NBER WORKING PAPER SERIES

MORE THAN AN IVORY TOWER:
THE IMPACT OF RESEARCH INSTITUTIONS ON THE QUANTITY AND QUALITY OF ENTREPRENEURSHIP

Valentina Tartari
Scott Stern

Working Paper 28846
http://www.nber.org/papers/w28846

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
May 2021

We would like to thank Pierre Azoulay, Maryann Feldman, Jeff Furman, Jorge Guzman, Fiona Murray, and seminar participants at Copenhagen Business School, DRUID, the AIEA-NBER Conference, the 2021 ASSA meetings, the NBER Productivity Seminar, and participants in the MIT Regional Ecosystem Acceleration Program for useful comments and feedback, and Luca Gius for excellent research assistance. This work was supported by the Jean Hammond (1986) and Michael Krasner (1974) Entrepreneurship Fund and the Edward B. Roberts (1957) Entrepreneurship Fund at MIT, and the Kauffman Foundation. Valentina Tartari gratefully acknowledges the support from the Independent Research Fund Denmark through a Sapere Aude Research Talent Grant (DFF –4003-00137B), and Stern gratefully acknowledges the hospitality and hosting of Copenhagen Business School. All errors and omissions are of course our own. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

At least one co-author has disclosed additional relationships of potential relevance for this research. Further information is available online at http://www.nber.org/papers/w28846.ack

More than an Ivory Tower: The Impact of Research Institutions on the Quantity and Quality of Entrepreneurship
Valentina Tartari and Scott Stern
NBER Working Paper No. 28846
May 2021
JEL No. L26,O3,R12

ABSTRACT

This paper provides systematic empirical evidence for the distinctive role of universities on local entrepreneurial ecosystems. Assessing the impact of research institutions on entrepreneurship is challenging, given that these institutions are often located in economic and innovation environments conducive to growth-oriented entrepreneurial activity, are themselves a source of local demand, and produce knowledge, which might serve as the foundation for new ventures. To overcome this inference challenge, we first combine comprehensive business registration records with a predictive analytics approach to measure both the quantity and quality-adjusted quantity of entrepreneurship at the zip-code level on an annual basis. We then link each location to the presence or absence of research-oriented universities or national laboratories. Finally, we exploit significant changes over time in Federal commitments to both universities and national laboratories. Our key finding is that changes in Federal research commitments to universities are uniquely linked to positively correlated changes in the quality-adjusted quantity of entrepreneurship. In contrast, increases in non-research funding to universities and funding to national laboratories is associated with either a neutral or negative impact on the quality-adjusted quantity of entrepreneurship. Research funding to universities seems to play a unique role in promoting the acceleration of local entrepreneurial ecosystems.

Valentina Tartari
Copenhagen Business School
Kilevej 14A
2000 Frederiksberg
Denmark
vt.si@cbs.dk

Scott Stern
MIT Sloan School of Management
100 Main Street, E62-476
Cambridge, MA 02142
and NBER
sstern@mit.edu

I. Introduction

A distinguishing feature of publicly funded research organizations such as universities or national laboratories is that they are situated in a single (or small number) of locations. As geographically bounded institutions, universities and national laboratories are often invoked, by both economists and policymakers, as drivers of local knowledge spillovers (Jaffe, 1989; Acs, Audretsch and Feldman, 1992; Jaffe, Trajtenberg and Henderson, 1993; Moretti, 2012) and local economic development (Adams, Chiang and Jensen, 2003; Kantor and Whalley, 2014; Hausman, 2020).

A particularly important route through which research institutions might influence local outcomes is through entrepreneurship. The role of universities and related institutions as drivers of local entrepreneurship (and subsequent economic development) is largely grounded in academic and policy discussions of leading entrepreneurial "ecosystems" such as Silicon Valley (anchored by institutions such as Stanford, Berkeley, and UC-SF), Boston (with at least 10 research institutions), and the Research Triangle Park in North Carolina (Feldman, 1994, 2014; Saxenian, 1996). For example, in the late 1980s, a group of five North Carolina State engineers conducting pioneering research on the advantages silicon carbide for light-emitting diode (LED) founded a start-up company, Cree. Given their ties to the local community, Cree was founded on the outskirts of the North Carolina State campus, and, over the next 25 years, established itself as a leader in the emerging LED industry, employing more than 6000 workers (mostly in the local region).

The purpose of this paper is to assess the impact of research institutions on the overall quantity and quality of local entrepreneurship. Focusing on the overall impact of research institutions on entrepreneurship in regions proximate to those research institutions allows us offer an integrated analysis of questions of theoretical and policy interest. On the one hand, both universities and national laboratories serve as a source of local demand and produce frontier knowledge that might form the basis of new companies. But, while the impact of research institutions as a source of local demand is likely to enhance the quantity of entrepreneurs, the ability to leverage frontier research can play a critical role in the formation of companies with a high potential for significant growth at founding. In other words, public research institutions may not simply increase the quantity of entrepreneurs but also their average quality. As well, though both universities and national laboratories conduct research at significant scale, universities are distinctive in also producing students (who might launch careers as entrepreneurs or start-up employees in the local area) and being governed by policies and rules that encourage openness and diffusion to the local environment (Valero and Van Reenen, 2019).

While a significant body of work emphasizes the importance of geographically mediated academic entrepreneurship as an outcome of university and national laboratory research (Kenney, 1986; Feldman, 1994; Audretsch and Stephan, 1996; Zucker, Darby and Armstrong, 1998), this prior by and large focuses on establishing the linkage between the presence of a specific type of research (e.g., life sciences) conducted at a university or national laboratory and subsequent entrepreneurship either initiated by individuals directly connected to the research activities or in industries that draw directly on knowledge related to that research (Stern, 1995; Zucker, Darby and Armstrong, 1998; Audretsch, Lehmann and Warning, 2005; Baptista and Mendonça, 2010; Babina *et al.*, 2020; Hausman, 2020). Though informative about the potential role that specific areas of research play in shaping subsequent local entrepreneurship, these focused sectoral studies do not provide evidence about the overall impact of research institutions on the overall level and nature of entrepreneurship. Specifically, research institutions such as universities engage in many different activities, including hosting a diverse research portfolio over many fields, educating students, and hosting sporting events and conferences. Focusing exclusively on the impact of a particularly fertile research area (such as biotechnology) may be evidence for the impact of a particular area of research at a particular time without providing evidence for the overall impact of research institutions on the local economic environment. In other words, the average impact of research activities by public research institutions on entrepreneurship may be very different than the impact linked to specific discoveries or inventions.

Rather than track the impact of specific research trajectories (or individuals), we focus on the impact of the resources of research institutions on the level and nature of entrepreneurship proximate to that research institution. Specifically, we assess how the activities of a research institution in a particular location influence both the rate and nature of entrepreneurship adjacent to that location. To do so we must address two broad challenges. First, while both theoretical and policy discussions of the economic developments of entrepreneurship focus on the impact arising from a small number of highly skewed growth outcomes (e.g., firms such as Cree that grow to become large employers within the local economy), most entrepreneurs found their firm with neither the intention nor the resources for significant growth (Schoar, 2010; Hurst and Pugsley, 2011). As a source of local economic demand, research institutions may encourage local entrepreneurship without advancing the potential for economic development. However, some of the activities of research institutions – such as the production of scientific discoveries, technological innovations, and students with frontier knowledge – may induce not only a higher quantity of entrepreneurship but also entrepreneurship with a higher quality. Capturing these broader impacts requires capturing not only the quantity of entrepreneurship – the number of new businesses within a given location – but also the quality of entrepreneurship – the potential for growth of these new businesses at their founding.

Second, even if one is able to measure both the quantity and quality of entrepreneurship, the co-location of research institutions and the strength of local entrepreneurial ecosystems is not random. Most notably, there may be common infrastructure, amenities, or knowledge that influence both the activities of the research institution as well as local entrepreneurship. Simply comparing the quantity and quality of entrepreneurship in locations proximate to research institution with locations that are not proximate to a research institution cannot provide direct evidence for the impact of research institutions on local entrepreneurship. Moreover, even if one tracks the activities of research institutions over time, it is possible that the activities of the entrepreneurial ecosystem are itself influencing the activities of the research institution. As such, an estimate of the causal impact of research institutions on local entrepreneurship requires both variation impacting the level of activities of the research institution independently of local entrepreneurship, as well as an approach that accounts for the dynamic impact of local entrepreneurship itself. At a broad level, the experiment we would like to assess is whether activity at the research institution (e.g., the development of a new research trajectory whose quality merits Federal funding) itself influences the quantity and quality of subsequent entrepreneurship in the ecosystem surrounding that research institution.

We tackle these challenges in three interrelated steps, First, leveraging data from the Startup Cartography Project (Andrews *et al.*, 2020), we develop a consistent set of zip-code-level measures of both the quantity and quality-adjusted quantity of entrepreneurship across 34 U.S. states from 1988-2012. Second, for each of these zip codes, we identify both the presence and absence of universities and national laboratories co-located nearby to that zip code, as well as a rich set of covariates that allow us to observe zip codes that are otherwise “similar” to those that are located next to a research institution. Finally, we take advantage of the fact that, for a given research institutions (university or national lab), there are quantitatively significant changes over time in the allocation of Federal funds (both research-oriented and more general) to that institution, and the allocation of the bulk of these funds is through competitive grant processes that is independent of future changes to the local entrepreneurial ecosystem (Federal research funding is driven by grants based on scientific merit). For example, after subtracting out research institution fixed effects and year fixed effects, more than one-third of the variance in research funds allocated to universities is idiosyncratic variation to that research institution over time (above and beyond any overall time trends).

We leverage these data to undertake two broad types of analyses. First, we undertake a cross-sectional evaluation of the link between research institutions and both the quantity and quality-adjusted quantity of entrepreneurship across locations. The cross-sectional analysis points to the broad colocation of university and a high quality-adjusted quantity of entrepreneurship, particularly in areas that are urban, in historically high-income zip codes, and for zip codes collocated with a top-tier university. Moreover, there has been a sizeable shift in the link between research institutions and entrepreneurship over time. First, regardless of whether there is a university or national lab in a location, the types of locations that host research institutions seem to have become more associated with the quantity of entrepreneurship over time. However, even after accounting for these underlying place-based characteristics, there is a sharp (and cyclical) uptick in the association between universities and the quality-adjusted quantity of entrepreneurship.

Building on these descriptive findings, we then turn to the heart of our analysis, where we focus on the causal impact of changes over time in the allocation of Federal resources (in terms of research and non-research funding) on the subsequent level of entrepreneurship. To address the main potential sources of bias, we implement both a fixed effects panel data estimator, as well as first-differences instrumental variables estimators that account for the potential for dynamics within a given entrepreneurial ecosystem. The results are striking. On the one hand, we find that a positive shock in the allocation of Federal resources (both research and non-research, to universities or national laboratories) is associated with a meaningful increase in the quantity of entrepreneurship. This is in line with the idea that, as large local economic institutions, universities induce local demand for entrepreneurship of all types.

However, when we turn to an analysis of the impact of Federal resources on the quality-adjusted quantity of entrepreneurship (i.e., accounting for the potential for growth and spillovers from these start-ups), we find a very sizeable positive impact of Federal research expenditures towards universities but a negative impact on non-research resources and a small (and negative) impact on research resources to national laboratories. Taken at face value, the estimates suggest that the ideation of an idea within a university that is able to attract \$1.56 million of Federal support induces an incremental level of quality-adjusted entrepreneurship to result in a one *successful* growth outcome from the cohort of start-ups founded in that given university ecosystem (i.e., within the zip codes immediately adjacent to that university) over the next six years. This quantitative impact is particularly striking given that, though highly skewed, the private returns to successful growth-oriented entrepreneurship (conditional on a successful exit (i.e., an IPO or acquisition worth more than \$10 million) are extremely high: a successful growth outcome has an average (though skewed) private value of more than $50 million) (Hall and Woodward, 2007). Moreover, given that the idiosyncratic component of Federal research support for a zip code within our dataset is more than $10 million, our findings suggest that the development of new research within universities that attract Federal research support have a meaningful near-term impact on the subsequent quality of entrepreneurship within local entrepreneurial ecosystems.

Overall, these findings document the distinct impact of universities and national laboratories on local entrepreneurial ecosystem. There are two key distinctions that are worthwhile to note. First, there seems to be a meaningful distinction between research activity and non-research activity at universities; whereas non-research activity raises the quantity of entrepreneurship (but lowers its average quality), research activity increases both the quantity and quality-adjusted quantity of entrepreneurship. Second, there is an important distinction between universities and national laboratories. While these are the two primary institutions conductions publicly funded research in the United States, university research activities seem to be more closely linked to spillovers to local entrepreneurial ecosystems, whereas the more cloistered nature of national laboratories is associated with a much more muted (or even perhaps negative) impact. These results highlight the important and unique role universities play in building and sustaining entrepreneurial ecosystems. Because of their norms and missions, universities seem to be able to spur economic development more widely than other research institutions, and also beyond the simple fact of being located in high-growth locations. For example, our findings offer broad support for recent policy proposals such as the 2021 Endless Frontiers Act that prioritize the role of university research in place-based policy (Gruber and Johnson, 2019; Senate of the United States, 2021).

II. The Impact of Public Research Institutions on Entrepreneurship

The central objective of this paper is to consider the empirical impact of research universities and national laboratories on entrepreneurship. We begin our analysis by briefly synthesizing g the distinctive economic properties of these public research institutions, and then draw out the implications of these properties for entrepreneurship in the context of a simple model.

It is useful to start by highlighting a central feature of the US national innovation system: the highly decentralized performance of publicly funded scientific and engineering research across a wide variety of institutions and domains (Mowery and Rosenberg, 1993; Rosenberg and Nelson, 1994). In particular, though the classical rationale for public investment in research arising from lack of appropriability is consistent with simply funding the research activities of private firms (Nelson, 1959; Arrow, 1972), the majority of public funding of research (as opposed to development) is concentrated in public and non-profit institutions, most notably universities and national laboratories. There are more than 260 research universities scattered across the United States receiving more than 30 billion USD annually from public research funds, and 46 national laboratories engaging in research with more than 19 billion USD research funds per year. When research is performed in an institution such as a national laboratory or university that only conducts research but face significant restrictions in direct commercialization, the social return from that public research investment depends crucially on whether research is transferred to the private sector. Though direct transfer to established firms is an important channel for such knowledge transfer (Furman and MacGarvie, 2007), the establishment of new firms leveraging new knowledge may be a particularly important channel for realizing the benefit of publicly funded research (Baptista, Lima and Mendonça, 2011; Kim, Kim and Yang, 2012).

As research-enhancing economic institutions (Furman and Stern, 2011; Valero and Van Reenen, 2019), universities and national labs share a number of distinctive properties and also display some key differences that shape their potential impact on entrepreneurship. On the one hand, both universities and national laboratories engage in a wide variety of different types of research, including both basic and applied research, and there are examples of important fundamental breakthroughs (as well as more modest discoveries) from both types of institutions. As well, both employ a workforce highly skewed towards those with a high level of formal training in scientific and engineering fields, the bulk of public research funding is assigned to individualized researchers pursuing well-defined research projects at a moment in time, and, to first approximation, researchers are allowed to disclose their research findings to the broader scientific and engineering community through publication and other types of disclosure (Dasgupta and David, 1994; Stephan, 2012).

Importantly, both universities and national laboratories are geographically bounded, with physical campuses in specific locations (and usually only a small number of satellite locations). Interestingly, the place-based nature of these research-enhancing institutions suggests that they will serve potentially as a source of knowledge for potential local spillovers. As well, universities and national laboratories are important economic agents within a given geography (e.g., as a source of local demand). As a source of economic density, universities concentrate demand in a particular place, thus encouraging the formation of new businesses near that place. This channel is of course particularly salient at the earliest stages of a university (Andrews, 2020) or in response to a change in the scale of university activities (i.e., in equilibrium, there would be a sustained flow of new businesses, but not necessarily different from other places).

Despite these significant similarities between universities and national laboratories, a number of key distinctions point towards the potential for differences in their likely impact on entrepreneurship in the local environment. Most notably, whereas universities combine the education and training of students alongside the pursuit of research, national labs by and large are focused more narrowly on research (Jaffe and Lerner, 2001; Adams, Chiang and Jensen, 2003). Not simply a difference in the scope of their activities, students may play an important (and often underappreciated) role in transferring knowledge produced by universities into private sector activity, including the founding of new firms by students directly as well as the hiring of students into start-ups for the purpose of leveraging their frontier knowledge.

Though the presence or absence of students and an educational function is likely the primary contrast between universities and national laboratories, universities also likely maintain practices and offer incentives that enhance the fluidity between research and industry, and may focus on a broader range of research problems that may be more applicable for local economic development. The most obvious channel through which knowledge flows from universities and other research organizations involves scientific research published in academic journals. This research is produced locally but it is distributed globally, and it is in principle available for anyone to use, independent of their geographical location. Yet, empirical evidence points to a different pattern. Jaffe and colleagues (1993) show that patents tend to cite other patents produced by organizations (both universities and corporations) that are located nearby. Even more strikingly, Adams (2002) finds that knowledge flows from universities tend to be much more local in nature than spillovers from firms, highlighting the apparent paradox that institutions whose mandate is to produce public knowledge, such as universities, tend to benefit disproportionately local businesses. Adams goes on arguing that it is precisely because of the open nature of the knowledge that is produced by universities, that we observe firms gravitating around academic institutions: as knowledge and information, especially if they are highly tacit in nature, do not transmit without costs, firms locate close to universities to absorb knowledge which is “reasonably current and not propriety” (Adams, 2002, p. 274). The importance of oral transmission of knowledge is indeed one of the main reasons why people tend to cluster in cities notwithstanding the increased costs (Lucas, 1988). This channel, which is present for all firms, may be even more relevant for start-ups (Audretsch and Thurik, 2001) as they rely more heavily on externally produced knowledge (Hall, Link and Scott, 2003) due to the lower amount of resources they can devote to internal R&D. In other words, while both universities and national labs conduct research, the dual mission of universities and related practices may have the consequence that universities are more likely to serve as a source of knowledge for start-up activity in the local environment, at least in the short to medium term.

These distinctive properties of universities and national laboratories as local economic institutions have distinctive implications for both the quantity and quality-adjusted quantity of local entrepreneurship. Consider a simple stylized model in which within a location *r* there is a potential set of risk-neutral entrepreneurs, normalized to be equal to a unit mass. Each entrepreneur *i* considers two alternative opportunities, *L* and *H*. *L* is a local opportunity whose net operating profits are determined by the size of the local market, d_r, while *H* is a knowledge-based opportunity whose value is determined by θ_r^H (the quality of the accessible local knowledge stock) and p^H (the probability of success in the knowledge-based opportunity conditional on pursuing that opportunity). For simplicity, the quality of the local opportunity does not depend on the quality of the local knowledge stock, and the quality or probability of success of the knowledge-based opportunity does not depend on the size of the local market. Finally, entrepreneurs differ in the skills and resources they bring to entrepreneurship, and so each entrepreneur faces an idiosyncratic fixed cost of pursuing entrepreneurship, κ_i, identical for both *L* and *H*, with $\kappa_i \sim f(\kappa;\bar{\kappa})$.[1]

Each potential entrepreneur chooses whether to pursue the local opportunity, the knowledge-based opportunity, or neither opportunity (e.g., paid employment). Normalizing the returns to paid employment equal to 0, the relative returns to *L* are simply $d_r - \kappa_i$ while the returns to *H* are $\theta_r^H p^H - \kappa_i$. Before turning to the choice of whether to choose entrepreneurship, this very simple model yields a sharp result in terms of whether to pursue *L* or *H*: if $d_r > \theta_r^H p^H$, then entrepreneurs will pursue the local opportunity and, in the converse, the knowledge-based opportunity. The quantity of entrepreneurship will then be determined by the value of the preferred opportunity versus the fixed cost of entry with a cut-off equal to to $\kappa_i < \kappa_r^* = \max\{d_r, \theta_r^H p^H\}$. This cut-off determines the overall quantity of entrepreneurship, $\varphi_r = \int_0^{\kappa_r^*} f(\kappa)d\kappa$. However, the quality-adjusted quantity of entrepreneurship, ρ_r, will be 0 if $d_r > \theta_r^H p^H$ and $\rho_r = \varphi_r = \int_0^{\theta_r^H p^H} f(\kappa)d\kappa$ if $d_r \leq \theta_r^H p^H$. Though simplified in the extreme, this structure highlights that while the quantity of entrepreneurship is shaped by the overall attractiveness of both local and knowledge-based opportunity, the quality-adjusted quantity of entrepreneurship (i.e., the incidence of entrepreneurship including the probability

---

[1] It is likely that the initial sunk costs of pursuing a local opportunity are different than that of pursuing a knowledge-based opportunity; however, as long as $\kappa_i^H = \chi\kappa_i^L$ (i.e., the idiosyncratic sunk costs are proportional), the analysis is identical (e.g., this would be the same as a "scaling" of $\theta^H = 1/\chi$).

of significant "growth") will be shaped by both the overall quality of growth opportunities (i.e., $\theta_r^H p^H$) but also the relative returns to local versus knowledge-based entrepreneurship.

We can then use this model to identify the impact of shifts in the level of resources available to a local entrepreneurial ecosystem, either in terms of the overall level of local demand or the level of knowledge itself. First, a shift in d_r or θ_r^H (weakly) increases the quantity of entrepreneurship by increasing the size of the local market or the quality of knowledge-based opportunities, respectively. Essentially, regardless of whether the ecosystem is in the *L* or *H* regime, expanding the level of resources or knowledge makes entry more attractive for marginal entrant, and so increases the overall quantity of entrepreneurship. And, since an increase in the θ_r^H only induces entry into the *H* opportunity, also increases the quality-adjusted quantity of entrepreneurship $(\frac{\partial \rho_r}{\partial \theta_r^H} \geq 0)$. However, because ρ_r is shaped not only by the absolute level of resources and knowledge but also the relative returns to each, an increase in the size of the local market, d_r, has the potential to shift the ecosystem from the *H* to the *L* regime $(\frac{\partial \rho_r}{\partial d_r} \leq 0)$. In other words, a simple increase in local resources without an increase in knowledge production itself may actually result in substitution away from knowledge-based start-ups and towards local opportunities.

Finally, increasing θ_r^H requires not only increasing the level of knowledge produced in ecosystem *r* but also ensuring that this knowledge is accessible to potential entrepreneurs in *r*. As discussed earlier, two distinctions between universities and national laboratories make it more likely that universities may exert a higher level of knowledge transfer on their local entrepreneurial ecosystem. First, universities not only produce research but also train and produce students, who may be important conduits for knowledge transfer into the local entrepreneurial ecosystem. As well, while many national laboratories maintain a certain level of insularity, universities are by and large more "open" in the sense of engaging with their local ecosystems and promoting local economic development and entrepreneurship. As such, we hypothesize that the impact of a shift in research resources for a public institutions on local quality-adjusted entrepreneurship will be larger for universities than for national laboratories.

III. The Measurement of Entrepreneurship[2]

---

[2] This section builds on Andrews, et al (2020), which itself builds on Guzman and Stern (2015, 2017, 2020), which introduces the details regarding state-level business registration records and the predictive analytics approach used to measure the quantity

The first task that we face is how to develop consistent measures of entrepreneurial quantity and quality on a consistent and granular basis so that we can examine how the presence or absence of a research institution, or changes in the economic and research activities of research institutions over time, influence subsequence entrepreneurship within a given local area. To do so, we leverage data from the Startup Cartography Project (Andrews *et al.*, 2020). Building on Guzman and Stern (2015, 2017), the Startup Cartography Project data leverage a systematic approach for measuring the founding quality of startups using predictive analytics and business registration records. The approach combines three interrelated insights. First, as the challenges to reach a growth outcome as a sole proprietorship are formidable, a practical requirement for any entrepreneur to achieve growth is business registration (as a corporation, partnership, or limited liability company). This practical requirement allows to form a population sample of entrepreneurs “at risk” of growth at a similar (and foundational) stage of the entrepreneurial process. Second, it is possible to distinguish among business registrants through the measurement of founding choices observable at or close to the time of registration. For example, Guzman and Stern measure start-up characteristics such as whether the founders name the firm after themselves (eponymy), whether the firm is organized in order to facilitate equity financing (e.g., registering as a corporation or in Delaware), or whether the firm seeks intellectual property protection (e.g., a patent or trademark). Third, though rare, some firms experience observable meaningful growth outcomes (such as achieving an IPO or high-value acquisition). Combining these insights, it is possible to estimate entrepreneurial quality by estimating the relationship between observed growth outcomes and start-up characteristics.

That is, for a firm *i* born in region *r* at time *t*, with at-birth start-up characteristics K_{irt} , we observe growth outcome $g_{ir(t+l)}$ *l* years after founding and estimate:

$$\theta_{irt} = P(g_{ir(t+l)} = 1 | K_{irt}) = f(\psi K_{irt})$$

$$\theta_{irt} = P(g_{ir(t+l)} = 1 | K_{irt}) = f(\psi K_{irt}) \tag{1}$$

and use the *predicted* value of this regression the measure of entrepreneurial quality. As long as the process by which start-up characteristics map to growth remain stable over time (an assumption which is itself testable), this mapping allows to form an estimate of founding characteristics to entrepreneurial quality for any business registrant within any given sample.

---

and quality of entrepreneurship.

$$\hat{\theta}_{irt} = f(\hat{\psi} K_{irt}) \tag{2}$$

These estimates can be used to generate three core entrepreneurship statistics capturing the level of entrepreneurial quantity, and the potential for growth entrepreneurship within a given geographical area and start-up cohort.

Entrepreneurial Quantity and Quality Indices (N and EQI). With a set of business registrants over a given time period within a given geographic region, it is straightforward to measure entrepreneurial quantity, N_{rt}, which is simply equal to the count of all new business registrations in ZTCA *r* and year *t.*

The predictive analytics approach described above is then combined with this population of business registrants to create an index of entrepreneurial quality for any group of firms (e.g., all the firms within a particular cohort or a group of firms satisfying a particular condition). Specifically, the Entrepreneurial Quality Index (EQI) is defined as an aggregate of quality at the ZTCA-year level by simply estimating the average of $\hat{\theta}_{irt}$ over that ZTCA:

$$EQI_{rt} = \frac{\sum_{i=1}^{N_{rt}} \hat{\theta}_{irt}}{N_{rt}} \tag{3}$$

To ensure that the estimate of entrepreneurial quality for a region *r* reflects the quality of start-ups in that location rather than simply assuming that start-ups from a given location are associated with a given level of quality, location-specific measures are excluded $H_{r,t}$from the vector of observable start-up characteristics.

The Regional Entrepreneurship Cohort Potential Index (RECPI). From the perspective of a given geographical area, the overall inherent potential for a cohort of start-ups combines both the quality of entrepreneurship in a region and the number of firms in such region (a measure of quantity). *RECPI* is therefore defined as simply *EQI* multiplied by the number of firms in that region-year:

$$RECPI_{rt} = EQI_{rt} N_{rt} \tag{4}$$

Since this index multiplies the average probability of a firm in a region-year to achieve growth (quality) by the number of firms, it is, by definition, the expected number of growth events from a region-year given the start-up characteristics of a cohort at birth. This measure of course abstracts away from the ability of a region to realize the performance of start-ups founded within a given cohort (i.e., its ecosystem performance), and instead can be interpreted as a measure of the "potential" of a region given the "intrinsic" quality of firms at birth, which can then be affected by the impact of the entrepreneurial ecosystem, or shocks to the economy and the cohort between the time of founding and a growth outcome. Together, this use of business registration data allows the construction of measures of both entrepreneurial quality and quantity-adjusted quality, giving us the ability to undertake a time-sensitive and granular evaluation of the interplay between research institutions and entrepreneurial activity.

IV. Empirical Approach

The primary objective of our analysis is to leverage the SCP measures to examine the impact of public research resources on the quantity and quality-adjusted quantity of entrepreneurship. Our data are composed of zip-code level measures of the quantity and quality-adjusted entrepreneurship over time at a granular geographic level, these geographic units are co-located (or not) with a public research institution (university or national laboratory or both), and measures of the resources utilized by those public research institutions (including their overall budget, but also specific time-varying measures of resources such as the annual level of Federal research funding allocated to a given university). In an ideal case, we would be simply able to allocate public research resources randomly (e.g., to some locations but not others) and then measure the quantity and quality-adjusted quantity of entrepreneurship over time for those locations that had received public research resources versus those that had not. Indeed, the first stage of our empirical analysis is purely descriptive, and simply documents the differences in the quantity and quality of entrepreneurship for locations that are proximate or not to universities, with a focus on how this cross-sectional empirical relationship has changed over time.

However, to assess the causal impact of universities and national laboratories on their local entrepreneurial ecosystem (i.e., on entrepreneurship proximate to that research institution), we must overcome three distinct but related empirical obstacles. First, the co-location of public research institutions and vibrant entrepreneurial ecosystems is not random. While it is instructive in a descriptive sense to simply examine the cross-sectional correlation between the presence of public research institutions and entrepreneurship, it is possible that the presence of a positive correlation may be driven by potentially unobserved (to the econometrician) location-specific infrastructure or amenities. For example, the area surrounding Boulder, Colorado contains both the University of Colorado (which might influence that ecosystem) but also has specific private sector initiatives (such as the entrepreneurial accelerator TechStars) as well as other nearby institutions (e.g., the National Renewable Energy Laboratory is only 25 miles away). Both public research institutions and strong entrepreneurial ecosystems are likely elements of the "Creative Economy" (Florida, 2002), and the correlation between public research activity and entrepreneurship may be driven by selection rather than a causal link between the two.

Second, even if one controls for selection of public research institutions into favorable locations (e.g., through location-specific fixed effects), the activities of these public research institutions themselves may be endogenous to the activities of the local entrepreneurial ecosystem. Knowledge produced by firms within a local entrepreneurial ecosystem might simultaneously promote entrepreneurial activity within that ecosystem, and also exert spillovers onto university or national laboratory research itself. For example, the strategic shift by Monsanto (headquartered in St. Louis) into agricultural biotechnology research not only impacted that firm but also resulted in a shift towards agricultural biotechnology research at co-located research institutions such as Washington University of St. Louis (Sohn, 2020). More generally, a discovery or invention made by the entrepreneurs in a given location might facilitate both a higher level of research output at the university and at the same time serve as a spur for subsequent local entrepreneurship. The potentially bidirectional nature of impact between public research institutions and local entrepreneurial activity means that a simple examination of positive correlation over time of public research institutions research and local entrepreneurial activity cannot identify the causal linkage of public research institutions on entrepreneurship.

And, finally, entrepreneurial ecosystems are themselves dynamic. A positive shock to the quantity and quality of entrepreneurship in a given location (e.g., as the result of favorable economic conditions) may not only influence the current rate of entrepreneurial activity but may also exert a positive influence on the follow-on period. This would occur, for instance, if there was a potential for a two-period lag between the an initial unobserved shock and its influence on the rate of entrepreneurship (e.g., increasing the rate most saliently in the period immediately following the period but also having a smaller influence on entrepreneurship in the second period after the shock). Accounting for this dynamic would result in a lagged dependent variable structure. Specifically, if the lagged dependent variable is itself correlated with the changing level of local university research, abstracting away from the dynamic impact of entrepreneurial ecosystems themselves might result in overestimating the impact of university research itself.

To address these interrelated empirical challenges, we leverage the fact that the level of Federal research resources allocated to any given public research institution varies considerably over time.

As we discuss in more detail in Sections V and VI, while the mean level of Federal research expenditures to a university within our sample is \$77 million, the standard deviation of those expenditures is \$112 million. More subtly, the standard deviation of the "residual" based on a regression of Federal research expenditures that includes a complete set of university-level and year fixed effects is \$38 million. In other words, even after accounting for differences across universities and secular trends across time, there is significant variation within a given university in the overall level of research funding received in a given period.

While the receipt of Federal research funding is of course not randomly allocated (indeed, it is dependent on the quality of the ideas and the grantsmanship of public researchers), our key assumption is that the *changes* over time in the level of overall Federal funding (research and otherwise) to a given university or national laboratory is exogenous to the existing quantity and quality of the local entrepreneurship. Essentially, our key assumption is that, conditional on the time-invariant quality level of a given ecosystem, there is random variation over time in the ability of individuals within that research institution (e.g., university faculty) to develop and attract Federal funding for their activities. Put another way, our analysis is premised on the idea that the ability to attract a Federal grant to a particular institution indicates an increase in both the resources and the knowledge produced by that institution. As such, a consistent estimate of the impact of those additional resources on subsequent local entrepreneurship can be interpreted as the impact of the time-varying strength of that institution on the local environment.

To take advantage of this variation over time in the level of Federal resources allocated to universities and national laboratories, we first link our zip-code-level measures of the quantity and quality-adjusted quantity of entrepreneurship to the presence or absence of universities themselves. To do so, we define a zip-code r in state m to be proximate to a university or national laboratory if the center of that zip code is within five miles of the university or national laboratory (the next section details this procedure in greater detail). For each university or national laboratory, we observe a vector of inflation-adjusted annual Federal resource commitments, X_{rt}, which includes the level of research funding to a university, the level of non-research funding to the university, and the overall Federal budget allocation to a given national laboratory. In order to allow a meaningful gap of time to elapse between the allocation of Federal funding and its impact on the local entrepreneurial ecosystem, we group the annual zip-code-level data from the SCP, as well as measures of Federal resource allocation, into a smaller number of multi-year periods (either 2 years or 3 years). Denoting each of these multi-year periods by s, we can thus observe a multi-year measure of local entrepreneurial activity, Y_{rs} (measured as either N_{rs} or $RECPI_{rs}$), for each zip code, as well as measures of the level of resources provided by the Federal government to universities or national laboratories in that zip code, X_{rs}, and other time-varying location-specific characteristics, Z_{rs}.

We implement two empirical approaches to estimate the impact of changes over time in the Federal resources allocated to research institutions, X_{rs-1}, on the quantity and quality-adjusted rate of entrepreneurship in *s*. We first examine a simple fixed effects panel estimator:

$$Y_{rs} = \alpha_r + \beta_s + \delta X_{rs-1} + \gamma Z_{rs-1} + \varepsilon_{rs} \tag{5}$$

This simple specification allows us to estimate the relationship between entrepreneurship and the lagged level of resources, while controlling for fixed differences between locations and year in terms of the rate of entrepreneurship, and controlling for other observable location-specific measures. This specification goes beyond a simple cross-sectional correlation between the colocation of research institutions and entrepreneurship by focusing exclusively on variation in resources that are structurally independent from local entrepreneurship: changes in the level of Federal resources are the result of grant writing rather than the result of a flow from entrepreneurship towards the university. However, this specification does not address the potential for dynamics within an entrepreneurial ecosystem insofar as it is possible that the rate of entrepreneurship is not only impacted by the level of resources from the university but also the existing level of entrepreneurship itself. In principle, we can account for this possibility by including the lagged level of dependent variable:

$$Y_{rs} = \alpha_r + \beta_s + \delta X_{rs-1} + \gamma Z_{rs-1} + \varepsilon_{rs} \tag{6}$$

Though an appealing method for accounting for dynamics, the assumptions required for consistent estimation of (6) using a simple fixed effects panel estimator are demanding (Angrist and Pischke, 2008, p. 243): (a) the elements of *X* and *Z* must themselves by serially independent (i.e., conditional on α_r, there can be no correlation over time in the *X* and *Z* themselves), and (b) the elements of ε_r must be independent of each other (i.e., the shocks are not only independent of *X* and Z, but there are also no dynamics to the shocks over time), and. Specifically, in the case where there is positive correlation across sequential time periods (e.g., an AR(1) process)), then (6) is likely to result in an overestimate of (the absolute value of) δ (our parameter of interest).

We can significantly relax the assumptions required for identification of δ by transforming (6) through first-differences:

$$\Delta Y_{rs} = \Delta\beta_s + \lambda\Delta Y_{rs-1} + \delta\Delta X_{rs-1} + \gamma\Delta Z_{rs-1} + \eta_{rs} \tag{7}$$

Specifically, this specification allows us to assess how *changes over time* in the resources allocated to a research institution influences subsequent *changes* over time in the quantity or quality-adjusted quantity of entrepreneurship. And, since we are focusing only on changes in the level of X and Z, this specification allows for correlation over time in the level of X and Z. However, both ΔY_{rs-1} and η_{rs} contain ε_{rs-1} and so there is a mechanical (negative) correlation between ΔY_{rs-1} and η_{rs}; as such, the estimates from an OLS implementation are biased (Nickell, 1981; Arellano and Bond, 1991). To address this, we can use the level of Y_{rs-2} as an instrument for ΔY_{rs-1} to implement a consistent estimator for (3) (Anderson and Hsiao, 1981).[3] As highlighted by Angrist and Pischke (2008), this approach may lead to an underestimate of the (absolute value of) δ due to attenuation bias.

Putting these ideas together, the key empirical challenges of estimating the impact of universities on local entrepreneurial ecosystems arises from the fact that the colocation of universities and start-ups is neither random nor unidirectional, and local entrepreneurial ecosystems are themselves dynamic. To address these concerns, we take advantage of the fact that Federal resources towards local research institutions are essentially independent of start-ups in that region (conditional on the pre-existing level of entrepreneurship), and we will take advantage of this insight by considering the impact of changes over time in Federal resources to local research institutions. To account for different potential biases, our analysis compares both fixed effects and panel data estimators to assess the robustness of our core results to different sets of assumptions required for consistent estimation of this fixed effects dynamic panel data model.

V. Data

Implementing our approach requires the combination of systematic measures of quantity and quality-adjusted quantity of entrepreneurship at a granular geographic level with measures of universities and national laboratories activities over time, as well as locational characteristics (e.g., urban versus rural, income, etc). We therefore combine data at the zip code level from the Startup Cartography Project, the National Center for Education Statistics, the NSF Survey of R&D

---

[3] Accounting for the interplay between dynamics over time within a panel and the presence of fixed differences is complex and involves tradeoffs (Angrist and Pischke, 2009, Section 5.3). To simplify exposition and focus on estimates that most directly connect to our research question, we focus on the fixed effects panel estimator and OLS and instrumental variables first-differences specifications. We have also explored using an Arellano-Bond GMM estimator (which allow for more efficiency by using all the lags of the levels of dependent and independent variables as instruments for the changes in those variables over time); however, we found that the estimates varied considerably depending on the precise specifications (and in ways that were not easily interpretable except perhaps reflecting the imprecision arising from a "weak instruments" problem).

Expenditures, Census socioeconomic data, as well as auxiliary sources concerning research university quality and type.

A. *The Startup Cartography Project*[4]

The first building block of our dataset consists of systematic and granular measures of the quantity and quality-adjusted quantity of entrepreneurship drawn from the Startup Cartography Project (Andrews *et al.*, 2020). Building on Guzman and Stern (2015, 2017, 2020), these measures are calculated by aggregating, to the zip code level, state-level business registration data with the predictive analytics approach discussed in Section 2 allowing for an estimate of entrepreneurial quality for all new registrants at founding. The data are drawn from 32 states (corresponding to more than 81% of the national GDP) from 1988 to 2012, and include, for any year, all new business registrants that are either a for-profit firm in the local jurisdiction, or a for-profit firm registered in Delaware but whose principal office address is in the focal state.[5]

For each business registrant, the SCP constructs variables related to: (i) growth outcomes (IPO or significant acquisition); (ii) firm characteristics based on business registration observables; (iii) firm characteristics based on secondary data that can be directly linked to the firms itself, such as patents and trademarks.

i. *Growth Outcomes*. The growth outcome, *Growth,* is a dummy variable equal to 1 if the firm has an initial public offering (IPO) or is acquired at a meaningful positive valuation within 6 years of registration as reported in the Thomson Reuters SDC database.[6]

ii. *Firm Characteristics Based on Business Registration Data.* Two binary measures relate to how the firm is registered: *Corporation*, which captures whether the firm is a corporation rather than an LLC or partnership, and *Delaware*, equal to one if the firm is registered in Delaware. Five additional measures are furthermore based directly on the name of the firm. *Eponymous* is equal to 1 if the first, middle, or last name of the top managers is part of the name of the firm itself, which may be associated with lifestyle businesses rather than growth-oriented entrepreneurs (Belenzon, Chatterji and Daley, 2017, 2020), and *Short Name* is equal to one if the entire firm name has three or less words, and zero otherwise (most growth-oriented ventures have names of three words or

---

[4] This section builds on Andrews, et al (2020), which itself builds on Guzman and Stern (2015, 2017, 2020), which introduces the details regarding state-level business registration records and the predictive analytics approach used to measure the quantity and quality of entrepreneurship.

[5] While the updated version of SCP (Andrews, et al, 2020) contains data for 48 states, we focus on the 32 states for which data were available in 2019 (in line with Guzman and Stern (2020)).

[6] Although the coverage of IPOs is essentially comprehensive, the SDC data set excludes some acquisitions. Guzman, et al (2020) and Andrews, et al (2020) discuss robustness to alternative datasets and measures of growth outcomes.

fewer). Seven measures are based on how the firm name reflects the industry or sector within which the firm is operating, taking advantage of the industry categorization of the US Cluster Mapping Project ("US CMP") (Delgado, Porter and Stern, 2016) and a text analysis approach. The first three are associated with broad industry sectors and include whether a firm can be identified as local (*Local*), traded (*Traded*) or resource intensive (*Resource Intensive*). The other five industry groups are narrowly defined high technology sectors that are typically associated with high growth firms, including whether the firm is within the biotech (*Biotech Sector*), e-commerce (*E-Commerce*), other information technology (*IT*), medical devices (*Medical Devices*) or semiconductors (*Semiconductor*) space.

iii. *Firm Characteristic Based on External Observables.* Two measures are related to intellectual property measures based on data from the U.S. Patent and Trademark Office. *Patent* is equal to 1 if a firm is associated with a patent application within the first year and 0 otherwise, including patents filed by the firm within the first year of registration and patents that are assigned to the firm within the first year from another entity (e.g., an inventor or another firm). The second measure, *Trademark*, is equal to 1 if a firm applies for trademark protection within a year from registration.

Table A.1 in Appendix A reproduces the core logistic regression model from Andrews et al. (2020) which estimates how the presence or absence of a startup characteristic correlates with the probability of growth (conditioning on the presence or absence of other startup characteristics). There is an extremely strong (and robust) correlation between startup characteristics and the probability of growth. Substantial changes in the predicted likelihood of a growth outcome are associated with characteristics observable in real time from business registration records as well as characteristics observable with a lag (e.g., patent and trademark applications). On the one hand, startups founded as corporations are almost 220% more likely to grow. Similarly, firms with short names are 79% *more* likely to grow, while eponymous firms are 69% *less* likely to grow. Finally, the interplay between corporate form and formal intellectual property protection is particularly predictive: a startup that registers in Delaware and applies for a patent in its first year is 83 times more likely to grow than a firm that only registers in its home state and does not apply for intellectual property protection.[7] For each firm, these estimates can be used to calculate the probability of growth at founding. While this measure is low on average (on the order of one in

---

[7] It is important to emphasize that these startup characteristics are not the *causal* drivers of growth, but instead are "digital signatures" that allow to distinguish firms in terms of their entrepreneurial quality. Registering in Delaware or filing for a patent will not guarantee a growth outcome for a new business, but the firms that historically have engaged in those activities have been associated with skewed growth outcomes

1400), firms with particular combinations of startup characteristics are far more likely to grow (e.g., firms in thet top 1 percent of the out-of-sample estimated quality distribution have a better than 1 in 100 chance of achieving a growth outcome).

These firm-level estimates can then be aggregated at an arbitrary level of geographic and temporal granularity to form annual measures of entrepreneurial quantity and quality-adjusted quality. To capture the local relationship of university and national laboratory activity on local entrepreneurship, we focus our geographic analysis at the ZCTA level.[8] As reported in Table 1, for each year and ZCTA, we construct two core measures: *# of Ventures*$_{rt}$ is simply equal to the number of newly formed ventures in ZCTA *r* in year *t* (mean = 41.8), while *RECPI*$_{rt}$ is *# of Ventures*$_{rt}$ multiplied by the average estimated probability of growth within that cohort (*EQI*$_{rt}$) multiplied by 10,000 (for interpretability). The average probability of a growth event of the startups within a ZTCA-year cohort is 0.029.[9]

Insert Table 1 about here

B. University and National Laboratory Measures

The second component of the data include measures related to universities and national laboratories. We first develop a list and location for all higher education institutions from the National Center for Education Statistics. From this list, we define a Research University (*RU*) those universities rated as either “high research activity” or “very high research activity” (the two top categories of research orientation) according to the Carnegie Classification system (Shulman, 2001). For each research university *j* and year *t*, we gather measures of both Federal and non-Federal sources of funding on an annual basis (calculated in 2012 dollars), including *RU Federal R&D Funds*$_{jt}$ and *RU Federal Non-R&D Funds*$_{jt}$ (from NSF Survey of Federal Funds for Research and Development[10]), and *RU Institutional Funds*$_{jt}$, *RU State Funds*$_{jt}$, *RU Business Funds*$_{jt}$ and *RU Other Funds*$_{jt}$ (from NSF Higher Education Research and Development Survey[11]). As well, for each research university, we include several dummies for university characteristics. We define a

---

[8] Zip Code Tabulation Areas (ZCTA) are generalized area representations of USPS Zip Code service areas (United States Census, https://www.census.gov/programs-surveys/geography/guidance/geo-areas/zctas.html). While standard zip codes represent geographic areas with well-defined geographical boundaries, zip code boundaries are potentially changed over time, and also there are zip codes for PO Boxes, military, and large customers. To fix zip code boundaries over time, we rely on the 2015 Zip Code to ZCTA crosswalk by the HRSA UDS mapper (https://www.udsmapper.org), which contains 41251 unique ZIP Codes, the correspondent ZCTAs, the type of ZIP Code, city and State. All but two ZIP codes in the SCP were matched using the cross walk (which were then assigned manually).

[9] We then group these annual measures into two-year or three-year aggregates when we conduct fixed effects panel and first-differences estimation of the causal impact of changes in Federal resources towards research institutions on the subsequent changes in N and RECPI.

[10] https://www.nsf.gov/statistics/srvyfedfunds/#sd

[11] https://www.nsf.gov/statistics/srvyherd/

research university as being in the *Top 10* or *Top 50* based on the 2017 Times Higher Education rankings (these broad rankings are quite stable over time). We then collected measures of the activities of each university from Wikipedia, including a separate dummy variable for *Business School*, *Medical School*, and/or *Law School*. Finally, we use the NSF NCSES Survey of R&D Expenditures at Federally Funded R&D Centers[12] to both construct a list and location data for all national laboratories and, for each national laboratory *k* in year *t*, measure *Lab Federal R&D* $Funds_{kt}$.

We then aggregate these university and national laboratory measures into measures of the presence and level of activities of research institutions proximate to a given ZCTA. Specifically, for each ZCTA, we create a circle with a 5 mile radius centered at the ZCTA centroid; we then aggregate the number, nature and funding of research universities that fall within the circle for each ZCTA. The median university is proximate to eight zip codes (i.e., half of all universities are proximate to eight or more zip codes), and the median national laboratory is proximate to 8.5 zip codes. As such, any estimate of the impact of the impact of a research institutions on a proximate zip code will then be scaled by the median number of zip codes impacted by institutions of that type. Overall, 6.4% of ZTCA are proximate to a research university, while only 0.6% of ZTCA are proximate to a national laboratory.

Finally, the incidence rate of professional schools is between 0.05 and 0.076, in part because there exist many professional schools that are independent of research universities (and are not themselves research-oriented) and also because professional schools often have separate campuses, often in urban locations. The average ZCTA-level aggregate of *RU Federal R&D* $Funds_{rt}$ is \$10.5 million (with a very large variance across locations), while the average of *RU Federal R&D* $Funds_{rt}$ is only about 11% of this level. *Lab Federal R&D* $Funds_{kt}$ registers and average of \$3.3 million.

C. ZCTA Characteristics

While also report cross-sectional estimates of the university and national laboratory “premium” in terms of the level of entrepreneurial quantity and RECPI (and examine how this premium has changed over time). We include a set of ZCTA characteristics, drawn from the 1990 Census, to control for the presence of locational characteristics that might also explain the premium. These include measures of population, density, male percentage, urban, fraction aged between 18 and 65, white, born outside the U.S., percentage of population holding a college degree, private versus public sector employment, income per sector, and housing value (See Appendix Table B.1 for

---

[12] https://www.nsf.gov/statistics/srvyffrdc/

precise variable definitions and names, means and standard deviations). When undertaking the matching between ZCTA and the 1990 Census, we fail to match 14% of ZCTAs, and so exclude these zip codes throughout the analysis (all of the panel and long-differences results are robust to their inclusion or exclusion). Overall, the final sample consist of 15950 ZCTAs in the time period 1988-2012.

VI. Results

We now proceed to assess the relationship between research institutions and regional entrepreneurship. Our analysis proceeds in two steps. We first consider how the quantity and quality-adjusted level of entrepreneurship varies in the cross-section with the presence or absence of a research institution, and how this relationship varies by institution type, geographic characteristics, and over time. A key insight from this analysis is that while the positive correlation between the presence of a university and entrepreneurial quality was negligible as of the late 1980s, there was a striking increase between 1988 and 2012 in the degree of colocation of research universities and high-quality ventures. We then turn to our core empirical analysis which focuses on how changes in resources to research institutions over time influence subsequent entrepreneurship. Across a broad range of approaches, we document that whereas an increase in any type of funding to research institutions increases the quantity of entrepreneurship, only research funding increases the quality-adjusted quantity of entrepreneurship. The effect sizes implied by these estimates are sizeable: a doubling of expenditures on university research would enhance the overall quality-adjusted quantity of entrepreneurship by 23% in proximate zip codes.

A. The Cross-Sectional Relationship Between Research Institutions and Entrepreneurship

We begin in Table 2 where we decompose *Number of Ventures* and *RECPI* by the presence or absence of at least one research university or national laboratory. Overall, there is a striking correlation between the presence of a research institution and entrepreneurial activity. The number of ventures more than doubles in the presence of a national laboratory and more than triples for those zip codes within a five-mile radius of a research university. Even more strikingly, RECPI shows a 350% increase in locations near national labs and more than a 700% boost in locations near research universities. Each of these conditional means is statistically significantly different than for locations that are not proximate to at least one research institution.

Insert Table 2 about here

Of course, the geography of universities or national laboratories is not random. Even when performing a cross-sectional analysis, it is important to account for locational differences that may also be correlated with the level and quality of entrepreneurship. More subtly, many national laboratories are themselves co-located with research universities, and so the presence of a national laboratory may in fact reflect the simultaneous presence of a research university rather than an independent correlation between national laboratories and entrepreneurship.

Table 3 explores this possibility by examining the relationship between entrepreneurship and the presence of research universities and national laboratories at the same time, both excluding and including a rich set of locational characteristics (population, density, income, etc), and also a set of year dummies. Three relationships stand out. First, regardless of the inclusion or exclusion of other measures, there is a quantitatively large and statistically significant relationship between both *Number of Ventures* and *RECPI* and the presence of a nearby research university. For example, in (3-4), where we control for the presence of national laboratories, as well as locational characteristics, the presence of a research university is associated with an increase of 351 in RECPI, corresponding to 40% of the standard deviation of that measure. Second, and in sharp contrast, the positive pairwise correlation between national laboratories and entrepreneurship is nullified out after controlling for the simultaneous presence of a research university, and there is actually a negative relationship between the presence of a national laboratory and the quantity of entrepreneurship after the inclusion of locational characteristics. Finally, it is useful to note that the inclusion of locational controls have a significant impact on the relationship between research institutions and entrepreneurship; for example, the relationship between the existence of a research university and *RECPI* is nearly halved after the inclusion of these controls. This is consistent with the idea that the areas where universities are located are indeed different from the areas without a research institution nearby, and they present features that are correlated with increased entrepreneurial activity (Florida, 2005).

Insert Table 3 about here

We emphasize that these relationships are not causal, but are simply intended to highlight the cross-sectional relationship between research institutions and both the quantity and quality-adjusted quantity of entrepreneurship. Figure 1 extends this logic by examining how the correlation between research institutions and *RECPI* has changed over time. To do so, we plot the regression coefficients of the interaction effects between Research University and *National Laboratory*, respectively, with a full set of year dummies (controlling for locational characteristics). There are two sharp insights from this figure. First, while the relationship between the presence of a research university and *RECPI* in the late 1980s was limited, there was a steady increase in the premium associated with university colocation through the 1990s, which has persisted through at least 2012. Put another way, the incidence for the colocation of universities and locational clusters of innovation and entrepreneurship, highlighted first in areas such as biotechnology (Zucker, Darby and Armstrong, 1998; Feldman, 2000), seems to have increased over time. The linkage between research universities and growth-oriented entrepreneurship is not an historical constant but has grown over the past thirty years. Second, and in sharp contrast, the overall relationship between *RECPI* and national laboratories is very noisy; there is no discernible robust relationship between these dedicated research facilities and entrepreneurial quality.[13]

Insert Figure 1 about here

B. The Impact of Changes in Federal Resources Towards Research Institutions on Local Entrepreneurial Ecosystems

We now turn to the centerpiece of our analysis where we move beyond cross-sectional correlations and turn to how changes in public funding to research institutions impact the subsequent quantity and quality of entrepreneurship. As highlighted in our descriptive findings, the key challenge of identifying the impact of research institutions on local entrepreneurship is that the location of universities and national laboratories is not random; research universities and national laboratories are associated with other locational characteristics (including the presence of vibrant entrepreneurial clusters) that themselves encourage entrepreneurship. Following the logic laid out in Section IV, we focus on leveraging variation in the level and nature of activities in research institutions over time that is independent of those locational characteristics. To do so, we first group each of the annual measures into two-year or three-year aggregates (depending on the specification), and then focus on the impact of changes in Federal resources to research institutions on first the quantity and then quality-adjusted quantity of entrepreneurship.

---

[13] We explore the interplay between university, locational characteristics and entrepreneurship in Appendix Table B.3, including the relationship to "elite" universities (Top 10 or Top 50) and professional schools. Interestingly, there is a significant relationship between *RECPI* and Top 10 universities (reflecting the importance of entrepreneurial clusters surrounding institutions such as Stanford and MIT), and the presence of a law school. As well, among other results, high-income zip codes near universities are associated with an increase in the *Number of Ventures.*

A key element of our empirical approach is leveraging variation over time in Federal funding to a given research institution to identify the impact of resource allocation on subsequent entrepreneurship. We therefore begin our analysis in Table 4 by considering the degree to which there is idiosyncratic variation within research institutions over time. For each of our three main measures of Federal resources (*RU Federal R&D Funds*, *RU Federal Non-R&D Funds*, and *Lab Federal R&D Funds*), we decompose the variance of each measure by considering the overall standard deviation, and the standard deviation of the residual after subtracting out institution-level fixed effects, year-level fixed effects, and institution-level and year-level fixed effects. Overall, while there are large and persistent differences across institutions (e.g., more than 40% of the overall standard deviation in *RU Federal R&D Funds* is accounted for by university-level fixed effects), and there is a temporal trend over time (e.g., *RU Federal R&D Funds* exhibits about a 5% per year average increase over the sample), there is a large level of idiosyncratic variation within research institutions over time. For example, the annual average of *RU Federal R&D Funds* at the university level is $76.7 million with a standard deviation of that measure of $111.6 million. However, when we subtract institution-level and year-level fixed effects, the average standard deviation of this "residual" variation is still $38.5 million. Relative to their respective means, there is a similar degree of idiosyncratic variation for both *RU Federal Non-R&D Funds* and *Lab Federal R&D Funds*. For example, relative to a mean of $215.5 and an unconditional standard deviation of $425.5, the "residual" standard deviation for *Lab Federal R&D Funds* is $192.5. Taken together, Table 4 suggests that the level of Federal resources allocated to individual research institutions has a meaningful idiosyncratic component – while a significant level of variation in resources is accounted for by institution-level and overall temporal effects, there is also significant variation in the ability of research institutions (through the grant writing of individual researchers, or the initiatives of administrators) to obtain Federal support. It is this source of variation which we will exploit to then examine the impact of shifts in the level of activity by research institutions on the subsequent quantity and quality of entrepreneurship.

Insert Table 4 about here

Our analysis begins in Table 5 with a focus on the overall quantity of entrepreneurship. We begin in (5-1) with a panel fixed effects specification, where we examine the impact of funding on the *Number of Ventures* controlling for year and ZTCA fixed effects (with standard errors clustered at the zip code level). There is a large and statistically significant impact of both *RU Federal R&D Funds* and *RU Federal Non-R&D Funds* with a much lower impact for *Lab Federal R&D Funds*.

Indeed, the quantitative estimate in (5-1) seems implausibly large: the estimates suggests that a \$1,000 increase in Federal funds to a university (either research-oriented or not) is associated with ~0.3 new ventures in the subsequent time period in each zip code co-located with that university. Of course this specification abstracts away from the potential dynamics within locations, where there might be contemporaneous increases in the level of the entrepreneurial ecosystem and the ability of universities within that ecosystem to attract Federal resources. As such, we then include the lagged level of entrepreneurship (*Number of Ventures*$_{s-1}$) in (5-2) (once again with ZTCA and year fixed effects). The coefficients on resources towards universities remain positive and statistically significant but are much smaller (though still sizeable) in magnitude. For R&D funds to universities, an increase of \$70,000 is associated with a subsequent increase of one business registrant in each zip code co-located to that university (recall that the median number of zip codes adjacent to a university is eight), and the estimated impact for non-R&D funds to universities is estimated to be nearly three times as large. This evidence is consistent with the idea that expenditures by institutions serve as a source of local demand (Andrews, 2020), and so an increase in university expenditures might encourage a higher quantity of entrepreneurship (without having an impact on quality).

Insert Table 5 about here

Given the inclusion of ZTCA fixed effects and the inclusion of *Number of Ventures*$_{s-1}$, a causal interpretation of (5-2) depends on assuming that changes in the levels of Federal resources over time within a location are independent of each other. If there are trends over time in the Federal resources (and the impact of those resources extend over time, then the (absolute value) of the coefficients on the Federal resource measures in (5-2) will be positively biased, leading to an overestimate of the impact of Federal resources on the quantity of entrepreneurship.

We therefore turn in the remaining columns of Table 4 to a first-differences specification, including both an OLS and instrumental variables implementation (Anderson and Hsiao, 1981). We exclude ZTCA fixed effects, but include county-level fixed effects to account for differences in the trend in entrepreneurship across regions, and state-year fixed effects to account for changing regional economic conditions (all of our broad findings are robust to inclusion or exclusion of state-year fixed effects). The results in (5-3) through (5-6) provide support for the hypothesis that resources directed at local research institutions encourage a higher level of local entrepreneurship. While the coefficient on *RU Federal Non-R&D Funds* is somewhat reduced (but remain quantitatively and statistically significant for two-year first-differences), the coefficients on *Federal*

R&D Funds and *Lab Federal R&D Funds* actually increase relative to (5-2). Moreover, these results are robust to the use of the Anderson-Hsiao first-differences instrumental variables approach (as discussed in Section IV), suggesting that the results are not simply driven by autocorrelation within zip codes over time. As well, the coefficient on Δ*Number of Ventures*$_{s-1}$ is positive, and is of similar magnitude in the case of both the OLS and Anderson-Hsiao IV estimator. According to these estimates, a $100,000 "shock" to university resources in a two-year period is associated with between 4 and 8 additional business registrants in adjacent zip codes in the subsequent two-year period (accounting for the overall trend in entrepreneurship in that county and the overall shift in entrepreneurship in that state-year). It is also useful to note that the impact of resources seems to be relatively rapid: the magnitudes of the effect are somewhat smaller for the case where we group the data into three-year time periods (and the coefficient on *RU Federal Non-R&D Funds* becomes insignificant).

While Table 5 focuses on the impact of resources on the quantity of entrepreneurship, Table 6 follows the same empirical logic as Table 5 but focuses attention on potentially a more critical measure, RECPI, which directly accounts for the growth potential of start-ups at the time of their founding. The first two specifications (6-1 and 6-2) present the fixed effects panel estimator (excluding or including a lagged dependent variable), the second two specifications (6-3 and 6-4) examine a first-differences specification (implemented with OLS and an Anderson-Hsiao instrumental variables estimator), and (6-5) and (6-6) present the first-differences specification with the data grouped into three-year rather than two-year intervals.

Insert Table 6 about here

The results are striking and robust. First, across both the fixed effects and first-differences specifications (both OLS and IV), there is a positive and quantitatively significant impact of *RU Federal R&D Funds* on the subsequent change in *RECPI*. Moreover, after controlling for *RECPI*$_{s-1}$, the coefficient on *RU Federal R&D Funds* varies from 0.81 to 1.26. From a quantitative perspective (and taking 0.81 as the baseline estimate), this implies that an increase of $1 million of Federal research support (approximately the size of 2 additional average-sized NIH grants) is associated with a shift of 810 in *RECPI* for each zip code proximate to that university. Recalling that *RECPI* is defined as the quantity of entrepreneurship, weighted by the growth probability for each firm multiplied by 10,000, this can be interpreted as increasing the potential of the birth of a single growth firm by 0.08 within that zip code over that two-year time period. However, each university is proximate on average to eight zip codes. As such, the value of RECPI increases by

10,000 (or the equivalent of one *successful* growth outcome from the cohort of start-ups founded in a given university ecosystem) for each $1.56 million of incremental Federal research support. To put this in context, consider Weber State University in Ogden, Utah, where a 249% increase in research funding between 1997 and 1998 was accompanied by a 40% subsequent increase in RECPI (including a 17% increase in # of Ventures) in zip codes less than five miles from campus (see Appendix D). Interestingly, this burst in entrepreneurial activity was not experienced in more distant zip codes in the state (e.g., zip codes between 5 to 50 miles (e.g., Salt Lake City).

Put another way, an increase by one average standard deviation of the idiosyncratic component of *RU Federal R&D Funds* (~ $10 million) is associated with the founding of ~ 6 additional firms that will experience significant growth over the subsequent six years. It is of course important to interpret this finding carefully: our estimates identify the impact of the ideation and attraction of Federal research support by a university researcher and not simply the random allocation of dollars to a random institution. In other words, while we assume that variation over time among university researchers in the arrival rate of new ideas that merit Federal funding is conditionally independent of the quality of their local entrepreneurial ecosystem, we cannot interpret these findings as a measure of the value of a randomly allocated dollar of public research funding. At the same time, the magnitude of these estimates are large given their follow-on economic impact. Though highly skewed, the private returns to successful growth-oriented entrepreneurship (i.e., the value of a successful exit such as an IPO or acquisition worth more than $10 million) are, on average, high. For example, the median returns to a *single* successful growth outcome (either an IPO or successful acquisition > $10M within six years of founding) are estimated to be more than $50 million of private value to the entrepreneur and investors (Andrews *et al.*, 2020), exclusive of additional local economic impact in terms of employment growth or subsequent regional knowledge and supply chain spillovers.[14] While a complete analysis of the economic value to that ecosystem would require not only observing the return to successes but also the investments in firms that do not succeed (which will dramatically reduce the overall return of the marginal Federal research funding), our estimates nonetheless highlight the potential value of local research institutions in playing a meaningful role in establishing and accelerating local entrepreneurial ecosystems.

In contrast to the positive impact of Federal funds for research, the estimated impact of *RU Federal Non-R&D Funds* is negative across all specifications. Moreover, the absolute value of these coefficients are from 2 – 6 times as large (ranging from -2.19 in (6-2) to -7.34 in (6-3)). Though a large estimated impact, it is useful to note that both the mean and standard deviation of

---

[14] See Figure D for calculations.

this measure is about 20% as large as *RU Federal R&D Funds*. Given the results from Table 5 suggest that the impact of non-R&D funding increases the *quantity* of entrepreneurship, these results are consistent with the idea that increasing local demand enhances entrepreneurship, but only research-oriented expenditures also increase the quality-adjusted quantity of entrepreneurship. Compared to research funding, non-research funding may simply result in substitution for the marginal entrepreneur towards more local business formation. Finally, there is a negative and much smaller in magnitude estimated coefficient across all specifications for *Lab Federal R&D Funds.* Despite being focused on research expenditures, this small and negative finding suggests that the cloistered nature of national laboratories (less open to local knowledge spillovers, less involved in teaching) results in a much more limited impact of these institutions on their local entrepreneurial ecosystems.[15]

Insert Table 7 about here

In Table 7 we examine the interplay between federal research funding to universities, and characteristics of the university and location. The interactions terms are added to the baseline first-differences models corresponding to (5-4) and (6-4), and include state-year and county fixed effects. When we interact *RU Federal R&D Funds* with *Top 50 University*, there is a negative coefficient on this interaction effect both for Δ*Number of Ventures* and *RECPI*. We find a similar negative effect for an interaction effect with *# of Research Universities* (and no consistent impact for universities located in more dense urban areas). These results are consistent with the idea that the marginal returns to Federal research funding is larger in environments with a less robust research and innovation infrastructure: environments with more resources (either through the underlying strength of a single university or the presence of multiple universities) may benefit less from a marginal addition of research resources into that local entrepreneurial ecosystem.

Insert Table 8 about here

One of the most distinctive results is the relative impact of changes in R&D resources for universities versus national laboratories on subsequent entrepreneurship. As discussed earlier, one of the primary mechanisms that universities might have an advantage in encouraging start-up formation within their local geography is through the concurrent production of students. We

---

[15] We have performed several robustness tests examining different time windows and excluding Stanford and MIT from the analysis. Results are presented in Appendix C1 and C2.

investigate this possibility in an exploratory way in Table 8, where we examine the impact not only of changes in funding allocation, but also changes in graduation rates by different types of students. Specifically, we consider, for both *# of Ventures* and *RECPI* , the impact of changing numbers of graduates of proximate universities, separated out by the change in students in natural sciences versus social sciences/humanities ((8-1) and (8-3)) and disaggregated further by graduation level (bachelors versus masters versus PhD). Overall, there is a large and significant effect of graduates in tghe natural sciences, particularly for the production of bachelor's degrees in terms of impact on *RECPI*; in contrast, changes over time in the number of social sciences and humanities degrees are associated with much smaller (and noisier) negative impacts on subsequent medium-term entrepreneurship (over the next two-year window).

Finally, it is useful to note that we additionally explored the impact of other types of funding to universities, including state research and non-research funding, and also funds from private sector businesses. Relative to the allocation of Federal R&D resources (which are predominantly the result of competitive grant processes through NSF, NIH and other Federal funding agencies), these other sources of funding are highly endogenous, as they are often explicitly premised on promoting local economic development and local entrepreneurship. For example, companies such as Monsanto invest heavily in research and innovation centers at universities close to their main corporate research facilities with the explicit aim of encouraging knowledge transfer and potential collaborative spin-out behavior from the university (Sohn, 2020). With that said, it is useful to note that the estimated coefficient on *Business Funds to Universities* is large and positive, and also results in an insignificant estimate for coefficients of interest such as *RU Federal R&D Funds*.

Overall, these results provide evidence about the relative salience of alternative channels of impact of research institutions on entrepreneurship. Specifically, the framework in Section II highlights that one channel by which research institutions might impact entrepreneurship is through their role as a source of local demand (Andrews, 2020). And, that would have the consequences that increased economic activity by research institutions might encourage an increase in the quantity of entrepreneurship. However, a second channel by which research institutions might impact entrepreneurship is through the generation of knowledge, and that channel is likely to be salient for research institutions where it is possible for potential entrepreneurs to carry that knowledge outside the boundary of the institution (or, similarly, if the boundaries of the institution are porous and open enough to support local knowledge spillovers). The broad pattern of findings in Tables 4-7 are consistent with the idea that research funding to universities plays a special role in shaping local entrepreneurship by inducing the creation of ventures with high potential for growth (and local economic impact), while non-research funding may actually increase the quantity of entrepreneurship but divert potential entrepreneurs towards the founding of more locally focused ventures.

VII. Conclusions

This paper provides specific evidence about the role of university research in shaping place-based entrepreneurial ecosystems. Our analysis considers the impact of both universities and national laboratories, as well as both research and non-research funding. Our approach allows us to identify the impact of the generation of an incremental research "idea" or "project" from a university that is of sufficient quality to attract Federal research funding: in other words, the key estimates provide evidence for the impact of the marginal funded idea or project on subsequent local entrepreneurship. Our key finding is that while Federal research funding of all types is linked to a higher quantity of entrepreneurship, only research funding to universities is associated with a meaningful increase in the quality-adjusted quantity of entrepreneurship.

Taken at face value the estimated impact of university research is very large: developing 2-3 successful research grants that attract an incremental $1.6 million of Federal funding induces a sufficient level of quality-adjusted quantity of entrepreneurship to eventually result in one successful local growth outcome. Our approach purposefully abstracts away from the precise mechanism underlying the impact of the knowledge produced by university research on subsequent entrepreneurship. On the one hand, universities can directly influence their entrepreneurial ecosystem through the establishment of spin-offs and technology transfer activities to industrial partners(Mowery *et al.*, 2001; Bercovitz and Feldman, 2008; Hausman, 2020). Second, it is possible that the impact on local entrepreneurial ecosystems arises from the impact of students and other non-permanent research staff involved in the research itself (Saxenian, 1996). In other words, it is possible that by graduating students who have been exposed to the most recent research activities of the university, the near-term spillovers to the local entrepreneurial ecosystem arises from the knowledge being "wrapped up in a person" (Moretti, 2004; Stephan, 2006; Abramovsky, Harrison and Simpson, 2007). Recent work by Babina et al. (2020) provide evidence for the role of individual in shaping the near-term commercialization impact of Federally funded research. Finally, it is possible that technology and knowledge transfer can also happen through more general spillovers to the local environment, perhaps mediated by the social networks of inventors and entrepreneurs. While published research is in principle available globally, there is significant evidence that location mediates the impact of spillovers from university research (Jaffe, Trajtenberg and Henderson, 1993; Adams, 2002). An important agenda for future research is integrating the assessment of the overall impact of universities and related research institutions on their local economic environment with the study of specific mechanisms that directly influence the process of knowledge transfer and local commercialization.

Beyond the precise mechanism underlying our findings, the overall estimates provide support for the potential for enhancements to Federal funding of university research to accelerate local entrepreneurial ecosystems. In an influential policy analysis, Gruber and Johnson (2019) argue for the establishment of a critical set of regional innovation hubs with the objective of "jump-starting" regional entrepreneurial ecosystems through public support of regional universities and research institutions (outside of established "superstar" locations such as the Bay Area and Boston). Our analysis provides insights into both the potential efficacy and necessary conditions that would be required for such a proposal to be successful. First, this paper provides the first direct empirical support for the hypothesis that Federal funding of university research (irrespective of the type of research) directly induces a higher level of growth-oriented local entrepreneurship. However, we also find that the impact of Federal research is limited to university research funding (and is not present for research funding to national laboratories or to non-research-related funding). As such, the evidence also suggests that the impact of Federal funding is realized in an environment characterized by a high degree of openness to the local environment that also allows for mobility (e.g., by students) from the university to the local ecosystem. Realizing the local economic benefits of a policy that enhances Federal research spending in certain geographic regions is likely to depend on proactive investment in shaping local entrepreneurial ecosystems themselves.

References

Abramovsky, L., Harrison, R. and Simpson, H. (2007) ‘University research and the location of business R&D’, *Economic Journal*, 117(519), pp. 114–141.

Acs, Z. J., Audretsch, D. B. and Feldman, M. P. (1992) ‘Real Effects of Academic Research: Comment’, *The American Economic Review*, 82(1), pp. 363–367.

Adams, J. D. (2002) ‘Comparative localization of academic and industrial spillovers’, *Journal of Economic Geography*, 2(3), pp. 253–278.

Adams, J. D., Chiang, E. P. and Jensen, J. L. (2003) ‘The influence of federal laboratory R&D on industrial research’, *Review of Economics and Statistics*, 85(4), pp. 1003–1020.

Anderson, T. W. and Hsiao, C. (1981) ‘Estimation of dynamic models with error components’, *Journal of the American statistical Association*, 76(375), pp. 598–606.

Andrews, M. (2020) *How do institutions of higher education affect local invention? Evidence from the establishment of US colleges*.

Andrews, R. J. *et al.* (2020) *The Startup Cartography Project*. Working Paper.

Angrist, J. D. and Pischke, J.-S. (2008) *Mostly harmless econometrics: An empiricist’s companion*. Princeton university press.

Arellano, M. and Bond, S. (1991) ‘Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations’, *The review of economic studies*, 58(2), pp. 277–297.

Arrow, K. J. (1972) ‘Economic welfare and the allocation of resources for invention’, in *Readings in industrial economics*. Springer, pp. 219–236.

Audretsch, D. B., Lehmann, E. E. and Warning, S. (2005) ‘University spillovers and new firm location’, *Research Policy*, 34(7), pp. 1113–1122.

Audretsch, D. B. and Stephan, P. E. (1996) ‘Company-scientist locational links: The case of biotechnology’, *The American economic review*, 86(3), pp. 641–652.

Audretsch, D. B. and Thurik, A. R. (2001) ‘What’s New about the New Economy? Sources of Growth in the Managed and Entrepreneurial Economies’, *Industrial and Corporate Change*, 10(1), pp. 267–315.

Babina, T. *et al.* (2020) *The Color of Money: Federal vs. Industry Funding of University Research*. Cambridge, MA.

Baptista, R., Lima, F. and Mendonça, J. (2011) ‘Establishment of higher education institutions and new firm entry’, *Research Policy*, 40(5), pp. 751–760.

Baptista, R. and Mendonça, J. (2010) ‘Proximity to knowledge sources and the location of knowledge-based start-ups’, *Annals of Regional Science*, 45(1), pp. 5–29.

Belenzon, S., Chatterji, A. K. and Daley, B. (2017) ‘Eponymous entrepreneurs’, *American Economic Review*, 107(6), pp. 1638–1655.

Belenzon, S., Chatterji, A. K. and Daley, B. (2020) ‘Choosing between growth and glory’, *Management Science*, 66(5), pp. 2050–2074.

Bercovitz, J. and Feldman, M. (2008) ‘Academic entrepreneurs: Organizational change at the individual level’, *Organization Science*, 19(1), pp. 69–89.

Dasgupta, P. and David, P. A. (1994) ‘Toward a new economics of science’, *Research policy*, 23(5), pp. 487–521.

Delgado, M., Porter, M. E. and Stern, S. (2016) ‘Defining clusters of related industries’, *Journal of Economic Geography*, 16(1), pp. 1–38.

Feldman, M. P. (1994) *The geography of innovation*. Springer Science & Business Media.

Feldman, M. P. (2000) ‘Location and innovation: the new economic geography of innovation, spillovers, and agglomeration’, in *The Oxford handbook of economic geography*.

Feldman, M. P. (2014) ‘The character of innovative places: entrepreneurial strategy, economic development, and prosperity’, *Small Business Economics*, 43(1), pp. 9–20.

Florida, R. (2002) ‘Bohemia and economic geography’, *Journal of economic geography*, 2(1), pp. 55–71.

Florida, R. L. (2005) *Cities and the creative class*. Psychology Press.

Furman, J. L. and MacGarvie, M. J. (2007) ‘Academic science and the birth of industrial research laboratories in the US pharmaceutical industry’, *Journal of Economic Behavior & Organization*, 63(4), pp. 756–776.

Furman, J. L. and Stern, S. (2011) ‘Climbing atop the shoulders of giants: The impact of institutions on cumulative research’, *American Economic Review*, 101(5), pp. 1933–1963.

Gruber, J. and Johnson, S. (2019) *Jump-starting America: How breakthrough science can revive economic growth and the American dream*. Hachette UK.

Guzman, J. and Stern, S. (2015) ‘Innovation economics. Where is Silicon Valley?’, *Science*, 347(6222), pp. 606–9.

Guzman, J. and Stern, S. (2017) ‘Nowcasting and Placecasting Entrepreneurial Quality and Performance’, in Haltiwanger, J. et al. (eds) *Measuring Entrepreneurial Businesses: Current Knowledge and Challenges*.

Hall, B. H., Link, A. N. and Scott, J. T. (2003) ‘Universities as Research Partners’, *Review of Economics and Statistics*, 85(2), pp. 485–491.

Hall, R. E. and Woodward, S. E. (2007) *The incentives to start new companies: Evidence from venture capital*. National Bureau of Economic Research.

Hausman, N. (2020) ‘University Innovation, Local Economic Growth, and Entrepreneurship’, *Review of Economics and Statistics*.

Hurst, E. and Pugsley, B. W. (2011) *What do small businesses do?* National Bureau of Economic Research.

Jaffe, A. B. (1989) ‘Real effects of academic research’, *The American Economic Review*, 79(5), pp. 957–970.

Jaffe, A. B. and Lerner, J. (2001) ‘Reinventing public R&D: Patent policy and the commercialization of national laboratory technologies’, *RAND Journal of Economics*. JSTOR, pp. 167–198.

Jaffe, A. B., Trajtenberg, M. and Henderson, R. (1993) ‘Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations’, *The Quarterly Journal of Economics*. Oxford University Press, 108(3), pp. 577–598.

Kantor, S. and Whalley, A. (2014) ‘Knowledge Spillovers from Research Universities: Evidence from Endowment Value Shocks’, *Review of Economics and Statistics*, 96(1), pp. 171–188.

Kenney, M. (1986) *Biotechnology: The university-industrial complex*. Yale University Press.

Kim, Y., Kim, W. and Yang, T. (2012) ‘The effect of the triple helix system and habitat on regional entrepreneurship: Empirical evidence from the U.S.’, *Research Policy*, 41(1), pp. 154–166.

Lucas, R. E. (1988) ‘On the mechanics of economic development’, *Journal of Monetary Economics*, 22(1), pp. 3–42.

Moretti, E. (2004) ‘Estimating the social return to higher education: evidence from longitudinal and repeated cross-sectional data’, *Journal of Econometrics*, 121(1–2), pp. 175–212.

Moretti, E. (2012) *The new geography of jobs*. Houghton Mifflin Harcourt.

Mowery, D. C. *et al.* (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(1), pp. 99–119.

Mowery, D. C. and Rosenberg, N. (1993) ‘The US national innovation system’, *National innovation systems: A comparative analysis*. Oxford University Press New York, pp. 29–75.

Nelson, R. R. (1959) ‘The simple economics of basic scientific research’, *Journal of political econom*, 67(3), pp. 297–306.

Nickell, S. (1981) ‘Biases in dynamic models with fixed effects’, *Econometrica: Journal of the econometric society*, pp. 1417–1426.

Rosenberg, N. and Nelson, R. R. (1994) ‘American universities and technical advance in industry’, *Research Policy*, 23(3), pp. 323–348.

Saxenian, A. (1996) *Regional Advantage: Culture and Competition in Silicon Valley and Route 128* . American Council of Learned Societies.

Schoar, A. (2010) ‘The divide between subsistence and transformational entrepreneurship’, *Innovation policy and the economy*, 10(1), pp. 57–81.

Senate of the United States (2021) *S.1260 - 117th Congress (2021-2022): Endless Frontier Act.*

Shulman, L. S. (2001) ‘The Carnegie classification of institutions of higher education’, *Menlo Park: Carnegie Publication.*

Sohn, E. (2020) ‘How Local Industry R&D Shapes Academic Research: Evidence from the Agricultural Biotechnology Revolution’, *Organization Science*

Stephan, P. (2006) ‘Wrapping it up in a person: The mobility patterns of new PhDs’, *Innovation policy and the economy*. MIT Press, 7, pp. 71–98.

Stephan, P. E. (2012) *How economics shapes science*. Harvard University Press Cambridge, MA.

Stern, S. (1995) ‘Incentives and focus in university and industrial research: the case of synthetic insulin’, in Rosenberg, N., Gelijns, A., and Dawkins, H. (eds) *Sources of Medical Technology: Universities and Industry*. National Academy of Sciences.

Valero, A. and Van Reenen, J. (2019) ‘The economic impact of universities: Evidence from across the globe’, *Economics of Education Review*, 68, pp. 53–67.

Zucker, L. G., Darby, M. R. and Armstrong, J. (1998) ‘Geographically localized knowledge: spillovers or markets?’, *Economic Inquiry,* 36(1), pp. 65–86.

TABLE 1

Variable Definition and Summary Statistics

	Definition	Source	Mean	Std. Dev.
Outcome Variables				
# of Ventures	Number of newly-formed ventures in each ZCTA, by year (1989-2012)	Guzman and Stern (2016)	41.78	101.20
RECPI	[10,000 * average quality * # of Ventures], within a ZCTA and by year (1989-2012)	Guzman and Stern (2016)	262.50	1,074.00
Public Research Measures				
National Lab	1 if at least one National Research Laboratory is in a 5-miles radius of the ZCTA	National Science Foundation	0.006	0.079
Research University	1 if at least one university classified as “Doctoral/Research Universities” or “Research Universities – (very) high research activity" is in a 5-miles radius of the ZCTA	National Center for Education Statistics, Carnegie Classification	0.062	0.241
# of Research Universities	Number of Research Universities in a 5-miles radius of the ZCTA	Own elaboration	0.098	0.471
Top50	1 if at least one of the universities in a 5 miles radius of the ZCTA is in the top 50 of the *Times Higher Education 2017 ranking* of US universities	Times Higher Education	0.018	0.133
Top10	1 if at least one of the universities in a 5 miles radius of the ZCTA is in the top 10 of the Times Higher Education 2017 ranking of US universities	Times Higher Education	0.006	0.080
Business School	1 if there is at least one business school in a research university in a 5 miles radius of the ZCTA	Wikipedia	0.048	0.214
Medical School	1 if there is at least one medical school in a research university in a 5 miles radius of the ZCTA	Wikipedia	0.036	0.187
Law school	1 if there is at least one law school in a research university in a 5 miles radius of the ZCTA	Wikipedia	0.040	0.197
R&D funds to universities	Total amount of federal funds ($k) earmarked for R&D activities received by the research universities in a 5 miles radius of the ZCTA, per year	National Science Foundation	10,490	68,942
Non-R&D funds to universities	Total amount of federal funds ($k) earmarked for non R&D activities received by the research universities in a 5 miles radius of the ZCTA, per year	National Science Foundation	1,247	7,912
R&D funds to national laboratories	Total amount of federal funds ($k) received by the national laboratories in a 5 miles radius of the ZCTA, per year	National Science Foundation	3,319	68,956
	Total number of ZCTA		16,320	
	Total number of observations (panel)		408,000	

Note: All dollar amounts are inflation-adjusted and standardized to 2012 dollars.

TABLE 2

Quantity and Quality-Adjusted Quantity of Entrepreneurship: With and Without a Research Institution

Research University = 0 (15,307 ZCTAs)			**Research University = 1 (1,013 ZCTAs)**			
Variable	*Mean*	*Std. Dev.*	*Variable*	*Mean*	*Std. Dev.*	*T-test of means*
# of Ventures	35.6	92.0	# of Ventures	134.6	167.5	***
RECPI	200.0	751.7	RECPI	1208	3015.1	***
National Laboratory = 0 (16,220 ZCTAs)			**National Laboratory = 1 (100 ZCTAs)**			
Variable	*Mean*	*Std. Dev.*	*Variable*	*Mean*	*Std. Dev.*	*T-test of means*
# of Ventures	41.5	101.2	# of Ventures	89.6	89.7	***
RECPI	258.4	1069.9	RECPI	929.6	1452.1	***

* p<0.05 ** p<0.01 *** p<0.001

TABLE 3

Number of Ventures and RECPI By Presence of Research Institutions: Year and Zipcode Controls

	# of Ventures		*RECPI*	
	(1)	(2)	(3)	(4)
Research University	99.67***	22.20***	1,003.23***	399.98***
	(4.74)	(4.84)	(81.14)	(74.57)
National Laboratory	-10.87	-39.05***	77.70	-235.72*
	(9.15)	(6.77)	(123.38)	(112.88)
Constant	12.54***	-87.88***	79.28***	-585.07***
	(0.38)	(10.47)	(4.41)	(142.31)
Observations	408,000	408,000	408,000	408,000
Number of ZCTAs	16,320	16,320	16,320	16,320
Census 1990 Controls		Yes		Yes
R-squared	0.0896	0.362	0.0577	0.236
Years FE	Yes	Yes	Yes	Yes
			w	

All specifications include year fixed effects. Columns 2 and 4 also include controls for the 1990 census variables presented in Appendix B, Table B.1. Robust standard errors in parenthesis, clustered at the ZCTA level. * $p<0.05$ ** $p<0.01$ *** $p<0.001$

TABLE 4

Time Variation of Federal Funding to Universities and National Laboratories: With and Without Institution and Time Fixed Effects

	Definition	Mean	Std. Dev.
University-level measures			
# of ZCTA per university *j*	Number of zip codes within 5 miles of university *j*	8.3	8.2
R&D funds	Average annual Federal Funds ($K) for R&D activities for university *j*	76.7	111.6
De-Meaned R&D Funds	Average annual Federal Funds ($K) for R&D activities for university *j*, subtracting university-level fixed effects		48.1
De-Trended R&D Funds	Average annual Federal Funds ($K) for R&D activities for university *j*, subtracting year-level fixed effects		107.3
"Residual" R&D Funds	Average annual Federal Funds ($K) for R&D activities for university *j*, subtracting year-level fixed effects and institution-level fixed effects		38.5
Non-R&D funds	Average annual Federal Funds ($K) for non-R&D activities for university *j*	8.7	12.8
De-Meaned Non-R&D Funds	Average annual Federal Funds ($K) for non-R&D activities for university *j*, subtracting university-level fixed effects		6.6
De-Trended Non-R&D Funds	Average annual Federal Funds ($K) for non-R&D activities for university *j*, subtracting year-level fixed effects		11.8
"Residual" Non- R&D Funds	Average annual Federal Funds ($K) for non-R&D activities for university *j*, subtracting year-level fixed effects and institution-level fixed effects		6.1
National Laboratories-level Measures			
R&D funds	Average annual Federal Funds ($K) for R&D activities for National Lab *k*	215.5	425.5
De-Meaned R&D Funds	Average annual Federal Funds ($K) for R&D activities for National Lab *k*, subtracting national lab-level fixed effects		219.8
De-Trended R&D Funds	Average annual Federal Funds ($K) for R&D activities for National Lab *k*, subtracting year-level fixed effects		408.5
"Residual" R&D Funds	Average annual Federal Funds ($K) for R&D activities for National Lab *k*, subtracting year-level fixed effects and institution-level fixed effects		192.5

TABLE 5

The Impact of Funding on the Quantity of Entrepreneurship: Fixed Effects, Long First Differences and Anderson-Hsiao Instrumental Variable Specification

	Long Panel with FE [2 years]		*Long First Differences [2 years]*		*Long First Differences [3 years]*	
				With Instrument		With Instrument
	$(\text{\# of Ventures})_t$ (1)	$(\text{\# of Ventures})_t$ (2)	$\Delta(\text{\# of Ventures})_t$ (3)	$\Delta(\text{\# of Ventures})_t$ (4)	$\Delta(\text{\# of Ventures})_t$ (5)	$\Delta(\text{\# of Ventures})_t$ (6)
Covariates in levels						
$(\text{\# of Ventures})_{t-1}$		0.981***				
		(0.003)				
Federal funds						
$(\text{R\&D funds to universities})_{t-1}$	0.348***	0.018***				
	(0.030)	(0.002)				
$(\text{Non R\&D funds to universities})_{t-1}$	0.362***	0.056***				
	(0.090)	(0.017)				
$(\text{R\&D funds to national laboratories})_t$	0.014*	0.001				
	(0.006)	(0.001)				
Covariates in first differences						
$\Delta(\text{\# of Ventures})_{t-1}$			0.646***	0.537***	0.714***	0.574***
`			(0.009)	(0.021)	(0.012)	(0.020)
Federal funds						
$\Delta(\text{R\&D funds to universities})_{t-1}$			0.102***	0.120***	0.049***	0.081***
			(0.011)	(0.013)	(0.008)	(0.010)
$\Delta(\text{Non R\&D funds to universities})_{t-1}$			0.079*	0.087*	0.027	0.043
			(0.040)	(0.041)	(0.043)	(0.046)
Δ(R&D funds to national laboratories			0.016***	0.017***	0.011***	0.013***
			(0.002)	(0.003)	(0.002)	(0.002)
Constant	15,100***	-253.88**	9,173***	9,520***	130.71	19,722***
	(521.51)	(97.83)	(1,049.34)	(1,122.13)	(2,136.67)	(2,008.57)
Observations	375,360	359,040	103,230	103,230	57,350	57,350
R-squared	0.169	0.935	0.594	0.589	0.636	0.629
Number of ZCTAs	16,320	16,320	11,470	11,470	11,470	11,470
Year FE	YES	YES				
ZCTA FE	YES	YES				
Counties FE			YES	YES	YES	YES
State*Year FE			YES	YES	YES	YES

Note: In column (1), both endogenous variable and covariates are expressed as 2-year averages of the original variables: $\overline{X}_t = \frac{x_t+x_{t-1}}{2}$. In columns (2), (3), we reduce the 25-year panel into 12 periods and difference out the resulting panel: $\Delta\overline{X}_t = \overline{X}_t - \overline{X}_{t-4}$. In column (4), we re-transform all our original variables as 3-year averages ($\overline{\overline{X}}_t = \frac{x_t+x_{t-1}+x_{t-2}}{3}$), reduce the 25-year panel into 8 periods and difference out: $\Delta\overline{\overline{X}}_t = \overline{\overline{X}}_t - \overline{\overline{X}}_{t-6}$. Notation simplified and Number of Ventures scaled by 1000 for exposition. Robust standard errors in parenthesis. * p<0.05 ** p<0.01 *** p<0.001

.

TABLE 6

The Impact of Funding on RECPI: Fixed Effects, Long First Differences and Anderson-Hsiao Instrumental Variable Specification

	Long Panel with FE [2 years]		*Long First Differences [2 years]*		*Long First Differences [3 years]*	
				With Instrument		With Instrument
	$(\text{RECPI})_t$	$(\text{RECPI})_t$	$\Delta(\text{RECPI})_t$	$\Delta(\text{RECPI})_t$	$\Delta(\text{RECPI})_t$	$\Delta(\text{RECPI})_t$
	(1)	(2)	(3)	(4)	(5)	(6)
Covariates in levels						
$(\text{RECPI})_{t-1}$		0.862***				
		(0.017)				
Federal funds						
$(R\&D\ funds\ to\ universities)_{t-1}$	6.134***	1.132***				
	(0.734)	(0.137)				
$(\text{Non R\&D funds to universities})_{t-1}$	-5.470**	-2.185**				
	(1.917)	(0.775)				
$(\text{R\&D funds to national laboratories})_{t-1}$	-0.199**	-0.039				
	(0.074)	(0.021)				
Covariates in first differences						
$\Delta(\text{RECPI})_{t-1}$			0.346***	-0.044	0.436***	0.230*
			(0.024)	(0.118)	(0.038)	(0.09)
Federal funds						
$\Delta(\text{R\&D funds to universities})_{t-1}$			0.806*	1.257***	2.034***	2.509***
			(0.336)	(0.363)	(0.461)	(0.47)
$\Delta(\text{Non R\&D funds to universities})_{t-1}$			-7.335***	-7.270***	-13.652***	-13.060***
			(1.921)	(1.806)	(3.199)	(3.11)
$\Delta(\text{R\&D funds to national laboratories})_{t-1}$			-0.233*	-0.312*	-0.297**	-0.322**
			(0.098)	(0.125)	(0.111)	(0.117)
Constant	96,977***	31,599***	28,433**		21,010***	95,049***
	(8,208)	(5,802)	(10,458)		(19,717)	(20,057)
Observations	375,360	359,040	103,230	103,230	57,350	57,350
R-squared	0.057	0.719	0.165	0.041	0.239	0.216
Number of ZCTAs	16,320	16,320	11,470	11,470	11,470	11,470
Year FE	YES	YES				
ZCTA FE	YES	YES				
Counties FE			YES	YES	YES	YES
State*Year FE			YES	YES	YES	YES

Note: In column (1), both endogenous variable and covariates are expressed as 2-year averages of the original variables: $\overline{X}_t = \frac{x_t+x_{t-1}}{2}$. In columns (2), (3), we reduce the 25-year panel into 12 periods and difference out the resulting panel: $\Delta\overline{X}_t = \overline{X}_t - \overline{X}_{t-4}$. In column (4), we re-transform all our original variables as 3-year averages ($\overline{\overline{X}}_t = \frac{x_t+x_{t-1}+x_{t-2}}{3}$), reduce the 25-year panel into 8 periods and difference out: $\Delta\overline{\overline{X}}_t = \overline{\overline{X}}_t - \overline{\overline{X}}_{t-6}$. Notation simplified and RECPI scaled by 1000 for exposition. Robust standard errors in parenthesis.

* $p<0.05$ ** $p<0.01$ *** $p<0.001$.

TABLE 7

The Impact of Funding on RECPI and Number of Ventures: Interactions with Institutional and Geographic Characteristics

	$\Delta(\text{\# of Ventures})_t$			$\Delta(\text{RECPI})_t$		
	(1)	(2)	(3)	(4)	(5)	(6)
$\Delta(\text{\# of Ventures})_{t-1}$	0.535*** (0.022)	0.512*** (0.023)	0.530*** (0.022)			
$\Delta(\text{RECPI})_{t-1}$				-0.082 (0.122)	-0.086 (0.124)	-0.075 (0.121)
Federal Funds						
$\Delta(\text{R\&D funds to universities})_{t-1}$	0.169*** (0.027)	0.075 (0.041)	0.180*** (0.024)	1.987** (0.708)	1.882** (0.686)	1.551* (0.767)
$\Delta(\text{Non R\&D funds to universities})_{t-1}$	0.071 (0.042)	0.091* (0.041)	0.079 (0.040)	-7.821*** (1.851)	-7.222*** (1.802)	-7.368*** (1.813)
$\Delta(\text{R\&D funds to national laboratories})_{t-1}$	0.016*** (0.003)	0.016*** (0.003)	0.016*** (0.003)	-0.350* (0.138)	-0.362** (0.130)	-0.382** (0.129)
Interactions						
$\Delta(\text{R\&D funds univ})_{t-1} * \text{top50 Univ}$	-0.098** (0.030)			-4.728*** (1.123)		
$\Delta(\text{R\&D funds univ})_{t-1} * \text{Urban ZCTA}$		0.028 (0.044)			-1.235 (0.788)	
$\Delta(\text{R\&D funds univ})_{t-1} * \text{\# of Research Uni}$			-0.030*** (0.005)			-0.879*** (0.177)
Observations	103,230	103,230	103,230	103,230	103,230	103,230
R-squared	0.589	0.587	0.588	0.030	0.025	0.031
Number of ZCTAs	11,470	11,470	11,470	11,470	11,470	11,470
State*Year FE	YES	YES	YES	YES	YES	YES
Counties FE	YES	YES	YES	YES	YES	YES

Note: The baseline specification for columns (1)-(3) is in Table 5, col. (4). The baseline specification for columns (4) - (7) is in Table 6, col. (4). Main effects on Top 50 Univ, Urban ZCTA, # of Research Universities, and the Constant are omitted for exposition.
All endogenous and federal funds variables are in a long difference format. Notation simplified and endogenous variables scaled by 1000 for exposition. Robust standard errors in parenthesis.
* p<0.05 ** p<0.01 *** p<0.001
.

TABLE 8

The Impact of Awarded Degrees on RECPI and Number of Ventures: Natural Sciences, Social Sciences and Humanities

	Δ(# of Ventures)$_t$ (1)	Δ(# of Ventures)$_t$ (2)	Δ(RECPI)$_t$ (3)	Δ(RECPI)$_t$ (4)
Δ(# of Ventures)$_{t-1}$	0.537***	0.538***		
	(0.021)	(0.021)		
Δ(RECPI)$_{t-1}$			-0.051	-0.061
			(0.119)	(0.120)
Federal funds				
Δ(R&D funds to universities)$_{t-1}$	0.120***	0.133***	1.072**	1.307**
	(0.013)	(0.015)	(0.394)	(0.445)
Δ(Non R&D funds to universities)$_{t-1}$	0.097*	0.046	-6.582***	-6.268***
	(0.041)	(0.041)	(1.754)	(1.739)
Δ(R&D funds to national laboratories)$_{t-1}$	0.017***	0.015***	-0.328*	-0.276*
	(0.003)	(0.003)	(0.133)	(0.117)
Degrees Awarded				
Δ(Bachelor students natural science)$_{t-1}$		6.262**		318.719***
		(1.970)		(72.103)
Δ(Master students natural science)$_{t-1}$		5.452		-83.530
		(3.067)		(154.357)
Δ(PhD students natural science)$_{t-1}$		-13.911**		-45.943
		(4.973)		(228.721)
Δ(Total students natural science)$_{t-1}$	3.528**		231.991***	
	(1.257)		(57.938)	
Δ(Bachelor students social sciences and humanities)$_{t-1}$		-0.653		-73.245*
		(0.981)		(31.414)
Δ(Master students social sciences and humanities)$_{t-1}$		-1.105		4.776
		(0.788)		(20.504)
Δ(PhD students social sciences and humanities)$_{t-1}$		-4.949		1,059.452**
		(7.211)		(393.664)
Δ(Total students social sciences and humanities)$_{t-1}$	-0.804*		-34.959**	
	(0.396)		(13.259)	
Observations	103,230	103,230	103,230	103,230
R-squared	0.589	0.589	0.038	0.035
Number of ZCTAs	11,470	11,470	11,470	11,470
Counties FE	YES	YES	YES	YES
State*Year FE	YES	YES	YES	YES

Note: The baseline specification is in Table 5 and 6, col. (3). Constant omitted for exposition.
All variables are in a long difference format. Notation simplified and endogenous variables scaled by 1000 for exposition. Robust standard errors in parenthesis.
* $p<0.05$ ** $p<0.01$ *** $p<0.001$

FIGURE 1a

Impact of the Presence of a RESEARCH INSTITUTION on Number of Ventures over time

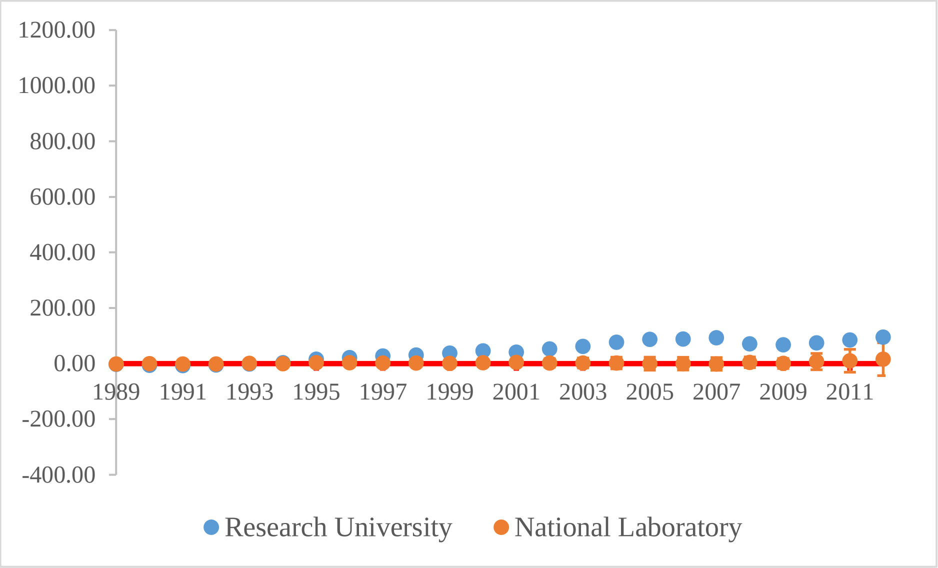
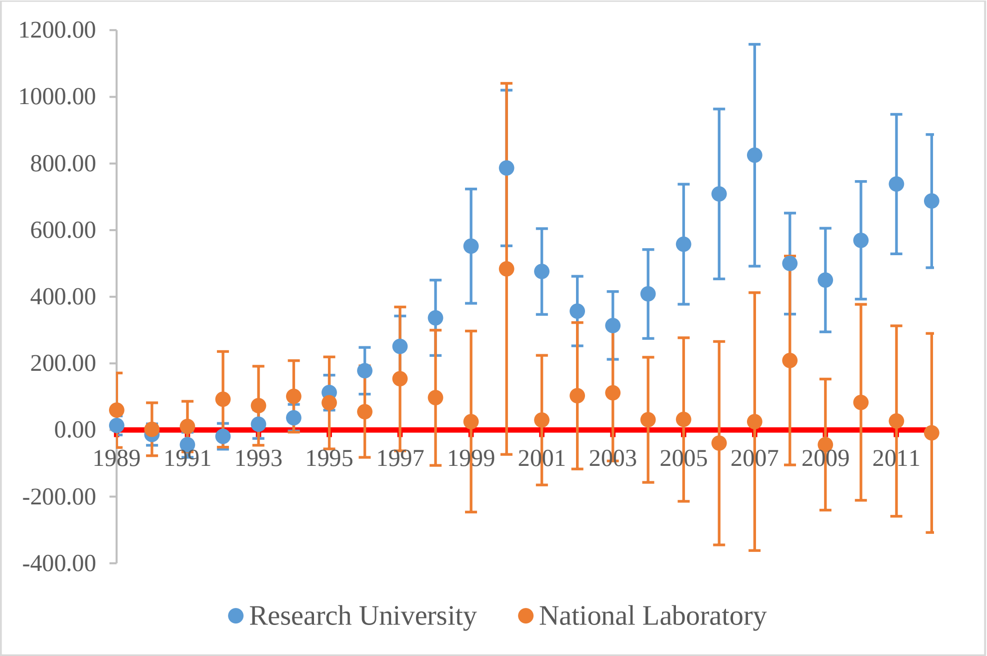


FIGURE 1b

Impact of the Presence of a RESEARCH INSTITUTION on RECPI over time



APPENDIX A
TABLE A.1

Entrepreneurial Quality Logit Model
Predictive Analytics Model of Equity Growth
Dependent Variable: Equity Growth
Logit model. Incidence Rate Ratios Reported

	Full model (1)
Corporate Governance Measures	
Corporation	3.202***
	(0.0650)
Name-Based Measures	
Short Name	1.786***
	(0.0208)
Eponymous	0.312***
	(0.0150)
Intellectual Property Measures	
Trademark	4.288***
	(0.200)
Patent - Delaware Interaction	
Patent Only	20.24***
	(0.847)
Delaware Only	15.26***
	(0.294)
Patent and Delaware	84.08***
	(2.720)
US CMP Clusters	Yes
US CMP High-Tech Clusters	Yes
N	26,051,461
R-squared	0.194

Reproduced from Andrews, et al (2020), Table 2, col (4). This regression is a logit model with Growth as the dependent variable. Growth is a binary indicator equal to 1 if a firm achieves IPO or acquisition within 6 years and 0 otherwise. Growth is only defined for firms born in the cohorts of 1988 to 2010. This model forms the basis of our entrepreneurial quality estimates, which are the predicted values of the model. Incidence ratios reported; Robust standard errors in parenthesis. * $p<0.05$ ** $p<0.01$ *** $p<0.001$

TABLE A.2

Funding and Degrees Controls: Definition and Summary Statistics

	Definition	Source	Mean	Std. Dev.
Funds to universities				
State funds to universities	Total amount of state funds ($k) received by the research universities in a 5 miles radius of the ZCTA, per year	National Science Foundation	981	7,389
Institutional funds to universities	Total amount of funds ($k) received from institutional donors by the research universities in a 5 miles radius of the ZCTA, per year	National Science Foundation	3,136	18,785
Business funds to universities	Total amount of funds ($k) received from business donors by the research universities in a 5 miles radius of the ZCTA, per year	National Science Foundation	1,148	8,175
Other funds to universities	Total amount of funds ($k) received from other sources by the research universities in a 5 miles radius of the ZCTA, per year	National Science Foundation	1,719	11,998
Degrees Awarded				
Bachelor students natural science	Total number of natural sciences bachelor degrees awarded in a 5 miles radius of the ZCTA, per year	National Science Foundation	108.1	446.4
Master students natural science	Total number of natural sciences Master degrees awarded in a 5 miles radius of the ZCTA, per year	National Science Foundation	44.34	241.7
PhD students natural science	Total number of natural sciences PhD degrees awarded in a 5 miles radius of the ZCTA, per year	National Science Foundation	13.6	97.1
Total students natural science	Total number of bachelor, Master and PhD degrees in natural sciences awarded in a 5 miles radius of the ZCTA, per year	National Science Foundation	166.0	753.7
Bachelor students social sciences and humanities	Total number of humanities and social sciences bachelor degrees awarded in a 5 miles radius of the ZCTA, per year	National Science Foundation	302.4	1,219
Master students social sciences and humanities	Total number of humanities and social sciences Master degrees awarded in a 5 miles radius of the ZCTA, per year	National Science Foundation	149.1	838.8
PhD students social sciences and humanities	Total number of humanities and social sciences PhD degrees awarded in a 5 miles radius of the ZCTA, per year	National Science Foundation	14.0	90.9
Total students social sciences and humanities	Total number of bachelor, Master and PhD degrees in humanities and social sciences awarded in a 5 miles radius of the ZCTA, per year	National Science Foundation	465.4	2,067.0
	Total number of ZCTA		16,320	
	Total number of observations (panel)		408,000	

Note: All dollar amounts are inflation-adjusted and standardized to 2012 dollars.

APPENDIX B

TABLE B.1

Census Controls: Summary Statistics

	Definition	Source	Mean	Std Dev
Census Variables by ZCTA (1990)				
Total population	Total population	Decennial Census	9,367	12,893
Density	Total population divided by the geographical area	Decennial Census	0.0004	0.001
Male	% of men in the total population	Decennial Census	0.496	0.038
Urban	% of the population residing in urban areas	Decennial Census	0.349	0.431
Age between 18 and 65	% of the population between 18 and 65 years old	Decennial Census	0.594	0.068
White	% of the population of white ethnicity	Decennial Census	0.884	0.182
Born outside the US	% of the population born outside of the United States	Decennial Census	0.042	0.077
College	% of the population with a college degree or higher	Decennial Census	0.210	0.131
Private sector	% of the population employed in the private sector	Decennial Census	0.720	0.118
Public sector	% of the population employed in the public sector	Decennial Census	0.153	0.083
Income per capita	Median per capita income	Decennial Census	12,506	5,780
Housing values	Median house value	Decennial Census	72,684	67,492
	Total number of ZCTA		16,320	
	Total number of observations (panel)		408,000	

TABLE B.2
of Ventures and RECPI By Presence of Research Institutions: Census Controls

	# of Ventures (1)	*RECPI* (2)
Research University	22.20***	399.98***
	(4.84)	(74.57)
National Laboratory	-39.05***	-235.72*
	(6.77)	(112.88)
Total population	0.00***	0.01***
	(0.00)	(0.00)
Density	-2,728.18*	-25,777.99+
	(1,086.45)	(13,878.21)
Male	-61.10**	351.21
	(21.98)	(250.15)
Urban	20.52***	42.62+
	(2.30)	(24.01)
Age 18 - 65	29.98*	810.16***
	(13.10)	(168.38)
White	16.24***	-148.40***
	(4.83)	(42.27)
Born outside the US	269.78***	2,033.19***
	(33.93)	(244.01)
College	84.88***	146.21
	(11.01)	(177.85)
Private sector	18.61***	-461.43***
	(4.88)	(57.39)
Public sector	-20.01**	-820.14***
	(7.13)	(84.63)
Income per capita	0.00***	0.02*
	(0.00)	(0.01)
Housing values	-0.00***	0.00***
	(0.00)	(0.00)
Constant	-87.88***	-585.07***
	(10.47)	(142.31)
Observations	408,000	408,000
Number of ZCTAs	16,320	16,320

All specifications include year fixed effects. Robust standard errors in parenthesis, clustered at the ZCTA level. * p<0.05 ** p<0.01 *** p<0.001

TABLE B.3

Number of Ventures and RECPI By Institutional and Geographic Characteristics

Columns 1 and 5 examine the effect of the quality of the universities in a 5 miles radius of the focal ZCTA. Columns 2 and 6 examine the effect of the presence of specialized schools in a 5 miles radius of the focal ZCTA. Columns 3 and 7 examine the interaction between the presence of at least one research university and the urbanization/income of the focal ZCTA. Columns 4 and 8 present the full specification.

	# of Ventures				**RECPI**			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Research university	24.40***	11.25	2.04	-7.32	260.19***	84.52	-167.90*	-166.48
	(5.14)	(7.47)	(7.01)	(8.88)	(71.13)	(160.68)	(67.68)	(145.72)
National Laboratory	-35.92***	-42.52***	-40.92***	-38.14***	-765.70***	-275.49*	-258.86*	-760.65***
	(7.38)	(7.02)	(7.37)	(7.82)	(184.71)	(118.28)	(118.11)	(184.13)
Quality of the university								
Top 50	-6.26			-10.01	-45.18			-116.26
	(11.82)			(12.14)	(152.09)			(162.51)
Top 10	-9.16			-22.16	2,229.9***			2,094.4***
	(17.56)			(17.73)	(445.11)			(438.82)
Specialized School								
Business School		-9.85		-7.18		74.20		-84.19
		(8.66)		(8.84)		(138.70)		(139.62)
Medical School		3.31		6.06		64.68		197.95
		(9.57)		(9.86)		(142.49)		(144.38)
Law School		27.19**		30.48**		366.37**		88.91
		(10.43)		(10.56)		(137.23)		(130.07)
Characteristics of the ZCTA								
Research University*Urban ZCTA			10.12	4.34			398.48***	177.84
			(8.57)	(9.01)			(106.04)	(120.50)
Research University*High Income ZCTA			40.53***	45.19***			948.25***	742.77***
			(11.78)	(12.13)			(204.00)	(192.95)
Constant	-87.49***	-89.43***	-92.36***	-92.99***	-645.46***	-612.27***	-615.81***	-664.86***
	(10.47)	(10.43)	(10.57)	(10.48)	(141.48)	(142.25)	(145.89)	(146.18)
Observations	408,000	408,000	408,000	408,000	408,000	408,000	408,000	408,000
R-squared	0.362	0.363	0.367	0.368	0.257	0.239	0.247	0.265
Number of ZCTAs	16,320	16,320	16,320	16,320	16,320	16,320	16,320	16,320
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Census 1990 Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Robust standard errors in parenthesis, clustered at the ZCTA level. * p<0.05 ** p<0.01 *** p<0.001

APPENDIX C

TABLE C1

The Impact of Funding on Number of Ventures: Robustness Checks

	Long Panel with FE [2 years]				*Long First Differences [2 years]*			
	Before 2000	After 2000	97-99 excluded	Stanford & MIT excluded	Before 2000	After 2000	97-99 excluded	Stanford & MIT excluded
	$(\#\text{Ventures})_t$ (1)	(#Ventures) (2)	(# of Ventures) (3)	(#Ventures) (4)	$\Delta(\#\text{Ventures})$ (5)	$\Delta(\#\text{Ventures})$ (6)	$\Delta(\#\text{Ventures})$ (7)	$\Delta(\#\text{Ventures})$ (8)
Covariates in levels								
$(\#\text{ of Ventures})_{t-1}$	0.992***	0.862***	0.979***	0.981***				
	(0.007)	(0.007)	(0.003)	(0.003)				
Federal funds								
$(\text{R\&D funds to univ.})_{t-1}$	0.120***	0.046***	0.023***	0.017***				
	(0.014)	(0.005)	(0.003)	(0.003)				
$(\text{Non R\&D funds to univ.})_{t\text{-}}$	0.037	0.088***	0.066***	0.091***				
	(0.026)	(0.025)	(0.020)	(0.024)				
$(\text{R\&D funds to NL})_{t-1}$	-0.003	0.000	-0.000	0.001				
	(0.008)	(0.001)	(0.001)	(0.001)				
Covariates in first differences								
$\Delta(\#\text{ of Ventures})_{t-1}$					0.729***	0.618***	0.620***	0.647***
`					(0.016)	(0.011)	(0.011)	(0.009)
Federal funds								
$\Delta(\text{R\&D funds to univ.})_{t-1}$					0.141***	0.102***	0.105***	0.127***
					(0.024)	(0.011)	(0.011)	(0.014)
$\Delta(\text{Non R\&D funds to univ.})$					-0.067	0.104*	0.091*	0.077
					(0.052)	(0.046)	(0.045)	(0.062)
$\Delta(\text{R\&D funds to NL})_{t-1}$					0.000	0.016***	0.016***	0.015***
					(0.016)	(0.003)	(0.002)	(0.003)
Constant	-1,160***	7,5130***	-270**	-206*	9,051***	-10,300***	9,235***	9,021***
	(190.112)	(268.117)	(100.423)	(97.071)	(1,068.782)	(1,990.329)	(1,071.081)	(1,048.849)
Observations	163,200	212,160	326,400	374,279	34,410	68,820	80,290	102,825
R-squared	0.874	0.797	0.933	0.935	0.650	0.591	0.587	0.595
Number of ZCTAs	16,320	16,320	16,320	16,273	11,470	11,470	11,470	11,425
Year FE	YES	YES	YES	YES				
ZCTA FE	YES	YES	YES	YES				
Counties FE					YES	YES	YES	YES
State*Year FE					YES	YES	YES	YES

Note: In columns (1-4), both endogenous variable and covariates are expressed as 2-year averages of the original variables: $\overline{X}_t = \frac{x_t + x_{t-1}}{2}$. In columns (5-8), we reduce the 25-year panel into 12 periods and difference out the resulting panel: $\Delta\overline{X}_t = \overline{X}_t - \overline{X}_{t-4}$. Notation simplified and Number of Ventures scaled by 1000 for exposition. Robust standard errors in parenthesis. * p<0.05 ** p<0.01 *** p<0.001

.

TABLE C2

The Impact of Funding on RECPI: Robustness Checks

	Long Panel with FE [2 years]				*Long First Differences [2 years]*			
	Before 2000	After 2000	97-99 excluded	Stanford & MIT excluded	Before 2000	After 2000	97-99 excluded	Stanford & MIT excluded
	$(\text{RECPI})_t$ (1)	$(\text{RECPI})_t$ (2)	$(\text{RECPI})_t$ (3)	$(\text{RECPI})_t$ (4)	$\Delta(\text{RECPI})_t$ (5)	$\Delta(\text{RECPI})_t$ (6)	$\Delta(\text{RECPI})_t$ (7)	$\Delta(\text{RECPI})_t$ (8)
Covariates in levels								
$(\text{RECPI})_{t-1}$	0.950***	0.659***	0.874***	0.858***				
	(0.030)	(0.046)	(0.017)	(0.020)				
Federal funds								
$(R\&D\ funds\ to\ univ.)_{t-1}$	3.530***	0.941**	1.354***	1.103***				
	(0.765)	(0.336)	(0.160)	(0.169)				
$(\text{Non R\&D funds to univ.})_{t-1}$	0.902	-4.558***	-2.816**	-1.076				
	(1.028)	(1.335)	(0.934)	(0.735)				
$(\text{R\&D funds to NL})_{t-1}$	0.203	-0.004	-0.025	-0.030				
	(0.288)	(0.037)	(0.021)	(0.021)				
Covariates in first differences								
$\Delta(\text{RECPI})_{t-1}$					0.753***	0.280***	0.281***	0.341***
					(0.073)	(0.026)	(0.026)	(0.029)
Federal funds								
$\Delta(\text{R\&D funds to univ.})_{t-1}$					4.106*	0.535	0.672*	0.669+
					(1.739)	(0.352)	(0.322)	(0.397)
$\Delta(\text{Non R\&D funds to univ.})_{t-1}$					0.448	-8.808***	-8.491***	-4.553**
					(2.024)	(2.300)	(2.174)	(1.580)
$\Delta(\text{R\&D funds to NL})_{t-1}$					0.267	-0.153*	-0.181*	-0.183*
					(0.624)	(0.067)	(0.077)	(0.079)
Constant	-24,044**	149,564***	6,425*	10,075**	30,146	94,235***	20,397*	33,540**
Observations	163,200	212,160	326,400	374,279	34,410	68,820	80,290	102,825
R-squared	0.612	0.383	0.737	0.709	0.347	0.135	0.133	0.160
Number of ZCTAs	16,320	16,320	16,320	16,273	11,470	11,470	11,470	11,425
Year FE	YES	YES	YES	YES				
ZCTA FE	YES	YES	YES	YES				
Counties FE					YES	YES	YES	YES
State*Year FE					YES	YES	YES	YES

Note: In columns (1-4), both endogenous variable and covariates are expressed as 2-year averages of the original variables: $\overline{X}_t = \frac{x_t + x_{t-1}}{2}$. In columns (5-8), we reduce the 25-year panel into 12 periods and difference out the resulting panel: $\Delta\overline{X}_t = \overline{X}_t - \overline{X}_{t-4}$. Notation simplified and RECPI scaled by 1000 for exposition. Robust standard errors in parenthesis. * $p<0.05$ ** $p<0.01$ *** $p<0.001$

APPENDIX D

Case Study: Weber State University (UTAH)

In 1998, WSU's R&D federal funding more than tripled. The Table highlights the substantial increase in RECPI and in the Number of Ventures in the immediate vicinity of the university.

Distance	Year	# of Ventures	RECPI	Lagged R&D fed funding	Growth # of Ventures	Growth RECPI	Lagged Growth R&D fed funding
5-50 miles	1998	9,479	14,093	2,148,481			
5-50 miles	1999	10,528	14,464	2,242,103	11.1%	2.6%	4.4%
<5 miles	1998	686	564	1,502			
<5 miles	1999	803	798	5,247	17.1%	**41.7%**	**249%**

Note: R&D funding to WSU is lagged one year to mitigate reverse causality concerns: the 249% increase in funding refers to the 1997-1998 period. Again, all funding amounts are in thousands (2012 baseline year).