Paper to be presented at DRUID18
Copenhagen Business School, Copenhagen, Denmark
June 11-13, 2018

The role of universities in local entrepreneurial ecosystems

Valentina Tartari
Copenhagen Business School
Department of Innovation and Organizational Economics
vt.ino@cbs.dk

Scott Stern
MIT

sstern@mit.edu

Abstract

While universities are often associated with strong entrepreneurial ecosystems, the underlying drivers of this relationship have proved more elusive. We combine comprehensive business registration records with a predictive analytics approach to estimate both the quantity and (growth-oriented) quality of entrepreneurship at the level of individual zip codes over time. Moreover, we link these locations to the presence or absence of research-oriented universities or national laboratories, and we construct comparison groups based on ex ante similarities. Finally, we take advantage of significant changes over time in Federal commitments to both universities and national laboratories, and in particular of the distinction between research-oriented versus more general financial support for university activities. Together, these building blocks allow us to highlight three core findings related to the role of universities in local entrepreneurial ecosystems. First, universities are associated with not only a higher level of entrepreneurship but also a higher level of quality-adjusted entrepreneurship, and this relationship has strengthened over time. Second, relative to the direct impact of universities, demographic and economic factors associated with the presence of a university are even more strongly associated with entrepreneurial ecosystems. Finally, even after controlling for such factors, changes over time in resources enhance entrepreneurship but only increases in research-oriented funding enhance entrepreneurial quality. Together, these findings suggest both that universities as large economic institutions play a critical (and often underappreciated) role in local economic development, but that the norms and governance of universities play a unique role in promoting growth entrepreneurship conducive to long-term economic growth.

The role of universities in local entrepreneurial ecosystems

1. Introduction

Since the seminal work of Jaffe (1989), economics of innovation literature has highlighted spillovers from universities as key sources in promoting firm innovation and performance (Hall, Link, and Scott 2003). In particular, empirical studies have found that the university contribution to industrial innovation is greater the higher the quality of academic research and the closer firms are to universities (Mansfield 1995). Moreover, there is a growing literature documenting how knowledge spillovers more in general play an important role in fostering not only innovative activities, but more in particular entrepreneurship.

From the founding of land grant state universities and colleges as the result of the Morrill Act to direct efforts to promote technology transfer such as the Bayh-Dole Act, both the mission of universities as well as specific policies and programs have been premised on the potential of universities to serve as a spur for local economic development (Jaffe 1989; Rosenberg and Nelson 1994). Anecdotal evidence suggests highly innovative and performing entrepreneurial clusters are located in the vicinity of research universities, such as Silicon Valley around Stanford and the technology cluster around the University of Cambridge (known as the Cambridge Phenomenon). The basic idea is that university spillovers tend to be spatially bounded so that new firms need to locate close to universities to gain better access to those spillovers, both in terms of research results and of highly qualified human capital.

There are at least three distinct channels by which universities influence nearby entrepreneurial activities. The first one is through 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 2017) 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). The second one is through the creation of human capital: universities graduate students are likely advantaged in terms of underlying capability relative to the population, and are a flow of new people into the local economy each year. If students tend to stick around the place where they graduate (Krabel and Flöther 2014), the flow of new graduates represent a potential source of entrepreneurial capacity for a particular place. Finally, the third channel works through knowledge spillovers, universities produce ideas through the research activities they undertake. University research – particularly in certain fields – might be disproportionate source for entrepreneurial ventures with a high potential for growth.

By and large, attempts to gauge the impact of universities on local economic development have conflated these three ways by which universities influence economic development, or have focused narrowly on how a change in the institutional arrangements governing university commercialization (such as the Bayh-Dole Act) influences overall patterns of economic development (Hausman 2012), or on the production of high-tech firms (David B. Audretsch, Lehmann, and Warning 2005; Baptista and Mendonça 2010; Bonaccorsi et al. 2014).

In this paper, we aim at understanding if there is distinctive role that universities play through the combination of their focus on research and their focus on institutional rules encouraging the use of that research in the local environment. Though arguably narrow, this linkage is nonetheless crucial as the fulcrum over which the spillover effects of the whole economy grows. In order to answer this question, what you would need to evaluate is not simply whether there was an association between the presence of a university and growth entrepreneurship, but instead to evaluate four further dimensions of that relationship. First, whether that association was related structurally to the presence of the university per se or whether it was associated with the type of environment in which universities were located. Second, even if one controlled for the environment in a careful way, to figure out whether it was research or knowledge production per se, or simply the scaling effect arising from a university that was driving the effect. Third, you would want to know that it was a type of effect that operated particularly in universities, or whether this was simply the result of any type of knowledge investment (e.g., whether corporate R&D or a grant to a "closed" institution might also work). Finally, you would want to distinguish between different types of entrepreneurial ventures in order to distinguish the role of research investment in open organizations in shaping growth entrepreneurship relative to perhaps shaping more plain entrepreneurship. This is a formidable set of requirements. Prior work in this area have often been able to address elements of these questions, but have not been able to offer causal evidence on the impact of universities on both the quantity and quality of entrepreneurship.

To break through this impasse, we develop an integrated estimate of the impact of universities on local entrepreneurship. Our approach involves three fundamental steps. First, we identify both the presence and absence of universities and national laboratories, and changes over time in university funding for both research universities and national laboratories. Second, as it has been difficult to disentangle the direct impact of universities on the local environment for entrepreneurship from the potential for confounding factors to give rise to both strong universities and a strong local economy, we identify covariates of places that allow us to see a useful control group for universities in terms of places that "should" have a university but may not (for historical reasons). Third, we are able to measure systematically the number and average quality of all new businesses within given places over time and to distinguish different types of ventures at a micro-level of geographical granularity identify, leveraging the zip-code level aggregates that result from the predictive analytics approach of Guzman and Stern (2017, 2015).

Through both a cross-sectional and a panel data analysis, we uncover three main findings. First, universities matter (both relative to national laboratories and relative to no research institution), but so does the environment in which universities are located. . In particular, the effect of universities on entrepreneurial quality in an area is particularly pronounced in urban and high-income zip codes, and in zip codes located near a top-tier university. Second, the association between universities and entrepreneurship seems to be changing over time, in at least two ways. On the one hand, the types of places that host universities seem to have become more associated with the quantity of entrepreneurship over time (though the distinctive role of universities is severely reduced once you control for the underlying place characteristics). On the other hand, even after accounting for underlying place-based characteristics, there is a sharp (and cyclical) uptick in the association between universities and quality-adjusted entrepreneurship. Third, we find a significant effect (both statistically and quantitatively) of increases in *research* funding to universities on both the quantity and quality of entrepreneurship, with a much larger impact on quality than on pure quantity, while there is no such an effect stemming from changes in research funding to national laboratories.

We believe our results clearly 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. We also highlight how universities have become more relevant in time for local entrepreneurship, suggesting the presence of very localized spillovers from universities to local businesses.

2. Conceptual framework

The mechanisms by which universities contribute to entrepreneurship is a topic of great interest both in the academic and policy community. In particular, in this paper we argue that universities have a distinctive role in spurring high quality entrepreneurial activities in their vicinity, and that this influence operates through at least three distinct channels.

The first channel operates through an increase in demand. Universities are in general large organizations with many employees and large student populations, which require a large variety of services such as restaurants, cafés, shops, etc. This means that we expect to find a large number of small businesses located around universities and serving the local population. However, this channel should be more salient at the earliest stages of the establishment of universities (Andrews, 2017) or subsequently to very significant changes in the scale of universities activities (such as a campus expansion, or the opening of a new department). This means that in equilibrium we expect to see a sustained flow of new businesses (also in terms of turnover of existing ones), but not necessary different from other places.

The second channel operates through the creation of highly skilled human capital, namely university graduates. As highlighted by Saxenian (1996), the mobility of human capital, as embodied by graduate student exiting universities and entering the labor market, is one of the most important channels that facilitate knowledge spillovers. Graduates increase the qualification of the local labor force and use their skills in local firms (Abramovsky, Harrison, and Simpson 2007; Moretti 2004), thereby improving local innovative and economic performance. Evidence on the patterns of mobility of university students after graduation are mixed. In one study in the United States, Stephan and colleagues (2005) find that roughly two out of three doctoral students leave the region where they received their training once they take a job. On the other hand, a study on German university students found that self-employed graduates are more likely to stay within their university region (Krabel and Flöther 2014), as it has been observed that entrepreneurs tend to start their businesses in the regions in which they have deep social networks (Dahl and Sorenson 2012).

This suggests that the flow of new graduates represent a potential source of entrepreneurial capacity for a particular place.

The third channel is represented by knowledge spillovers that are not embodied by students. 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 (D. B. 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.

3. Empirical approach

One of the challenges we face in estimating the casual effect of university activities on entrepreneurship is that these activities do not occur randomly. In particular, while universities have long been associated with a progressive, "creative class" lifestyle, not all "creative class" or entrepreneurial clusters are linked closely with research universities. Indeed, universities are generally located in places that are conducive to high levels of economic dynamism, independent of the universities themselves. Ignoring this issue would lead to an overestimate of the real effect of universities on local entrepreneurship. In order to tackle this issue, we make use of census data to characterize precisely the geographical areas with and without universities. This allows us to predict econometrically the probability that a specific area would indeed host a university, allowing subsequent estimations to be "corrected" by the endogenous features that make ZIP codes proximate to a university different from ZIP codes located far away from universities.

In particular, we characterize all ZIP codes in the dataset using data from the 1990 US decennial census and we use these characteristics as covariates in a probit model estimating the likelihood of a certain ZIP code to be located near a university or a national laboratory. We can then use these estimated probabilities as a control variable in the cross-sectional regressions investigating the impact of the presence of research institutions on local entrepreneurial activity, effectively correcting for certain features of a location that have an impact on entrepreneurship, independently of the presence of a research institution.

A second challenge related to the endogenous nature of university research activity arises as universities may be affected by the local business community. There may be bi-directional knowledge spillovers and the presence of highly productive firms may increase the local demand for workers trained in a university setting who transition to local jobs. Our estimation strategy seeks to isolate the spillover effects of research universities' activities on their local economies. Our strategy is to difference-out time-invariant characteristics of ZIP codes, effectively addressing a wide class of potential selection problems. In particular, this strategy controls for any permanent differences across ZIP codes that are correlated with the scale of university activities: our results are identified from within-ZIP code changes in university activity over time.

Another issue, which needs to be taken into account in our econometric model, is the volatility of both the dependent and independent variables across the year. All variables in our model can show large peaks and drops in certain years; however, for example, the effect of a large increase in R&D funding to a specific university in a certain year does not necessarily exhaust its effect on the local entrepreneurial environment in the subsequent year. To smooth the time pattern of the variables we run our models using both 2 years and 3 years averages. This means that when analyzing 2 years average we reduce our 25-years long panel to a panel consisting of 12 periods (8 periods in the case of 3 years averages).

Our goal is to estimate the responsiveness of changes in entrepreneurial activity to changes in funding to research institutions (universities and national laboratories) in a ZIP code using a long-differences specification. For 2-years averages, we estimate the model as:

$$\Delta Y_{it} = \beta_1 Y_{it-1} + \beta_2 \Delta R\&D_{it-1} + \beta_3 \Delta nonR\&D_{it-1} + \beta_4 \Delta NatLabs_{it-1} + T + C + \varepsilon_{it}$$

where ΔY_{it} is the long difference in the average number of ventures or their quality in ZIP code *i*, in period *t* relative to period *t – 2* ($\Delta Y_{it} - \Delta Y_{it-x}$), $\Delta R\&D_{it-1}$ is the long difference in average federal funding to universities for R&D in ZIP code *i* in period *t – 1* relative to year *t – 3* ($\Delta R\&D_{it-1}$-$\Delta R\&D_{it-3}$), $\Delta nonR\&D_{it-1}$ is the long difference in average federal funding to universities for non-R&D activities in ZIP code *i* in period *t – 1* relative to year *t – 3* 5 ($\Delta nonR\&D_{it-1}$-$\Delta nonR\&D_{it-3}$), $\Delta NatLabs_{it-1}$ is the long difference in the average federal funding to National Laboratories in ZIP code *i* in period *t – 1* relative to year *t – 3* ($\Delta Nat\ Labs_{it-1}$-$\Delta NatLabs_{it-3}$), T is a set of year fixed effects, C is a set of county (or metropolitan area) fixed effects and ε_{it} is the error term. Our parameters of interest areβ_2, β_3, and β_4.

The long-difference specification in this equation addresses concerns that time-invariant ZIP code characteristics might bias our estimates of the true impact of university activity. Moreover, the term Y_{it-1} provides a baseline for previous entrepreneurial activity in the area. We also cluster the standard errors at the ZIP code level to address the fact that federal funds are measured at the ZIP code level. We examine the change in the number of ventures and in their quality between *t* and *t – 2* as a function of the changes in federal funding received between *t – 1* and *t – 3*. This ensures that time periods are separate and therefore we can isolate the effect of funding on entrepreneurship, and not vice versa.

4. Data

In order to implement the empirical analysis we presented, we draw form a unique dataset that combines detailed information about entrepreneurial and universities activities in 28 US States from 1988 to 2012.

The first building block of our dataset consists of entrepreneurial activity and quality data at the ZIP code level. Building on Guzman and Stern (2017, 2015), this paper focuses not only on the quantity of entrepreneurship (nor on highly selective measures of the rate of successful entrepreneurs) but instead on the measurement and assessment of entrepreneurial "quality." Guzman and Stern's approach combines three interrelated insights. First, a practical requirement for any growth-oriented entrepreneur is business registration (as a corporation, partnership, or limited liability company). Through these public documents, it is possible to observe a "population" sample of entrepreneurs observed at a similar (and foundational) stage of the entrepreneurial process. Second, the methodology moves beyond simple counts of business registrants (Klapper, Amit, and Guillén 2010) by measuring characteristics related to entrepreneurial quality at or close to the time of registration. These characteristics include how the firm is organized (e.g., as a corporation, partnership, or LLC, and whether the company is registered in Delaware), how it is named (e.g., whether the owners name the firm eponymously after themselves), and how the idea behind the business is protected (e.g., through an early patent or trademark application). These start-up characteristics may reflect choices by founders who perceive their venture to have high potential, and may represent early-stage "digital signatures" of high-quality ventures. Third, some firms (although rarely) experience meaningful growth outcomes (e.g., an IPO or high-value acquisition within six years of founding), and therefore can be used to estimate the relationship between these growth outcomes and start-up characteristics. This mapping allows to form an estimate of entrepreneurial quality for any business registrant within the sample (even those in recent cohorts where a growth outcome (or not) has not yet had time to be observed). Through this predictive analytics approach, the authors propose a set of statistics for the measurement of growth entrepreneurship. In this paper we use two measures: the number of start-ups within a ZIP code and the Regional Entrepreneurship Cohort Potential Index (RECPI). RECPI is calculated as the product of the average quality within any given group of firms and the number of start-ups within a given geographical region (e.g., from a zip code or town to the entire five-state coverage of our sample). RECPI allows the direct calculation of the expected number of growth outcomes from a given start-up cohort within a given geographical boundary (in our case ZIP code).

In order to estimate the probability of a certain ZIP code hosting a research institution, we need to link the data on entrepreneurial quantity and quality with socioeconomic data from the census. Entrepreneurship data are organized at the ZIP code level (as indicated in the registration documents), which we use as the basic geographical unit of analysis in our paper. Some considerations about ZIP codes and using them as geographical entities are therefore necessary in order to explaining the linking process to the census data.

ZIP Code service areas are not geographical entities with boundaries, they instead represents routes for mail delivery. They are created by the US Postal Service (USPS) and they fall into four types: military, PO box, large volume customer, and standard. Excluding the special case of the military (the same ZIP code is assigned irrespective of geographical location, inside and outside the U.S.), only standard ZIP Code areas actually represent a geographical space, which may be characterized in terms of area and population (and all subsequent measures related to the population). PO box and large volume customers are basically points in the map and therefore do not possess any characteristic that can be linked to census data.

In 2000, the U.S. Census Bureau created a new statistic entity called ZIP Code Tabulation Areas n(ZCTAs). ZCTAs are generalized area representations of USPS ZIP Code service areas. ZCTAs follow census block boundaries and they are not stable over time, as they are computer-delineated based on the location of addresses at the time of the Census. In some cases, ZCTAs may even cross state boundaries. In the 1990 census, a limited amount of information was available at the ZIP Code level, while from 2000, information is only available at the ZCTA level. Unfortunately, neither USPS, nor the Census Bureau release “official” concordance table that allow making the link between ZIP Code area and ZCTAs. We therefore relied on ZIP Code to ZCTA crosswalk produced in 2015 by the UDS mapper (financed by the HRSA), which contains 41251 unique ZIP Codes, the correspondent ZCTA, the type of ZIP Code, city and State. All but two ZIP codes in the entrepreneurship dataset were matched using the cross walk (the remaining two were assigned a ZCTA manually). By transforming ZIP codes into ZCTAs, we are sure we are analyzing entities, which have geographical properties (they are not just a single point in space), and therefore can be matches to socioeconomic characteristics. The quantity and quality variables in the entrepreneurship data have then been aggregated at the level of the ZCTA. In particular, the number of ventures for each ZCTA is obtained by summing the number of establishments in all the ZIP codes included in the ZCTA, while RECPI at the level of the ZCTA is the sum of the RECPI in all ZIP codes include in the ZCTA. In the paper, we refer to the geographical unit of analysis as ZIP code.

The resulting ZIP codes have then been matched with geographical coordinates (latitude and longitude) and decennial census data. The match has been successful for 86% of the ZIP codes for the 1990 census, 94.4% for the 2000 census and 100% for the 2010 census. Table 1 lists the socioeconomic variables we collected and which constitute the second building block of our dataset.

Insert Table 1 about here

The third building block relates to the data about universities. We have constructed a detailed dataset consisting of all higher education institutions in the United States, as listed by the National Center for Education Statistics. For each higher education institution, we know its precise location (ZIP code, city, state) and its Carnegie Classification. In particular, we create a dummy variable *research university* for the subset of institutions representing leading research colleges and universities (Doctoral/Research Universities, Research Universities – very high research activity, Research Universities – high research activity). We have additionally collected data from the NSF Survey of Federal Funds for Research and Development about federal funding to universities in each year, both for R&D and non-R&D related activities. We have also characterized the quality of universities using the Times Higher Education ranking of US universities: we have created a dummy variable called *top10* indicating if a university is in the top 10 of the ranking, and a variable *top50* if it is in the top 50 of the ranking.

The last building block relates to national laboratories. We use data from the Survey of R&D Expenditures at Federally Funded R&D Centers administered by the National Center for Science and Engineering Statistics at the National Science Foundation. This dataset contain funding records for all national laboratories, along with their precise geographical location (ZIP code, city, state).

In order to explore the effect of research organization activities on entrepreneurship, the last step in constructing our dataset has been to find which institutions are located in the vicinity of our ZIP codes. In order to do so, we created 5 miles (we used 2 miles and 10 miles as robustness checks) radius circumferences around the centroid of each ZIP code, and we used the latitude and longitude data to identify which institutions would fall in the specified area. By doing so, we know for each ZIP code how many higher education institutions and national laboratories are present in a 5-miles radius, and their aggregated funding (which we can differentiate by their research university status).

The final sample consist of 15950 ZIP codes in the time period 1988-2012.

5. Results

Column 1 and 2 of Table 2 show the means and standard deviations of various ZIP code characteristics (computed in 1990), diving the sample by whether a ZIP code has a research university in a 5 miles radius. Column 3 presents t-statistics for a test of differences in the means between columns 1 and 2. The comparison produces a number of interesting results. ZIP codes located near research universities are indeed very different from ZIP codes that are not along all dimensions considered. In particular, university ZIP codes are more populated (both in absolute terms and density-wise) and more urban, they have a higher share of the population in working age (18 years old to 65 years old), they are more ethnically diverse (both in terms of race and place of birth), they have a more educated population who works more predominantly in the private sector. Finally, both median income per capita and house values are higher in ZIP codes in the vicinity of universities.

Insert Table 2 about here

Table 3 presents the logistic regressions predicting the likelihood of having a research university (model 1a and 1 b) or a National Laboratory (model 2a and 2b) in a 5 miles radius of the focal ZIP code. Rather surprisingly, not all predictors have an analogous effect on the location of research universities versus national laboratories. For example, ZIP codes with higher density are more likely to have a research university in a 5 miles radius, but not a national laboratory. On the other hand, a high percentage of college-educated people in the ZIP code is predictive of the location of both types of institutions. These regressions represent the first stage through which we extract the predicted probability that we will employ in subsequent regressions in order to control for ZIP code characteristics, which are highly correlated with the presence of research organizations and if unaccounted for may cause an overestimation of the actual effect of universities on local entrepreneurial activities.

Insert Table 3 about here

Table 4a shows the basic cross-sectional effect of the presence of a research university or a research laboratory on the number of ventures created and their quality, by ZIP code. We estimate an OLS model with years fixed effects, and errors clustered at the ZIP code level. The presence of a research university has a positive impact on both the number and the quality of the ventures created in a certain ZIP code, while the presence of a national laboratory has no statistically significant effect.

Insert Table 4a about here

Table 4b extends the previous table by controlling for ZIP code characteristics in several different ways. Columns 1 and 2 include as regressors the ZIP code characteristics we have used in the logistic models presented in Table 3. Columns 3 and 4 we use the predictions extracted from the models in Table 3 as regressors, while in columns 5 and 6 we also add State fixed effects. From these models, we can indeed see that there are certain socioeconomic characteristics (which are characteristics of places near research organizations) that are positively correlated with entrepreneurial activity. However, even after controlling for them, there is a significant residual effect of the actual presence of a research university on the number and the quality of the ventures created in a certain ZIP code.

Insert Table 4b about here

In Table 5a we examine the effect of co-location with high quality universities (in the top 10 or top 50): while there is no statistically significant effect on the number of ventures created, having a very high quality (top 10) research university in the vicinity has a very strong a positive effect on the quality of those ventures.

Insert Table 5a about here

In Table 5b we explore if the effect of proximity to a research university is influenced by certain specific characteristics, namely urbanization and wealth, of the ZIP codes themselves. In particular, we interact the variable indicating the presence of a research university with two dummy variables: *urban* takes the value of one if 100% of the population of the focal ZIP codes reside in a urban area, *high income* takes the value of one if the median income in the focal ZIP code is in the 95^{th} percentile of income distribution of all ZIP codes. Results reveal that the positive effect of proximity to a research university on entrepreneurial quality is disproportionately important in ZIP codes with high levels of income and where the population resides in urban areas.

Insert Table 5b about here

We then turn our attention to the temporal dynamics of entrepreneurial quality. In Figure 1 we plot the effect of the presence of a research university and of a national laboratory (with and without controlling for the predicted probability of having a research organization in the vicinity of the ZIP code) on RECPI by year. There are several interesting observations we can draw from this plot. First, the effect of the presence of national laboratories is negligible along the whole time period investigated. Second, while introducing the predicted probabilities indeed reduces the magnitude of the effect of the presence of research universities on entrepreneurial quality, the effects does not disappear. Third, it is clear that the importance of co-location nearby research universities has increased over tie: from being almost negligible until the mid-1990s, in 2012 the vicinity to a research university would have an impact on RECPI four times bigger than in 1997.

Insert Figure 1 about here

To sum up the evidence so far, our results indicate that ZIP codes located in proximity of a research university or a national laboratory are different from ZIP codes located further away from this kind of institutions. In particular, these ZIP codes tend to be densely populated urban areas, with a high share of college-educated inhabitants. We also observe that only research universities have a positive impact on both the number of ventures created and their quality, while national labs have a negligible effect, even after controlling for ZIP codes differences, state fixed effects and years fixed effects. In particular, the effect of universities on entrepreneurial quality in an area is particularly pronounced in urban and wealthy ZIP codes, and in ZIP codes located near a top-tier university. Finally, we observe that the role of universities in spurring entrepreneurial ecosystems seems to have changed over time, namely increasing in importance, while we do not observe any time trend for national laboratories. While being largely descriptive, these results clearly highlight the important and unique role universities play in building and sustaining entrepreneurial ecosystems. We therefore move to our main goal of estimating the responsiveness of changes in entrepreneurial activity to changes in funding to research institutions (universities and national laboratories) in a ZIP code using a long-differences specification.

Table 6 and 7 show the panel specifications. In both tables, Column 1 is a panel data model with ZIP code and year fixed effects, with lagged variables (1 year) for the main independent variables. In this model, we do not consider the issue of large yearly variations and we expect a rather quick effect of changes in federal funding to research institutions on the local entrepreneurial ecosystem. Federal funding for R&D activities in universities has a positive effect on both the number of ventures created in the subsequent year in the focal ZIP code and their quality, while other federal funds to universities have only an impact on the number of ventures. Federal funds to national laboratories have no statistically significant effect.

Columns 2 and 3 present the results of the long-differences specification for 2-years averages variables with year fixed effects. Column 2 also includes Metropolitan Area fixed effects, while column 3 includes County fixed effects. In this specification, only changes in federal funding for R&D activities in universities have a positive and statistically significant effect on both the quantity and quality of entrepreneurial activities. In column 4 we perform a robustness check and we employ 3-years averages of variables instead of 2-years averages: results are unchanged.

Insert Table 6 and 7 about here
----

6. Conclusion

In this paper we provide systematic empirical evidence for the unique role that universities play in shaping local entrepreneurial ecosystems, and in so doing allow for a reconceptualization of the underlying theoretical and policy drivers of the impact of universities on entrepreneurship. We present three main sets of results. First, universities are associated with not only a higher level of entrepreneurship but in particular a higher level of quality-adjusted entrepreneurship, and this relationship has strengthened over time. Second, relative to the direct impact of universities, demographic and economic factors associated with the presence of a university (income levels, diversity) are even more strongly associated with the presence of a strong entrepreneurial ecosystem. Finally, even after controlling for such factors, changes over time in resources enhance entrepreneurship but only increases in research-oriented funding enhance entrepreneurial quality. Together, these findings suggest not only that universities as large economic institutions have a critical (and often underappreciated) role in local economic development, but also that the specific norms and governance of universities play a unique role in promoting growth entrepreneurship conducive to long-term economic growth.

While we believe our works shows a clear link between university researcher activities and local high-growth entrepreneurship, future work would benefit from a careful micro-level analysis of exactly how these knowledge spillovers influence local ventures. As universities are increasingly asked to contribute to actively economic growth through commercialization activities, it would be interesting to understand the extent that faculty-generated academic spin-offs contribute to the high-growth entrepreneurial activities in the vicinity of universities. Moreover, future work could focus on potential industry-related contingencies of this link between universities and entrepreneurial activities. In particular, understanding how firms in certain industries may benefit from scientific specialization of universities activities would help shed further light on the role of universities in stimulating economic activity.

References

Abramovsky, Laura, Rupert Harrison, and Helen Simpson. 2007. "University Research and the Location of Business R&D." *Economic Journal* 117(519): 114–41.

Adams, J. D. 2002. "Comparative Localization of Academic and Industrial Spillovers." *Journal of Economic Geography* 2(3): 253–78.

Andrews, Michael. 2017. *The Role of Universities in Local Invention: Evidence from the Establishment of U.S. Colleges*.

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

Audretsch, David B., Erik E. Lehmann, and Susanne Warning. 2005. "University Spillovers and New Firm Location." *Research Policy* 34(7): 1113–22.

Baptista, Rui, and Joana Mendonça. 2010. "Proximity to Knowledge Sources and the Location of Knowledge-Based Start-Ups." *Annals of Regional Science* 45(1): 5–29.

Bonaccorsi, Andrea, Massimo G. Colombo, Massimiliano Guerini, and Cristina Rossi-Lamastra. 2014. "The Impact of Local and External University Knowledge on the Creation of Knowledge-Intensive Firms: Evidence from the Italian Case." *Small Business Economics* 43(2): 261–87.

Dahl, Michael S., and Olav Sorenson. 2012. "Home Sweet Home: Entrepreneurs' Location Choices and the Performance of Their Ventures." *Management Science* 58(6): 1059–71.

Guzman, Jorge, and Scott Stern. 2015. "Innovation Economics. Where Is Silicon Valley?" *Science* 347(6222): 606–9.

———. 2017. "Nowcasting and Placecasting Entrepreneurial Quality and Performance." In *Measuring Entrepreneurial Businesses: Current Knowledge and Challenges*, eds. John Haltiwanger, Erik Hurst, Javier Miranda, and Antoinette Schoar.

Hall, Bronwyn H., Albert N. Link, and John T. Scott. 2003. "Universities as Research Partners." *Review of Economics and Statistics* 85(2): 485–91.

Hausman, Naomi. 2012. "University Innovation, Local Economic Growth, and Entrepreneurship." *SSRN Electronic Journal* (October): 48.

Jaffe, A. B. 1989. "Real Effects of Academic Research." *The American Economic Review* 79(5): 957–970.

Jaffe, A. B., M. Trajtenberg, and R. Henderson. 1993. "Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations." *The Quarterly Journal of Economics* 108(3): 577–98.

Klapper, L, R Amit, and MF Guillén. 2010. NBER *Entrepreneurship and Firm Formation across Countries*..

Krabel, Stefan, and Choni Flöther. 2014. "Here Today, Gone Tomorrow? Regional Labour Mobility of German University Graduates." *Regional Studies* 48(10): 1609–27.

Lucas, Robert E. 1988. "On the Mechanics of Economic Development." *Journal of Monetary Economics* 22(1): 3–42.

Mansfield, Edwin. 1995. "Academic Research Underlying Industrial Innovations: Sources, Characteristics, and Financing." *The Review of Economics and Statistics* 77(1): 55.

Moretti, Enrico. 2004. "Estimating the Social Return to Higher Education: Evidence from Longitudinal and Repeated Cross-Sectional Data." *Journal of Econometrics* 121(1–2): 175–212.

Rosenberg, Nathan, and Richard R. Nelson. 1994. "American Universities and Technical Advance in Industry." *Research Policy* 23(3): 323–48.

Saxenian, AL. 1996. *Regional Advantage*.

Stephan, Paula. E., Albert J. Sumell, James D. Adams, and Grant C. Black. 2005. "Firm Placements of New PhDs: Implications for Knowledge Transfer." In *The Role of Labour*

Mobility and Informal Networks for Knowledge Transfer, Boston: Kluwer Academic Publishers, 123–46.

Table 1: Census data

Variable	**Description**
Total population	Total population
Density	Total population divided by the geographical area
Male	Percentage of men in the total population
Urban	Percentage of the population residing in urban areas
Age between 18 and 65	Percentage of the population between 18 and 65 years old
White	Percentage of the population of white ethnicity
Born outside the US	Percentage of the population born outside of the United States
College	Percentage of the population with a college degree or higher
Private sector	Percentage of the population employed in the private sector
Public sector	Percentage of the population employed in the public sector
Income per capita	Median per capita income
Housing values	Median house value

Table 2: Descriptive statistics

	No research university (14936 ZIP codes)		At least one research university in a 5 miles radius (1014 ZIP codes)		
Variable (Year 1990)	**Mean**	**Std. Dev.**	**Mean**	**Std. Dev.**	**Diff.**
Total population	8490	12055	24919	16128	***
Density	0,0003	0,0009	0,0023	0,0025	***
Men	0,4964	0,0368	0,4900	0,0463	***
Urban	0,3144	0,4149	0,9548	0,1575	***
Age between 18 and 65	0,5903	0,0643	0,6536	0,0839	***
White	0,8969	0,1678	0,7082	0,2654	***
Born outside the US	0,0372	0,0717	0,1190	0,1118	***
College education (or higher)	0,2040	0,1214	0,3407	0,1871	***
Empl. in the private sector	0,7169	0,1193	0,7723	0,0775	***
Empl. in the public sector	0,1536	0,0838	0,1616	0,0775	**
Median income per capita	12338	5262	16547	10364	***
Median house value	69603	61397	133006	115713	***

Table 3: Logistic regressions predicting the probability of a ZIP code to have a research university or a national laboratory in a 5 miles radius

VARIABLES	(1) Research University	(2) Research University - Odds Ratio	(3) National Laboratory	(4) National Laboratory - Odds Ratio
Population	-0.00	1.00	0.00+	1.00+
	(0.000)	(0.000)	(0.000)	(0.000)
Density	349.30***	5.00e+151***	-1.45	0.23
	(41.303)	(2.067e+153)	(47.288)	(11.076)
Men	-3.72*	0.02*	1.56	4.76
	(1.636)	(0.040)	(1.765)	(8.393)
Urban	3.96***	52.64***	3.88***	48.29***
	(0.304)	(16.016)	(0.883)	(42.631)
Age between 18 and 65	4.41***	82.08***	-2.79*	0.06*
	(0.904)	(74.185)	(1.292)	(0.080)
White	-2.26***	0.10***	-1.37**	0.26**
	(0.207)	(0.022)	(0.436)	(0.111)
Born outside US	-1.32*	0.27*	0.27	1.31
	(0.511)	(0.137)	(0.711)	(0.931)
College education (or higher)	3.55***	34.87***	5.09***	161.69***
	(0.531)	(18.522)	(0.826)	(133.599)
Empl. in the private sector	2.53*	12.57*	-2.75	0.06
	(1.194)	(14.998)	(1.796)	(0.114)
Empl. in the public sector	4.20**	66.66**	2.52	12.39
	(1.299)	(86.597)	(1.643)	(20.356)
Median income per capita	-0.00	1.00	-0.00**	1.00**
	(0.000)	(0.000)	(0.000)	(0.000)
Median house value	-0.00	1.00	0.00**	1.00**
	(0.000)	(0.000)	(0.000)	(0.000)
Constant	-8.35***	0.00***	-5.64***	0.00***
	(1.227)	(0.000)	(1.532)	(0.005)
Observations	15,950	15,950	15,950	15,950

Robust standard errors in parentheses
*** p<0.001, ** p<0.01, * p<0.05, + p<0.1

Table 4a

VARIABLES	(1) Number of ventures	(2) RECPI
Research university in 5 miles radius	91.46***	679.11***
	(4.978)	(70.935)
National Lab in 5 miles radius	-11.35	16.84
	(9.379)	(100.101)
Constant	12.29***	35.74***
	(0.425)	(3.737)
Observations	398,750	398,750
Number of zipcode	15,950	15,950
Year FE	YES	YES

Robust standard errors in parentheses
*** p<0.001, ** p<0.01, * p<0.05, + p<0.1

Table 4b

VARIABLES	(1) Number of ventures	(2) RECPI	(3) Number of ventures	(4) RECPI	(5) Number of ventures	(6) RECPI
Research Univ in 5 miles radius	15.58**	351.48***	12.84*	278.28***	15.70**	314.13***
	(5.279)	(71.555)	(6.282)	(74.015)	(5.641)	(74.320)
National Lab in 5 miles radius	-38.81***	-163.40+	-29.57**	-167.99	-22.59*	-198.20+
	(7.127)	(95.516)	(10.114)	(107.982)	(9.692)	(103.686)
Population	0.00***	0.00*				
	(0.000)	(0.001)				
Density	-4,381.17***	-28,798.92*				
	(1,054.243)	(11,454.959)				
Men	-63.86*	754.53+				
	(27.512)	(417.154)				
Urban	21.26***	50.40+				
	(2.486)	(26.075)				
Age between 18 and 65	33.03*	635.43***				
	(13.966)	(148.155)				
White	11.81*	-160.45***				
	(5.256)	(43.023)				
Born outside US	306.34***	1,231.99***				
	(36.248)	(174.484)				
College education (or higher)	70.79***	-29.96				
	(11.348)	(122.618)				
Empl. in the private sector	21.73***	-310.15***				
	(5.255)	(48.562)				
Empl. in the public sector	-20.14**	-529.50***				
	(7.798)	(86.722)				
Median income per capita	0.01***	0.01**				
	(0.001)	(0.005)				
Median house value	-0.00***	0.00***				

	(0.000)	(0.000)				
Research Univ - prediction			271.68***	1,048.21***	248.55***	931.22***
			(15.162)	(136.682)	(13.950)	(127.943)
National Lab - prediction			16.17	4,095.47**	-33.49	1,916.28+
			(110.880)	(1,353.141)	(102.420)	(1,157.912)
Constant	-90.67***	-642.20**	0.03	-29.94***	-16.91***	-83.65***
	(13.531)	(227.510)	(0.468)	(5.136)	(3.889)	(5.789)
Observations	398,750	398,750	398,750	398,750	398,750	398,750
Number of zipcode	15,950	15,950	15,950	15,950	15,950	15,950
Year FE	YES	YES	YES	YES	YES	YES
State FE	NO	NO	NO	NO	YES	YES

Robust standard errors in parentheses
*** p<0.001, ** p<0.01, * p<0.05, + p<0.1

Table 5a

VARIABLES	(1) Number of ventures	(2) RECPI
Research university in 5 miles radius	13.60*	184.28**
	(5.653)	(65.963)
National Lab in 5 miles radius	-23.74*	-540.34***
	(10.598)	(154.797)
University - Top 50	9.34	208.56
	(14.364)	(201.149)
University - Top 10	-0.56	1,365.98***
	(22.760)	(375.864)
Research Univ - prediction	247.73***	880.17***
	(13.969)	(129.486)
National Lab - prediction	-38.18	1,171.40
	(102.647)	(1,104.134)
Constant	-16.89***	-81.41***
	(3.890)	(5.516)
Observations	398,750	398,750
Number of zipcode	15,950	15,950
Year FE	YES	YES
State FE	YES	YES

Robust standard errors in parentheses
*** p<0.001, ** p<0.01, * p<0.05, + p<0.1

Table 5b

VARIABLES	(1) Number of ventures	(2) RECPI	(3) Number of ventures	(4) RECPI
Research university in 5 miles radius	29.61***	62.65	9.01+	138.42*
	(7.151)	(51.697)	(4.826)	(63.918)
Urban ZIP code	54.59***	177.93***		
	(4.090)	(32.532)		
Research Univ*Urban ZIP code	-28.34**	306.59**		
	(9.444)	(100.295)		
National Lab in 5 miles radius	-22.56*	-189.74+	-25.43**	-225.92*
	(9.711)	(106.048)	(8.607)	(103.429)
Research Univ - prediction	152.59***	517.92**	222.03***	817.60***
	(15.662)	(159.643)	(12.207)	(117.416)
National Lab - prediction	30.81	2,464.55*	-194.95*	893.76
	(96.715)	(1,174.128)	(81.635)	(1,027.934)
High income ZIP code			75.37***	347.43***
			(4.141)	(29.769)
Research Univ*High income ZIP code			26.76*	676.80***
			(12.721)	(188.958)
Constant	-15.86***	-79.69***	-19.80***	-96.45***
	(3.958)	(5.475)	(4.264)	(8.903)
Observations	398,750	398,750	398,750	398,750
Number of zipcode	15,950	15,950	15,950	15,950
Year FE	YES	YES	YES	YES
State FE	YES	YES	YES	YES

Robust standard errors in parentheses
*** p<0.001, ** p<0.01, * p<0.05, + p<0.1

Table 6 – Number of ventures

VARIABLES	(1) Panel with FE	(2) First diff - 2YAvg	(3) First diff - 2YAvg	(4) First diff - 3YAvg
2YAvg N of ventures – 2Y AVG – difference – lag2		0.0462	0.1344**	
		(0.045)	(0.043)	
2YAvg federal funds to univ - R&D - difference – lag1		0.0001***	0.0001***	
		(0.000)	(0.000)	
2YAvg federal funds to univ - non R&D - difference – lag1		0.0000	0.0001	
		(0.000)	(0.000)	
2YAvg federal funds to national lab - difference – lag1,		-0.0000	0.0000*	
		(0.000)	(0.000)	
Number of ventures - 2 years avg – lag2	0.9043***			
	(0.012)			
Federal funds to univ - R&D - 2 years avg – lag1	0.0000+			
	(0.000)			
Federal funds to univ - non R&D - 2 years avg – lag1	0.0003***			
	(0.000)			
Federal funds to national lab - 2 years avg – lag1	0.0000			
	(0.000)			
3YAvg N of ventures – 3Y AVG – difference – lag2				0.2107***
				(0.021)
3YAvg federal funds to univ - R&D - difference – lag1				0.0001***
				(0.000)
3YAvg federal funds to univ - non R&D - difference – lag1				0.0001
				(0.000)
3YAvg federal funds to national lab - difference – lag1				0.0000+
				(0.000)
Constant	0.6746**	6.4028***	11.0785***	16.5876***
	(0.257)	(1.416)	(0.832)	(1.334)

	(1)	(2)	(3)	(4)
Observations	350,900	75,376	89,688	44,844
R-squared	0.797			
Number of zipcode	15,950	9,422	11,211	11,211
Year FE	YES			
ZIP code FE	YES			
MA FE		YES		
County FE			YES	YES

Robust standard errors in parentheses
*** p<0.001, ** p<0.01, * p<0.05, + p<0.1

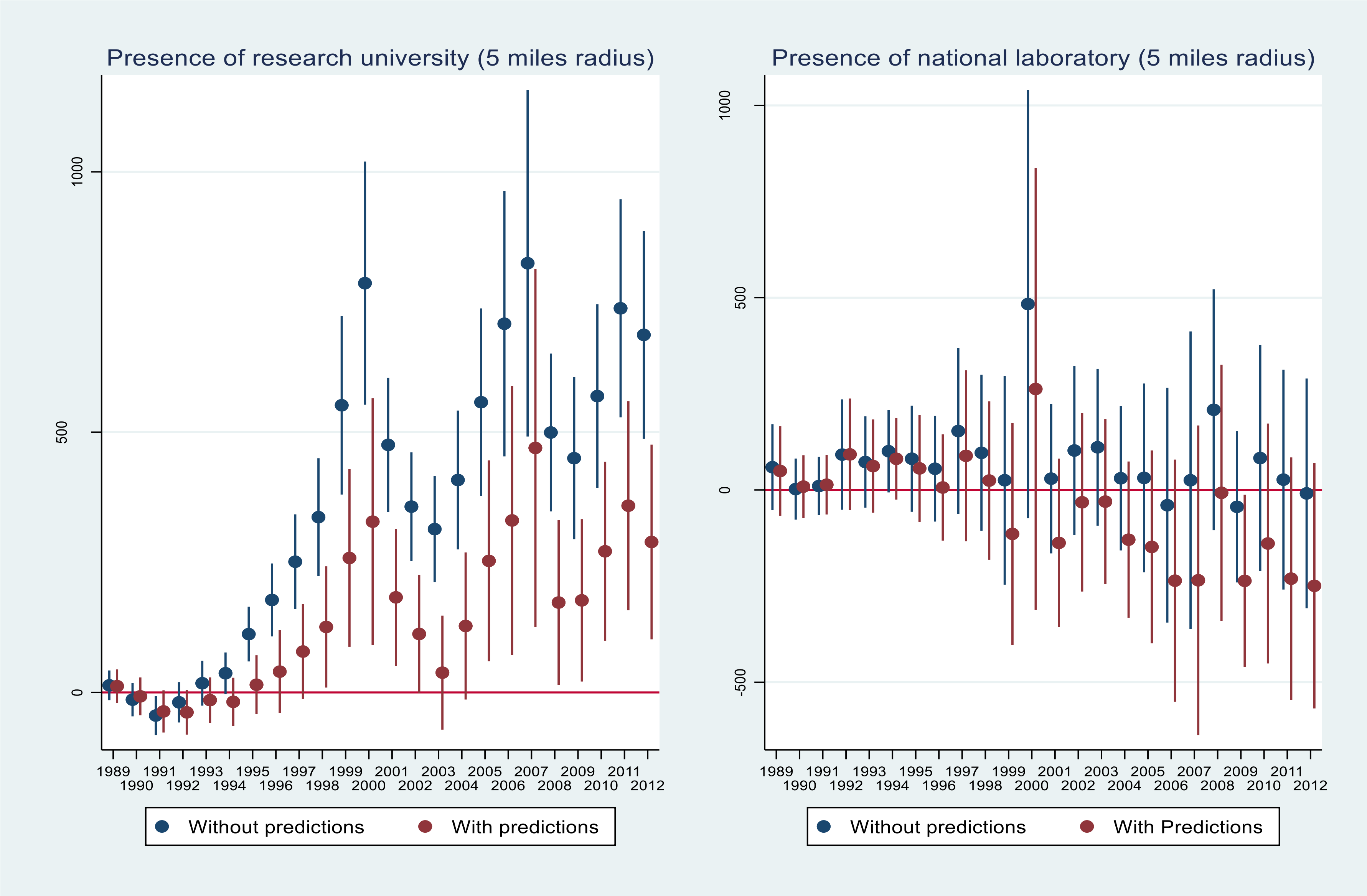
Table 7 - RECPI

VARIABLES	(1) Panel with FE	(2) First diff - 2YAvg	(3) First diff - 2YAvg	(4) First diff - 3YAvg
2YAvg RECPI – 2Y AVG – difference – lag2		-0.543***	-0.531***	
		(0.130)	(0.135)	
2YAvg federal funds to univ - R&D - difference – lag1		0.003***	0.003***	
		(0.001)	(0.001)	
2YAvg federal funds to univ - non R&D - difference – lag1		-0.002	-0.002	
		(0.002)	(0.002)	
2YAvg federal funds to national lab - difference – lag1,		-0.000*	-0.000*	
		(0.000)	(0.000)	
RECPI - 2 years avg – lag2	0.498***			
	(0.039)			
Federal funds to univ - R&D - 2 years avg – lag1	0.002***			
	(0.000)			
Federal funds to univ - non R&D - 2 years avg – lag1	-0.003+			
	(0.002)			
Federal funds to national lab - 2 years avg – lag1	-0.000**			

	(0.000)			
3YAvg RECPI – 3Y AVG – difference – lag2				-0.232***
				(0.055)
3YAvg federal funds to univ - R&D - difference – lag1				0.003***
				(0.001)
3YAvg federal funds to univ - non R&D - difference – lag1				-0.005
				(0.004)
3YAvg federal funds to national lab - difference – lag1				-0.000*
				(0.000)
Constant	29.728***	33.453***	90.653***	48.384**
	(3.958)	(7.308)	(18.396)	(15.963)
Observations	350,900	75,376	89,688	44,844
R-squared	0.266			
Number of zipcode	15,950	9,422	11,211	11,211
Year FE	YES			
ZIP code FE	YES			
MA FE		YES		
County FE			YES	YES

Robust standard errors in parentheses
*** p<0.001, ** p<0.01, * p<0.05, + p<0.1

Figure 1



Presence of research university (5 miles radius)
1000
500
0
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
Without predictions
With predictions
Presence of national laboratory (5 miles radius)
1000
500
0
-500
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
Without predictions
With Predictions