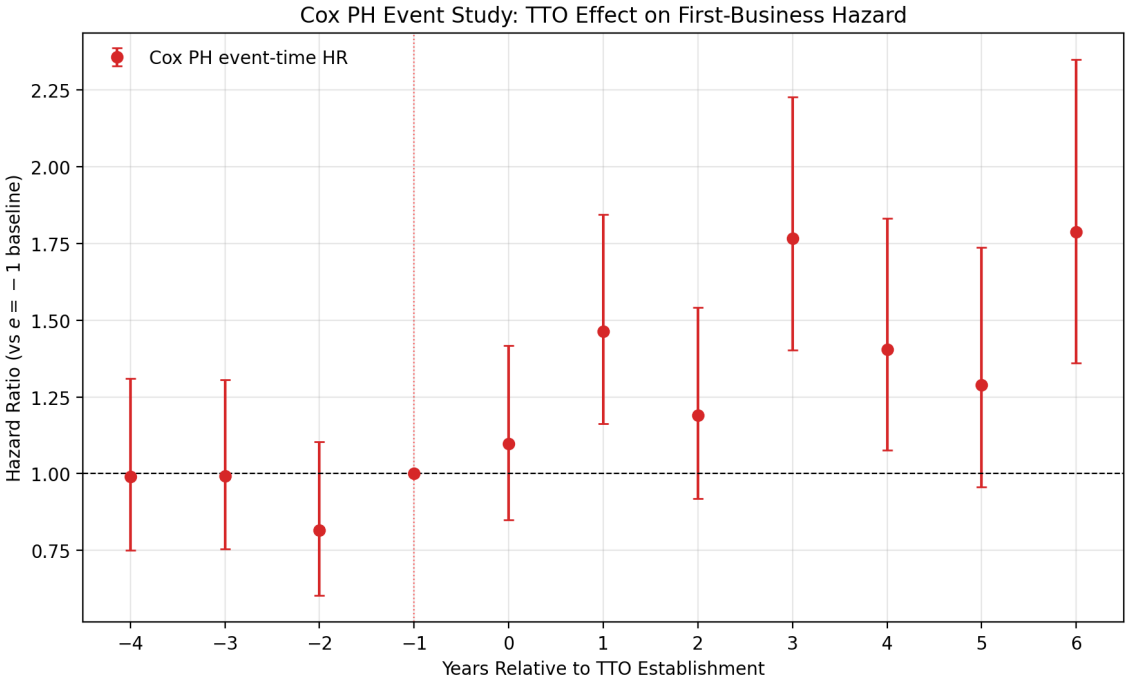
B Appendix Figures

Figure A1: Cumulative Business Formation Staggered DID



Notes: This figure plots the event-study estimates of TTO effect on the cumulative business formation indicator. CSDID with never-treated and not-yet-treated comparison groups (Callaway and Sant'Anna, 2021) and the Sun and Abraham (2021) interaction-weighted estimator. Vertical bars represent 95% confidence intervals. The dotted vertical line marks the last pre-treatment period.

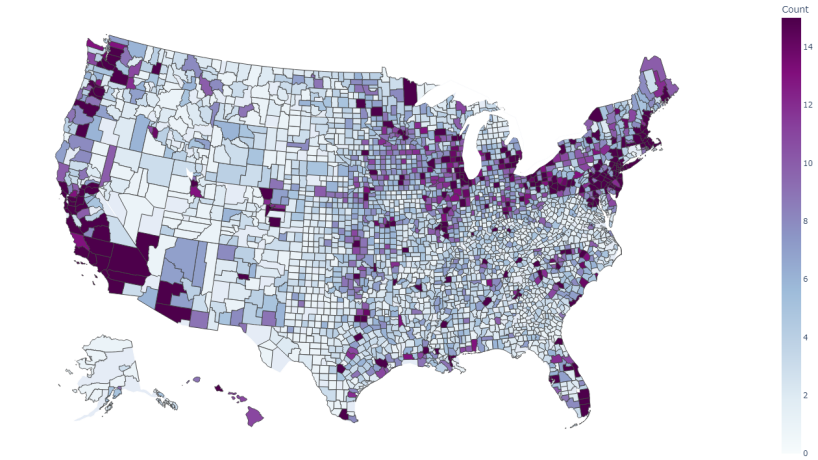
Figure A2: Cox Proportional Hazards Staggered DID



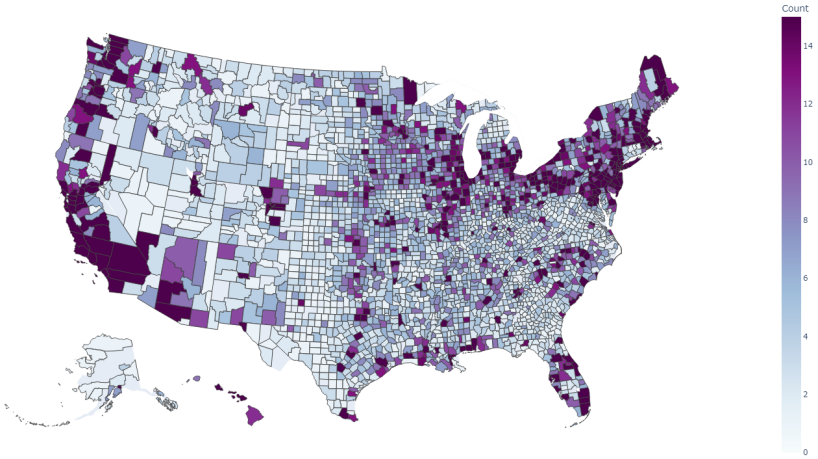
Notes: This figure plots the event-time hazard ratios from a Cox proportional-hazards staggered DID on first business formation. Each point reports $\exp(\widehat{\beta}_k)$ for event-time dummy $\mathbf{1}\{e_{it} = k\}$ in $k \in [-4, +6]$. Vertical bars represent 95% confidence intervals. The dotted vertical line marks the last pre-treatment period.

Figure A3: Geographic Distribution of FDIC Bank Branches, 1970 and 1979

(a) 1970

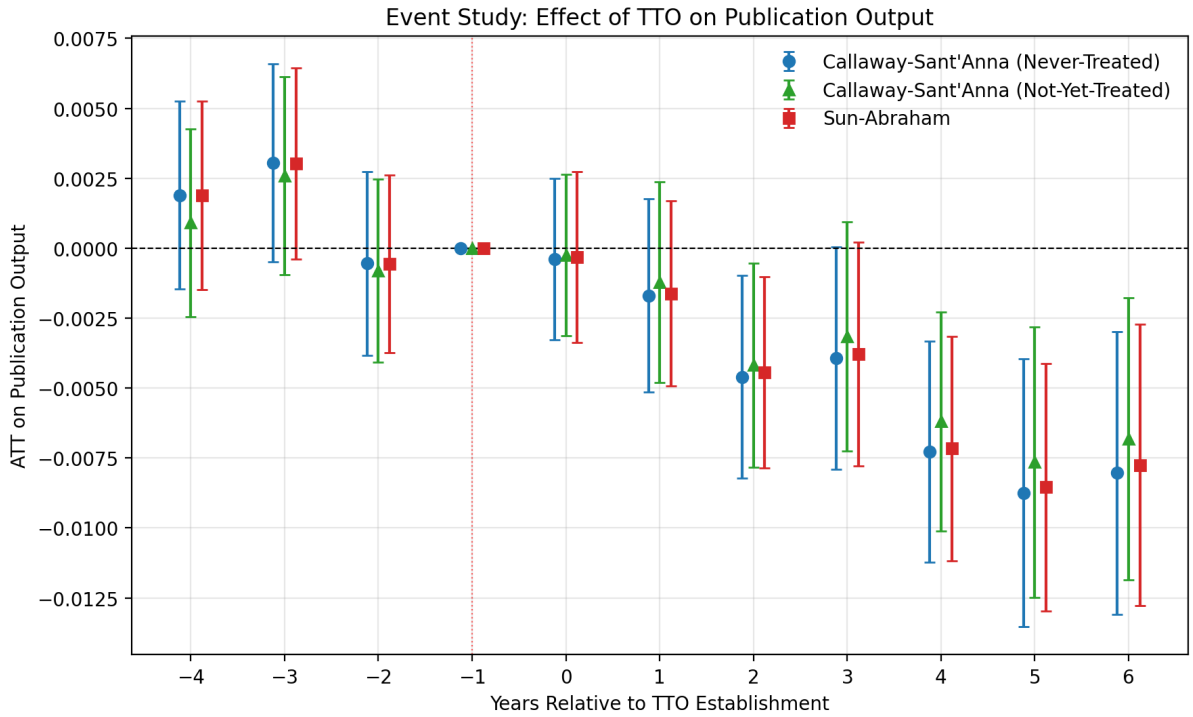


(b) 1979



Notes: This figure shows the geographic distribution of active FDIC-insured bank branches across U.S. counties in 1970 and 1979. The spatial distribution of branches is largely stable over time, supporting the use of the 1970 county-level branch count as a proxy for local credit supply.

Figure A4: Event Study: Effect of TTO Establishment on Publication Output



Notes: This figure plots dynamic treatment effect estimates from the Callaway and Sant'Anna (2021) and Sun and Abraham (2021) event study specifications on the full sample. The dependent variable is the number of publications by a scientist in a given year. Vertical bars represent 95% confidence intervals. The dotted vertical line marks the last pre-treatment period.