First Version: February 2007

This Version: October 2007

The Impact of Venture Capital on Innovation and the Creation of New Businesses[1]

Marcos A. Mollica

UBS

and

Luigi Zingales

University of Chicago, NBER & CEPR

Abstract

We exploit the cross section, cross industry, and time series variability of VC investments in the United States to study the impact of venture capital activity on innovation and the creation of new businesses. As a measure of the quality of research in a certain area we use the number of citations of academic papers produced by faculty in the area. As an instrument for the size of VC investments we use the size of state pension fund's assets. Even with these controls, we find that VC investments have a significant positive effect both on the production of patents and on the creation of new businesses. A one standard deviation increase in the VC investment per capita generates an increase in the number of patents between 4 and 15%. An increase of 10% in the volume of VC investment increases the total number of new business by 2.5%.

[1] Luigi Zingales thanks the Kauffman foundation, the Stigler Center, and the Initiative on Global Markets at the University of Chicago for financial support.

The American venture capital industry role in promoting technological innovation is often cited as a key source of comparative advantage for the U.S. economy. Inspired by the successes of venture capital (VC) financed successes, such as Google, Yahoo, and Microsoft, many other countries have tried to jump start a local VC industry in the hope of boosting their own economic development. Yet there is remarkably little systematic evidence documenting the effective contribution of VC to the development of new high tech businesses. Kortum and Lerner (2000) find that VC backed firms produce more and more valuable patents. Hellman and Puri (2000, 2001) show that VC backed firms are faster in developing products and introducing them in the market and they have a higher rate of executive turnover. Furthermore, the presence of VC sector seems to produce superior macroeconomic performance, Wasmer and Will (2000).

The crucial question, however, is to determine whether VCs naturally go where they are needed or whether the pre-existence of a VC industry fosters the implementation of innovation. In other words, did the Silicon Valley became such just because Stanford was there or also because some VCs early n sat up shops there.

This paper tries to identify what is the effect that an exogenous increase in the level of VC investments has on the number of patents (our proxy for innovation) and on the number of new businesses exploiting those patents. One problem in doing so is that a positive association of VC investment and number patents may be driven by the correlation of these two variables with a third one: for instance, the availability of opportunities for innovation in a particular state. Another problem is reverse causality:

the rise in the number of patents might attract the VC investments and not the other way around.

We address the omitted variable problem in two ways. First, to control for omitted factors related to the supply of high-tech ideas we use the production of academic papers in different scientific areas in different US regions. Second, we use the time series and cross industry differences in the quality of academic production to identify the effect of VCs. In other words, it might be hard to disentangle whether VCs located in the Bay area because of the blossoming computer industry or the other way around, but we can try to disentangle whether the Bay area has a comparative advantage vis-à-vis Madison (WI) in translating medical research into innovation as a result of the VCs who located there to cater to the computer industry.

As Lerner (1995) clearly shows, VC firms tend to invest in projects close to their headquarters, to minimize the cost of their advising and monitoring. Hence, we can safely divide the United States in small geographical units and consider them as separate VC markets. As a regional unit we use the Bureau of Economic Analysis (BEA) economic areas, which are composed by counties surrounding metropolitan areas. There are a total of 179 BEA areas. By using these more narrowly defined regional units, it is possible to have a more precise estimate of the regional impact of the VC investment on the number of patents and the creation of new businesses. The high number of BEA areas also substantially increases the degrees of freedom of our regressions, allowing for a more precise identification of the effects.

To address the potential reverse causality problem we use the total assets of local and state pension funds as an instrument for VC investments. The idea is that state pension funds are subject to political pressure to invest some of their funds in new businesses in the states. Hence, the size of the state pension fund triggers a shift in the local supply of VC investment that should help us identify the effect of VC on patents.

To this purpose, we use a two stage least square regression where in the first stage we run an OLS regression of VC investment on pension funds assets, time, region and sector fixed effects. The second stage is the Poisson model with VC investment predicted from the first stage plus other controls. To account for the generated regressor problem, in this specification we bootstrap the standard errors.

We find that VC investments have a significant positive effect on the production of patents. A one standard deviation increase in the VC investment per capita generates an increase in the number of patents between 4 and 15%. We also find that academic production of scientific papers is an important determinant of the number and the importance of patents.

Finally, we are interested to know whether the positive impact of VCs on patents translates to a greater number of new businesses. This question is not only interesting per se, but it is also a useful reality check regarding the relation between patents and VC investment. In fact, it could be that the positive association between patents and VC investment is present only because VCs tend to patent more. By contrast, we find that VC investments have an impact on new businesses even when the effect of patent is accounted for. An increase of 10% in the volume of VC investment increases the total number of new business by 2.5%.

Our results have important policy implications. Many states are thinking about subsidizing local venture capital activity. Our paper suggests that this could be a relatively high return activity. But it also suggests that subsidizing University research could be profitable. In future versions of this paper we will try to assess which one of the two channels is more cost effective.

In its analysis of the real effect of VC, this paper is most closely related to Kortum and Lerner (1998), Ueda et all (2003), and Darby and Zuker (2003). As in this last paper, we look at the effect of research and VC investment on new business entry. They focus only on the nanotechnology sector, while we explore this issue across sectors, locations, and time, reaching similar results on the importance of university research, but opposite results on the effect of VC investments

Our paper is also related to the literature on spillover effects of universities to the regional economies. Jaffe (1989), for instance, finds that university research has a significant positive effect on corporate patent and an indirect effect on local innovation by inducing R&D spending (see also Anselin, Varga and Acts (2002), Adams (2003), and Audretsch, Lehmann and Warning (2003).

The rest of the paper proceeds as follows. Section 1 describes the data and the methodology. Section 2 presents the statistical model used and Section 3 the results. Section 4 concludes.

1- Data Description and Methodology

The Bureau of Economic Analysis defines the BEA areas as areas that define the relevant regional markets surrounding metropolitan or micropolitan statistical areas. They consist of one or more economic nodes - metropolitan or micropolitan statistical areas that serve as regional centers of economic activity - and the surrounding counties that are economically related to the nodes. The economic areas were redefined on November 17, 2004, and are based on commuting data from the 2000 decennial population census, on redefined statistical areas from OMB (February 2004), and on newspaper circulation data from the Audit Bureau of Circulations for 2001. There are 179 regions in the present division used by the BEA. For a more detailed explanation of these areas see Johnson and Kort (2004).

Basically, there are three main reasons to use these geographical units instead of states. First, the VC firms, due to their advisor and monitor roles, tend to be located close to the business they are involved. Therefore, analyzing the VC investment by BEA regions is a better approximation to the reality of this business. Second, the BEA is composed by 179 regions as opposed to 50 states, significantly increasing the degrees of freedom and the reliability of the estimates. Finally, in spite of the fineness of this partition, we still can get, thanks to the Bureau of Economic Analysis, a large number of economic data at the BEA-region level, such as personal income, wages and employment, which could be used as controls in the empirical model.

To obtain a regional specific measure of scientific production we use the 1981-2003 University Science Indicators from Thomson scientific. This database contains the number of ISI-indexed papers from each university and the number of times the papers were cited through 2003 for 284 research universities in the United States. The use of citation data is especially important because it gives a measure of the quality of the scientific production. Unfortunately, citations have a problem of their own. Since they are cumulative, more recent papers get fewer cites because they have less years to be cited. To attenuate this problem, while we have citation up to 2003, we only use data up to 1999, leaving few years for citations to accumulate. In addition, we divide the number of citations by the number of years this paper can possibly be cited. Hence, all these figures should be interpreted as average citations per year.

To obtain a BEA specific measure of scientific production we use the total sum of the average yearly citation-weighted papers produced by faculty located in each of the 179 BEA areas. In the database there are 74 fields of research. For reasons that will be clearly momentarily, we group these 74 fields into six main groups. Table 1 presents the summary statistics for these six main groups across the 179 BEA regions and the eight years between 1992 and 1999.

As Table 1 shows, there is a lot of variability in average citation-weighted papers by field. While the media tends to be close to zero, the mean varies between the 6 yearly citation-papers in communication and media to the 1,085 in biotechnology. This variability is an important source of identification of region and sector specific effects.

In Table 2 we show the number of papers for the six tech sectors in the top 15 region for a particular year (1998). The salient feature here is the presence of substantial variability across regions and across sectors. While the Boston area is toward the top in every field, Washington Baltimore is at the top in the Computer area, but has only half of the papers in biotechnology with respect to Boston and New York.

We obtain the data on VC investments from Venture Economics. We use the firm ZIP code to map the investments into BEA regions. To express the value in constant dollar we deflate by the Producer Price Index (PPI), so that all the value are expressed in million of 1982 US$. We focus only on those sectors that could potentially be linked with academic research. Venture economics have 10 major industries: Biotechnology, Media and Communications, Computer hardware, Computer Software and Services, Industrial/Energy, Medical/Health, Semiconductors/Other Electrical, Consumer Related, Internet Specific, Other Products. We classify the first seven as Tech related, which might be affected by University research and patents, and the remaining three as non tech related. The Tech related investments account for about 65% of the VC investment in US. Since we cannot distinguish the patents in computer hardware and software, we club these two sectors together.

The overall mean of this variable showed in Table 1 is US$ 9.6 million and the median is zero, reflecting the large number of BEA region-years with no VC investments (around 50% of the observations). The sector with highest average investment is Computer ($19M), with a maximum of $2.5 billion.

In Table 3 we show in greater detail the distribution of VC investments in BEA regions by main investment ranges. Most of the areas have positive investment in all the periods. In the early 1980s only 21 BEA regions had more than 100M of VC investments.

By the end of the 1990s, this number had increased to 61 (a third of the sample). This change reflects the explosion of the VC activity during the sample period.

In Table 4, we show the VC investment by sector. What we call technological sector accounts for more that 70% of the investment in the period. Finally, in Table 5, we show the VC investment on the 6 tech sectors in the top 15 region for a particular year (1998). Also in this case we observe a substantial variability across regions and across sectors. Not surprisingly, the Bay area is first for most sectors, but not for Communication and Media, where the S. Louis area has twice as much investment, and not for the Medical area, where the San Antonio area is more important.

The patent data come from the NBER patent dataset, which includes all the citations received by US patents with application year between 1980 and 1999. As a main variable we are going to use citation adjusted patents. As an alternative one, the raw number of patents. While the latter is a rougher measure because it does not include an adjustment for quality, it does not suffer from the truncation problem that affect the number of citations, given that we do not observe the citations after 1999.

Patents are divided in 2 groups: assigned and unassigned patents. In the United States, a granted patent is assigned to the inventor. If the inventor works for a company, as part of their employment, the patent is generally assigned to the employer before it is granted. If the inventor is an individual, the patent remains assigned to her/him or may remain unassigned if the patent owner and the inventor are one and the same.[2]

---

[2] For a detailed exposition of the methodology employed in the construction of the NBER Patent Dataset see Hall, Jaffe and Trajtenberg (2000).

Since we are not interested in the innovation that takes place within large publicly traded companies, we eliminate all the patents assigned to companies with a COMPUSTAT code (mostly publicly traded companies) and we use only the residual. As we can see in Table 1, the average number of citation-patents is 32, with a range from 0 to 38,647.

We gather data on pension funds from the State and Local Government Employee-Retirements Systems annual survey conducted by the Census Bureau. The survey includes public employee retirement systems administered by state and local governments throughout the nation. We use deflate the total pension assets of these funds deflated by the PPI. The data is available from 1993 to 1999. From Table1 we see that the average asset of these funds is US$ 5.9 billions, the median is US$ 5.9 billions and the standard deviation is US$ 11.8 billions. Notice that this variable does not have sector variability.

To capture the effect of patents and VC investments on new businesses we would like to have a measure of start ups by BEA areas. Unfortunately, we do not have this exact measure. By BEA region we were only able to find the existing number of establishments every year from the US Census Bureau *County Business Pattern* dataset. Thus, we calculate the yearly change in the total number of establishments in each BEA area, where an establishment is defined as “a single physical location at which business is conducted”.

This proxy has two limitations. First, it is derived from the total number of establishments not the total number of firms. A firm can have many establishments, if it operates in different locations. Hence, a change in the number of establishments overestimates the change in the number of companies. Second, our measures is the net change in the number of establishments, not the new creation of establishments, hence it subtracts the number of establishments that were shut off during the year.

Table1 reports the numbers of new business by BEA region year. The grand average is 43, the median 6, and the standard deviation177. There is a large variability across sector. For example, while the average number of new business is -1 in the industry/energy sector it is 113 in the medical/health sector.

We had to construct a mapping from the patent, paper and new business datasets to the VC investment dataset**.** Ideally, to link these datasets we would need the information on all individual projects to determine which paper and patent were relevant to its development. Unfortunately, we do not have access to this data and had to construct an arbitrary link. The correspondence between patent and paper fields on one side and VC investment in the other is showed in Table 6. In some cases (as for the medical sector) the link is quite obvious, in others (like the industrial/energy sector) less so. As for the new business we match with VC investment dataset by looking at the establishment SIC code and try to find a reasonable map to VC sectors. While less than ideal, the noisiness in this mapping bias the results against finding any effect at which is The important point though is that the lack of precise link information introduces a noise in the connection between VC investment, academic research and patenting making it more difficult to establish the relation between these variables. However, it seems to me that it does not bias the result in one way or another.

4- Statistical Models

The dataset we use to test the impact of VC investment on patents have two important characteristics: First, the dependent variable, the total number of citations received by patents invented in the area, is a count variable. Thus, it is an integer bounded by zero. Second, there are many zeros in the sample. The appropriate statistical model to deal with this kind of data features is Poisson regression model presented in Hausman, Hall and Griliches (1984) which applies the model in a similar context, to analyze the relationship between research and development (R&D) expenditures of firms and the number of patents applied for and received by them. Also, Darby and Zucker (2003) applied this model to study the links between entry of firms into the nanotechnology and the strength of the local science base.

To identify the role of VC investment and paper citation on patenting, we implement three different model specifications, which explore different variability of the data. Let n_{ijt} be the number of patents in region i time t, X_{ijt} is the vector of explanatory variables and μ's fixed effects explained below. From the properties of the Poisson distribution we have that E n_{ijt} $=\lambda_{ijt.}$ Then, the first model, which uses as a group a sector/region pair, is

(1) $$prob(n_{ijt}) = \frac{e^{-\lambda_{ijt}} \lambda_{ijt}^{\eta_{ijt}}}{n_{ijt}!} \quad \text{where,} \quad \log\lambda_{it} = \mu_{ij} + X_{ijt}\beta$$

In this model biotech in Los Angeles is one group and biotech in Boston is another. Thus, in this case we use the time variability to identify the parameters.

The second model, which uses as a group a time/region pair, is

(2) $$prob(n_{ijt}) = \frac{e^{-\lambda_{ijt}} \lambda_{ijt}^{\eta_{ijt}}}{n_{ijt}!} \quad \text{where,} \quad \log\lambda_{it} = \mu_{it} + X_{ijt}\beta$$

In this model the Los Angeles area in 1992 is one group and the Boston area in 1992 another. Thus, we use the sector variability for identification.

The third model, where we employ time, region and sector fixed effects separately, is

(3) $$prob(n_{ijt}) = \frac{e^{-\lambda_{ijt}} \lambda_{ijt}^{\eta_{ijt}}}{n_{ijt}!} \quad \text{where,} \quad \log\lambda_{it} = \mu_{i} + \gamma_{j} + X_{ijt}\beta$$

In this model we have an effect for being in Boston, one for being in biotech, and one for being in 1992, but none of these is interacted. Hence, we use both the sector and the time variability for identification.

To control for differences in the size of different BEA regions, we divide both the number of papers and the VC investment by the population living present in the area.

The advantage of using all three models is to check the robustness of the results. Given the different amount of variability we have through time and across sectors, it is useful to know whether the results are driven from one or the other or both.

To test whether the results are driven by a possible reverse causality between patents and VC investments, we will use the total assets of local and state pension funds as an instrument for VC investment. The idea is that the pension funds produce a shift in the local supply of VC investments that help identify the effect of VC on patents. In this case, we use a two stage least square regression where in the first stage we run a linear panel regression of VC investment on pension funds assets and time, region and sector fixed effects. The second stage is the above Poisson model where we substitute VC investment by its predicted value from the first stage. We correct the standard errors by bootstrapping them.

As for the model used to test the impact of VC and patents on the creation of new business we have a different problem. VC investment is likely to affect the creation of new business with a lag. But our dataset has a very short time series (nine years). With such a short period it would be very difficult to reliably estimate this lag. Thus, we are not able to explore the time series variability to identify the parameters. As a result, we collapse the time dimension by averaging the changes in establishments over the nine years and then use only the region and sector variability to estimate the parameters. So the model used is:

$$\text{Number New Business}_{ij} = \mu_i + \gamma_j + \beta_1 \text{Patents}_{ij} + \beta_2 \text{Papers} + \beta_2 \text{ VC investment} + \varepsilon_{ij}$$

Notice that we cannot use the Poisson model here because the dependent variable, number of new business, takes negative values.

5- Results

5.1 Effects of VC and University research on patents

In Table 7 we present the results for the three models present in the previous section. In column I we regress the citation-weighted number of patents in each BEA-sector-year on the number of citations of academic papers in that BEA-sector-year and the total 1982 dollar value of VC investment in that BEA-sector-year. In the regression we control for sector-region fixed effects and for time fixed effects. In other terms, we allow each region to have a specific mean in each sector and we are really exploiting only the time series variation in these values. As column I shows, both paper citations and VC investment have a positive and statistically significant effect on the citation-weighted number of patents.

Column II runs the same regression, with sector fixed effects and region-time fixed effects. So here we allow the Boston area to have a different mean each period, but not to differ across sectors (except for a general specific sector mean). In other words, in this specification we are exploiting the cross industry variability. The results are very similar also from a quantitative point of view. .

Finally, in column III we insert sector, region, and time fixed effect, but no sector-region or time-region fixed effects. Hence, here we are using both the time series and the cross section to identify the effects. Once again the results are very similar.

These results suggest that VC investments have a direct impact on innovation even after controlling for the availability of new ideas, which is proxied here by the number of citations of academic papers. This result is in contrast with the findings of Darby and Zucker (2003), who find that the VC investments do not add any additional explanatory power to the quality of the scientific contribution in the nano biotech sector. The difference in the results might be due to the specificity of the nanobiotech sector or to the larger size of our dataset, which allows us to identify the effects more precisely.

In terms of economic magnitude, the impact of VC on patents is very significant. By the coefficients estimates from Table 7, a one standard deviations increase in VC investment per capita will produce an increase from 4 to 15% on the number of patents depending on the model used.

Given the high concentration of research and venture capital activity in two areas (Boston and the San Francisco area) in Table 8, we check the robustness of our result to excluding these two areas. With exception of the first model where VC investment is not significant, the overall picture is the same, both paper citation and VC investment are important to explain innovation. Also, the coefficients are about the same magnitude as those of Table 7.

One concern with our specification is that papers that produce patents tend to be more cited. Hence, the causality could go from patents to papers and not from papers to patents. To address this concern, in Table 9 we run the same regression using the raw number of papers, instead of paper citations, as explanatory variable. The results are very similar to those in Table 7, and hence they do not seem to be caused by this reverse causality.

To address the possibility of reverse causality for VC investments, in Table 10 we use instrumental variables. As an instrument for VC investment we use the total assets of local and state pension funds. Most states have laws mandating a minimum amount of local investments for every state pension fund. Consistent with this assumption, in Panel A we see that the instrument is statistically significant in the first stage regression. When we use it as an instrument in Panel B, we find that the VC investment remains significant.

5.2 Effects of VC and University research on new businesses

Is this effect of VC and university research limited to the number of patents, or does it translate also in more new ventures? In Table 11 we try to answer this question. Given the noisiness of this data, we aggregate it over time. So In Table 11 we exploit only the sector and region variability of the data.

As the first three columns of Table 11 show, VC investment, number of patents, and paper citation are individually statistically significant in explaining the creation of new businesses. When both patents and VC investments are inserted, the effect of patents is reduced. This might suggest that, consistent with Kortum and Lerner (2000), the VC are facilitating the creation of patents and thus the number of patents is partly proxying for the presence of VC. When this variable is inserted directly, the estimated effect of number of patents drops.

When all three variables are inserted (column 6), the coefficient on the patent variables decreases further, but all three variables are statistically significant. That the VC investment coefficient remains substantially unchanged suggest that the presence of VC continue to exert an important influence on the creation of new business even after controlling for the influence of ideas coming from university research (paper citation variable) and the innovations from the private sector (number of patent).

In order to have a sense of the economic magnitudes of these coefficients we can look at an estimate of the elasticity of the new business with respect to the VC investment. This is estimated by $\eta_{vc} = b_{vc}\frac{\overline{VCinvestement}}{\overline{NewBu\sin ess}}$, that is, the estimated coefficient for VC investment multiplied by the sample average of the variables. The sample estimate will converge in probability to the true elasticity. The estimated elasticity in our case is .25 which means that at the average value of the data an increase of 10% in the volume of VC investment would produce an increase of 2.5% in the total number of new business, given that the other variables remain constant.

6- Conclusions

This paper tries to assess the contribution of venture capital to the promotion of innovation and the creation of new businesses. The problem in identifying this contribution is to disentangle it from other factors that might drive both. We control for omitted factors related to the supply of high-tech ideas by using the production of academic papers in different scientific areas in different US regions. Furthermore, we control for endogenous effects by using the total assets of pension funds in a region as an instrument for VC investment. The idea here is that the presence of pension funds in an area tends to increase the supply of fund to VC because a small fraction of their assets are mandated to be invested in local projects.

The results support the idea that VC has a positive impact on innovation and the creation of new companies. For example, a one standard deviation increase in VC investment per capita will produce a 4 to 15% increase in the number of patents. Similarly, a 10% in VC investment increases the total number of new business by 2.5%.

These results seem to provide *prima facie* evidence of the local benefit of the presence of a venture capital industry. Before we jump to concluding that these results justify subsidization of this industry, however, two further aspects should be analyzed. First, what is the actual cost of promoting VC and what is its effectiveness? Second, our paper also shows that university research produces innovation and new businesses. If we want to promote theses two goals, is it more cost effective to do it by subsidizing VC or by subsidizing university research? Further research is needed to answer these important questions.

References

Adams, James D. (2002), .Comparative Localization of Academic and Industrial Spillovers,.Journal of Economic Geography, 2(3): 253-278.

Andersson, Roland, John M. Quigley, and Mats Wilhelmsson, (2004), .University Decentralization as Regional Policy: The Swedish Experiment,. Journal of Economic Geography, 4(4): 371-388.

Anselin, L., Varga, A. and Acs, Z. (1997), .Local Geographic Spillovers between University Research and High Technology Innovations,.Journal of Urban Economics, 42.

Audretsch D., Lehmman, E. and Warning, S. (2003). Universities Spillovers: Strategic Location and New Firms Performance. Working Paper.

Darby, M. and Lynne Zucker (2003). "Grilichesian Breakthroughs: inventions of methods of Inventing and firm entry in nanotechnology." NBER working paper 9825.

Florax, R. Folmer, H. (1992). Knowledge Impact of Universities on Industries: an Aggregate Simultaneous Investment Model. Journal of Regional Science, Vol. 32, No. 4, pp 437-466.

Hall, Bronwyn H., Adam B. Jaffe and Manuel Trajtenberg (2000). "The NBER Patent Citations Data File: Lessons, Insights and Methodological Tools." NBER unpublished paper.

Hellman, Thomas and Manju Puri (2000). “The Interaction between Product Market and Financing Strategy: The Role of Venture Capital.” Review of Financial Studies 13, 959-984.

Hellman, Thomas and Manju Puri (2002).” Venture Capital and Profissionalization of Start-ups: Empirical Evidence, Journal of Finance 57, 169-197.
Jaffe, A. (1989), .Real Effects of Academic Research. American Economic Review, 79:957-970.

Johnson, Kenneth P. and John R. Kort (2004). "2004 Redefinition of the BEA Economic Areas." Bureau of Economic Analysis technical report.

Kortum, S. and Josh Lerner (2000) "Assessing the Contribution of Venture Capital to Innovation." RAND Journal of Economics, vol 31 p.674-692.

Lerner, J., 1995. Venture capitalists and the oversight of private firms. Journal of Finance 50, 301-318.

Ueda, M. and Masayuki Hirukawa. "Venture Capital and Productivity." UPF Economics and Business Working Paper No. 522.

Wasmer, Etienne and Philipe Weil (2000). “The Macroeconomics of Labor and Credit imperfections.” Bonn: IZA DP 179.

Table 1 – Summary Statistics

VC investment is from Venture Economics and is computed as the sum of investment of VC funds in firms located in each Bureau of Economic Analysis region expressed in 1982 dollars. Number of Patents is from the NBER patent database and is the number of citations to U.S. patents not assigned to publicly traded companies between 1992 and 1999. Public pension assets is from the State and Local Government Employee-Retirement System survey. It is computed as the sum of State and Local Government pensions assets in 1982 dollars. New businesses is the change in the number of establishments from the US Census Bureau County Business Pattern dataset. Citations of papers is from the Thomson Scientific Science Indicator, which measures 284 research universities in the United States. The variable used is the number of times papers written by people on the faculty at institutions located in the BEA area were cited through 2003. In order to mitigate the truncation problem, the total number of citations is divided by the time the paper could have received citations. All the variables are fro the period 1992-1999.

Sector	Variable	mean	median	sd	min	max
Biotechnology	VC investment (in Million US$)	4.97	0.00	27.80	0.00	428.84
	Number of patents	12.35	0.00	48.44	0.00	578.00
	Pension assets (in Million US$)	5,857.31	2,284.64	11,800.00	15.95	130,000.00
	New businesses	3.30	1.00	12.25	-17.00	139.00
	Paper citations	1,085.46	80.56	2,768.47	0.00	23,449.86
Communications and Media	VC investment (in Million US$)	11.76	0.00	52.83	0.00	955.79
	Number of patents	17.00	1.00	59.14	0.00	691.00
	Pension assets (in Million US$)	5,857.31	2,284.64	11,800.00	15.95	130,000.00
	New businesses	21.77	8.00	67.52	-170.00	1,104.00
	Paper citations	4.82	0.00	14.73	0.00	157.78
Computer	VC investment (in Million US$)	19.39	0.00	126.76	0.00	2,573.82
	Number of patents	22.99	0.00	104.04	0.00	1,638.00
	Pension assets (in Million US$)	5,857.31	2,284.64	11,800.00	15.95	130,000.00
	New businesses	101.93	18.00	302.71	-18.00	4,924.00
	Paper citations	8.40	0.13	23.43	0.00	201.40
Industrial/Energy	VC investment (in Million US$)	7.61	0.00	47.13	0.00	968.18
	Number of patents	77.32	18.00	187.31	0.00	1,901.00
	Pension assets (in Million US$)	5,857.31	2,284.64	11,800.00	15.95	130,000.00
	New businesses	-1.42	0.00	43.79	-409.00	553.00
	Paper citations	335.68	15.00	841.65	0.00	6,337.00
Medical/Health	VC investment (in Million US$)	10.69	0.00	42.70	0.00	601.08
	Number of patents	42.81	5.00	136.39	0.00	1,363.00
	Pension assets (in Million US$)	5,857.31	2,284.64	11,800.00	15.95	130,000.00
	New businesses	112.99	47.00	270.90	-875.00	3,376.00
	Paper citations	1,768.50	23.00	4,515.03	0.00	38,646.57
Semiconductors/ Other Elect.	VC investment (in Million US$)	3.64	0.00	28.42	0.00	663.32
	Number of patents	21.27	2.00	71.47	0.00	733.00
	Pension assets (in Million US$)	5,857.31	2,284.64	11,800.00	15.95	130,000.00
	New businesses	17.24	6.00	54.63	-431.00	590.00
	Paper citations	15.06	0.14	40.18	0.00	323.11
Total	VC investment (in Million US$)	9.68	0.00	64.06	0.00	2,573.82
	Number of patents	32.29	2.00	114.29	0.00	1,901.00
	Pension assets (in Million US$)	5,857.31	2,284.64	11,800.00	15.95	130,000.00
	New businesses	42.77	6.00	177.06	-875.00	4,924.00
	Paper citations	536.32	0.72	2,288.64	0.00	38,646.57

Table 2- Number of Papers in the Top 15 BEA Regions in 1998

Citations of papers is from the Thomson Scientific Science Indicator, which measures 284 research universities in the United States. The variable used is the number of times papers written by people on the faculty at institutions located in the BEA area were cited through 2003. In order to mitigate the truncation problem, the total number of citations is divided by the time the paper could have received citations.

Sectors / Areas	Biotechnology	Communications and Media	Computer	Industrial/Energy	Medical/Health	Semiconductors/ Other Elect.	Total
Boston-Worcester-Manchester	4,903	74	173	2,950	9,031	209	17,340
New York-Newark-Bridgeport	5,152	81	172	2,770	8,750	176	17,101
Washington-Baltimore-Northern Virginia	2,554	70	181	1,997	6,405	209	11,416
San Jose-San Francisco-Oakland	2,950	68	117	2,114	4,514	214	9,977
Los Angeles-Long Beach-Riverside	2,750	70	112	2,499	4,210	254	9,895
Houston-Baytown-Huntsville	2,498	16	39	890	6,027	69	9,539
Seattle-Tacoma-Olympia	1,772	14	66	804	4,605	96	7,357
Detroit-Warren-Flint	1,639	16	64	1,218	3,176	117	6,230
Philadelphia-Camden-Vineland	1,656	14	33	661	3,677	33	6,074
Minneapolis-St. Paul-St. Cloud	1,230	10	41	485	3,950	51	5,767
Raleigh-Durham-Cary	1,569	15	45	838	3,200	93	5,760
Denver-Aurora-Boulder	1,368	7	40	715	2,640	69	4,839
Indianapolis-Anderson-Columbus	1,422	21	65	811	2,281	48	4,648
Chicago-Naperville-Michigan City	1,095	9	36	645	2,406	48	4,239
San Diego-Carlsbad-San Marcos	1,497	27	27	682	1,636	69	3,938
Total	34,055	512	1,211	20,079	66,508	1,755	124,120

Table 3- Frequency of Investment by BEA Areas

Each cell reports the number of BEA regions that in that period had a total VC investment in the range indicated on the left hand side. VC investment is from Venture Economics and is computed as the sum of investment of VC funds in firms located in each Bureau of Economic Analysis region expressed in 1982 dollars

	Period			
Investment Range (M US$)	**1981-1985**	**1986-1990**	**1991-1995**	**1996-1999**
0	74	64	63	41
0-50	79	75	74	64
50-100	5	13	13	13
100-200	11	7	8	16
200-500	6	10	5	15
500-1,000	1	4	11	8
1,000-5,000	3	6	4	17
>5,000	0	0	1	5
Total	179	179	179	179

Table 4- Investment of VC Funds by Sectors

Fraction of VC investments in each sector in each period. VC investment is from Venture Economics and is computed as the sum of investment of VC funds in firms located in each Bureau of Economic Analysis region expressed in 1982 dollars

	Period					
Sector	**1980-1984**	**1985-1989**	**1990-1994**	**1995-1999**	**2000-2004**	**Grand Tota**
Biotechnology	0.04	0.04	0.06	0.03	0.04	0.04
Communications and Media	0.13	0.13	0.16	0.15	0.13	0.14
Computer	0.35	0.15	0.12	0.14	0.16	0.15
of which Hardware	0.26	0.08	0.04	0.02	0.02	0.03
of which Software and Services	0.09	0.06	0.09	0.12	0.13	0.12
Industrial/Energy	0.11	0.06	0.10	0.07	0.06	0.07
Medical/Health	0.07	0.08	0.12	0.09	0.07	0.08
Semiconductors/Other Elect.	0.10	0.06	0.04	0.04	0.07	0.06
Tech Sectors:	0.80	0.52	0.60	0.71	0.74	0.71
Internet Specific	0.00	0.00	0.01	0.19	0.20	0.17
Consumer Related	0.11	0.25	0.17	0.13	0.08	0.11
Other Products	0.09	0.23	0.23	0.16	0.18	0.18
Non Tech Sectors:	0.20	0.48	0.41	0.48	0.46	0.46

Table 5- VC investment in the Top 15 BEA Regions in 1998

Sectors / Areas	Biotechnology	Communications and Media	Computer	Industrial/Energy	Medical/Health	Semiconductors/ Other Elect.	Total
San Jose-San Francisco-Oakland	349,237	1,162,226	3,594,666	55,908	515,398	541,020	6,218,455
Boston-Worcester-Manchester	354,852	1,102,286	1,208,376	126,507	232,968	51,587	3,076,576
St. Louis-St. Charles-Farmington	-	2,005,250	266,349	521,881	36,600	4,050	2,834,130
New York-Newark-Bridgeport	120,367	391,766	1,098,576	77,358	552,835	93,309	2,334,211
Los Angeles-Long Beach-Riverside	76,849	74,344	772,295	400,769	562,327	150,369	2,036,953
San Antonio	16,906	687,957	-	-	650,000	-	1,354,863
Washington-Baltimore-Northern Virginia	10,595	619,690	529,387	3,148	106,432	31,364	1,300,616
Dallas-Fort Worth	-	229,146	143,879	503,297	107,894	16,380	1,000,596
Denver-Aurora-Boulder	47,913	292,379	460,973	9,214	12,143	33,216	855,838
Seattle-Tacoma-Olympia	52,073	276,954	370,503	64,100	44,955	2,800	811,385
San Diego-Carlsbad-San Marcos	148,936	66,145	189,623	14,300	209,450	106,105	734,559
Houston-Baytown-Huntsville	44,891	13,789	66,245	448,845	99,550	2,000	675,320
Philadelphia-Camden-Vineland	65,888	196,500	146,127	80,500	84,074	32,853	605,942
Atlanta-Sandy Springs-Gainesville	12,872	77,498	393,247	2,000	64,704	4,400	554,721
Chicago-Naperville-Michigan City	-	111,967	181,352	96,730	157,843	6,500	554,392
Total	1,301,379	7,307,897	9,421,598	2,404,557	3,437,173	1,075,953	24,948,557

Table 6- Fields Links Across Datasets

VC	Patent	Papers
Biotechnology	Biotechnology	Biochemistry & Biophysics Biology, General Biotechnology & Applied Microbiology Endocrinology, Nutrition & Metabolism Experimental Biology Physiology Chemical Engineering Chemistry & Analysis Chemistry Inorganic & Nuclear Chemistry Organic Chemistry / Polymer Science Physical Chemistry / Chemical Physics Microbiology Cell & Developmental Biology Molecular Biology & Genetics
Communications and Media	Communications	Information Technology & Communications Systems
Computer Hardware and Software	Computer Hardware & Software Computer Peripherials Information Storage	Computer Science & Engineering
Industrial/Energy	Heating Pipes & Joints Coating Gas Organic Compounds Resins Miscellaneous-chemical Power Systems	AI, Robotics & Automatic Control Civil Engineering Environmental Engineering / Energy Mechanical Engineering Nuclear Engineering Geological, Petroleum & Mining Engineering Materials Science & Engineering Metallurgy Applied Physics / Condensed Matter / Materials Science Optics & Acoustics Physics Metal Working Miscellaneous-Mechanical
Semiconductors/Other Elect.	Receptacles Electrical Devices Electrical Lighting Semiconductor Devices Miscellaneous-Elec	Electrical & Electronics Engineering

Table 6- Field Links Across Datasets (continue)

Internet Specific	**No Link**	**No Link**
Medical/Health	Drugs Surgery & Med Inst. Miscellaneous-Drgs&Med	Anesthesia & Intensive Care Cardiovascular & Hematology Research Cardiovascular & Respiratory Systems Clinical Immunology & Infectious Disease Clinical Psychology & Psychiatry Dentistry / Oral Surgery & Medicine Dermatology Endocrinology, Metabolism & Nutrition Environmental Medicine & Public Health Gastroenterology & Hepatology General & Internal Medicine Health Care Sciences & Services Hematology Medical Research, Diagnosis & Treatment Medical Research, General Topics Medical Research, Organs & Systems Neurology Oncogenesis & Cancer Research Oncology Ophthalmology Orthopedics & Sports Medicine Otolaryngology Pediatrics Pharmacology/Toxicology Radiology, Nuclear Medicine & Imaging Reproductive Medicine Research/Lab Medicine & Medical Technology Rheumatology Surgery Urology & Nephrology Immunology Pharmacology & Toxicology
Other Products	No Link	No Link
Semiconductors/Other Elect.	Receptacles Electrical Devices Electrical Lighting Semiconductor Devices Miscellaneous-Elec	Electrical & Electronics Engineering

Table 7 - Poisson Regressions – Complete Sample

The dependent variable is number of citations on patents from the NBER patent dataset, where we eliminate patents assigned to companies with a COMPUSTAT code to isolate patents from non listed companies. The models are Poisson panels: prob(nit)=exp(λit).λit/(nit!) where, Model 1: logλit = μij + Xijtβ ; Model 2: logλit = μit + Xijtβ; Model 3: logλit = μi + γj + Xijtβ. VC investment is from Venture Economics and gives the total investment of VC funds in a particular region. Citations of papers is from the Thomson Scientific Science Indicator. The variable used is the number of times the papers were cited through 2003. In order to mitigate the truncation problem, the total number of citation is divided by the time the paper could have received citations. The sample period is 1982-1998 and includes all 179 BEA regions and 6 tech sectors. All variables are normalized by the respective BEA population. All standard deviations are bootstrapped. * indicates statistical significant at 10%, ** statistical significant at 5%, and *** statistical significant at 1%

	Model 1	**Model 2**	**Model 3**
Citation of Papers	2.43	1.40	1.87
	(0.43)***	(0.12)***	(0.19)***
VC Investments	2.83	3.40	5.05
	(1.23)**	(0.57)***	(0.84)***
Sector Fixed Effects	yes	yes	yes
Region Fixed Effects	no	no	yes
Time Fixed Effects	no	no	yes
Sector/Region Fixed Effects	yes	no	no
Region/Time Fixed Effects	no	Yes	no
Number obs.	17,085	17,436	18,258
Number of Groups	1,005	2,906	179
Log likelihood	-41,460	-48,689	-64,018

Table 8 - Poisson Regressions
Excluding Boston and San Francisco Areas

The dependent variable is number of citations on patents from the NBER patent dataset, where we eliminate patents assigned to companies with a COMPUSTAT code to isolate patents from non listed companies. The models are Poisson panels: prob(nit)=exp(λit).λit/(nit!) where, Model 1: $\log\lambda_{it} = \mu_{ij} + X_{ijt}\beta$; Model 2: $\log\lambda_{it} = \mu_{it} + X_{ijt}\beta$; Model 3: $\log\lambda_{it} = \mu_i + \gamma_j + X_{ijt}\beta$. VC investment is from Venture Economics and gives the total investment of VC funds in a particular region. Citations of papers is from the Thomson Scientific Science Indicator. The variable used is the number of times the papers were cited through 2003. In order to mitigate the truncation problem, the total number of citation is divided by the time the paper could have received citations. The sample period is 1982-1998 and includes all 179 BEA regions except Boston and San Francisco and 6 tech sectors. All variables are normalized by the respective BEA population. All standard deviations are bootstrapped. * indicates statistical significant at 10%, ** statistical significant at 5%, and *** statistical significant at 1%

	Model 1	**Model 2**	**Model 3**
Citation of Papers	2.52	1.68	2.25
	(0.59)***	(0.13)***	(0.76)***
VC Investments	0.97	5.56	2.28
	(0.77)	(2.55)**	(0.76)***
Sector Fixed Effects	Yes	yes	yes
Region Fixed Effects	no	no	yes
Time Fixed Effects	no	no	yes
Sector/Region Fixed Effects	Yes	no	no
Region/Time Fixed Effects	no	Yes	no
Number obs.	16,881	17,232	18,054
Number of Groups	993	2,872	179
Log likelihood	-38,699	-145,461	-57,486

Table 9 – Poisson Regressions –
Complete Sample using Number of Papers instead of Paper Citation

The dependent variable is number of citations on patents from the NBER patent dataset, where we eliminate patents assigned to companies with a COMPUSTAT code to isolate patents from non listed companies. The models are Poisson panels: prob(nit)=exp(λit).λit/(nit!) where, Model 1: $\log\lambda_{it} = \mu_{ij} + X_{ijt}\beta$; Model 2: $\log\lambda_{it} = \mu_{it} + X_{ijt}\beta$; Model 3: $\log\lambda_{it} = \mu_i + \gamma_j + X_{ijt}\beta$. VC investment is from Venture Economics and gives the total investment of VC funds in a particular region. Papers is from the Thomson Scientific Science Indicator. The variable used is the number of times the papers were cited through 2003. In order to mitigate the truncation problem, the total number of citation is divided by the time the paper could have received citations. The sample period is 1982-1998 and includes all 179 BEA regions except Boston and San Francisco and 6 tech sectors. All variables are normalized by the respective BEA population. All standard deviations are bootstrapped. * indicates statistical significant at 10%, ** statistical significant at 5%, and *** statistical significant at 1%

	Model 1	**Model 2**	**Model 3**
Number of Papers	4.92	8.31	4.47
	(5.1)	(0.35)***	(0.55)***
VC Investments	2.85	3.81	5.09
	(1.28)**	(0.69)***	(0.82)***
Sector Fixed Effects	Yes	Yes	yes
Region Fixed Effects	no	No	yes
Time Fixed Effects	no	No	yes
Sector/Region Fixed Effects	Yes	No	no
Region/Time Fixed Effects	no	Yes	no
Number obs.	17,085	17,436	18,258
Number of Groups	1,005	2,906	179
Log likelihood	-42,049	-151,325	-64,864

Table 10 – Instrumenting VC investment with Pension Funds Assets

The first stage is a linear panel of VC investment on Pension funds assets with region and sector fixed effects (robust standard deviations in parenthesis). The second stage is a Poisson panel of the predicted value from the first stage on the citation of papers (bootstrapped standard deviations in parenthesis). Pension fund assets are the total assets of Local and State Pension Funds from the US Census Bureau as an instrument for VC investment expressed in 1982 $. VC investment is from Venture Economics and gives the total investment of VC funds in a particular region. Paper citation is from the Thomson Scientific Science Indicator. The variable used is the number of times the papers were cited through 2003. In order to mitigate the truncation problem, the total number of citation is divided by the time the paper could have received citations. The sample period is 1982-1998 and includes all 179 BEA regions except Boston and San Francisco and 6 tech sectors. All variables are normalized by the respective BEA population. All standard deviations are bootstrapped. * indicates statistical significant at 10%, ** statistical significant at 5%, and *** statistical significant at 1%

1st Stage - Dependent Variable: VC Investment (M US$)

Independent Variables	
Total Real Assets of Local and State Pension Funds (by BEA population)	0.1
	(0.03)***
Region/Sector Fixed Effects	Yes
Time Fixed Effects	Yes

2nd Stage - Dependent Variable: Patents

Independent Variables	No Instrument	With Instrument
Citation of Papers (by BEA population)	2.43	2.96
	(0.43)***	(0.75)***
VC Investment	2.83	23.43
	(1.23)**	(9.12)***
Region/Sector Fixed Effects	Yes	Yes
Time Fixed Effects	Yes	Yes
Number obs.	5370	4485

Table 11 – OLS Estimation Results for New Businesses

The model is a linear fixed effect panel: Number New Businessesij= μi + γj + β1Patentsij + β2Papers + β2 VC investment + εij. The data for new businesses is the change in the number of establishments from the US Census Bureau County Business Pattern dataset. VC investment is from Venture Economics and gives the total investment of VC funds and firms in a particular region deflated by the Producer Price Index (PPI). Paper citations is from the Thomson Scientific Science Indicator. The variable used is the number of times the papers were cited through 2003. In order to mitigate the truncation problem, the total number of citation is divided by the time the paper could have received citations. Number of citations on patents is from the NBER patent dataset, where we eliminate patents assigned to companies with a COMPUSTAT code to isolate patents from non listed companies. The sample period includes all 179 BEA and 6 tech sectors. We collapse all variables along the region and sector dimensions by taking time series average on the sample 1988-1997.All variables are normalized by the respective BEA population. All standard deviations are bootstrapped. * indicates statistical significant at 10%, ** statistical significant at 5%, and *** statistical significant at 1%

	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
VC Investments		49.19			43.95	44.87
		(7.95)***			(7.95)	(7.91)***
Citation of Papers			18.19			17.55
			(9.49)**			(9.18)*
Number of Patents				66.00	40.94	37.08
				(20.69)***	(16.82)**	(16.60)**
Sector Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Region Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Number obs.	1074	1074	1074	1074	1074	1074
R^2	0.48	0.53	0.48	0.50	0.53	0.54

Figure 1 – VC Investment by BEA Regions (1981-1999)

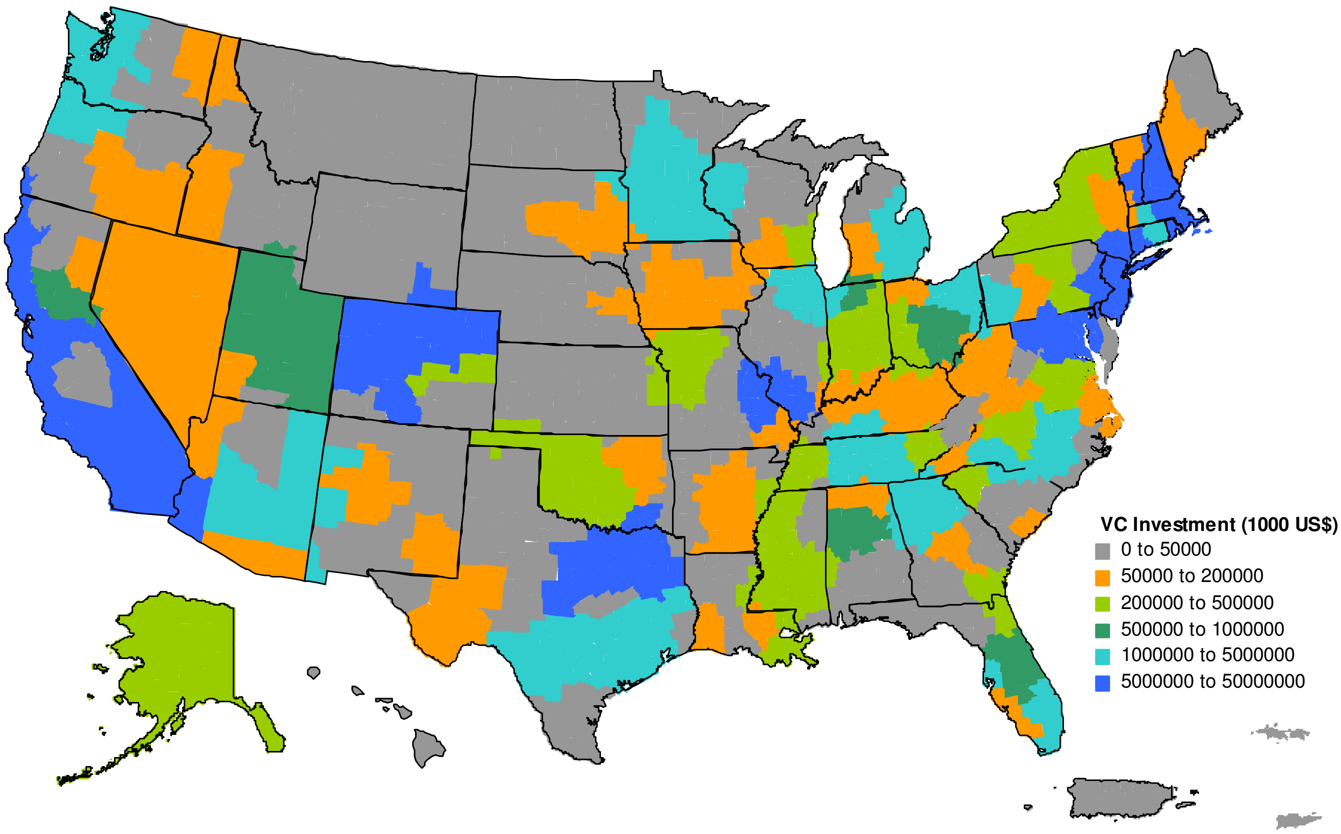

Figure 2 – Number of Patents by BEA Regions (1981-1999)

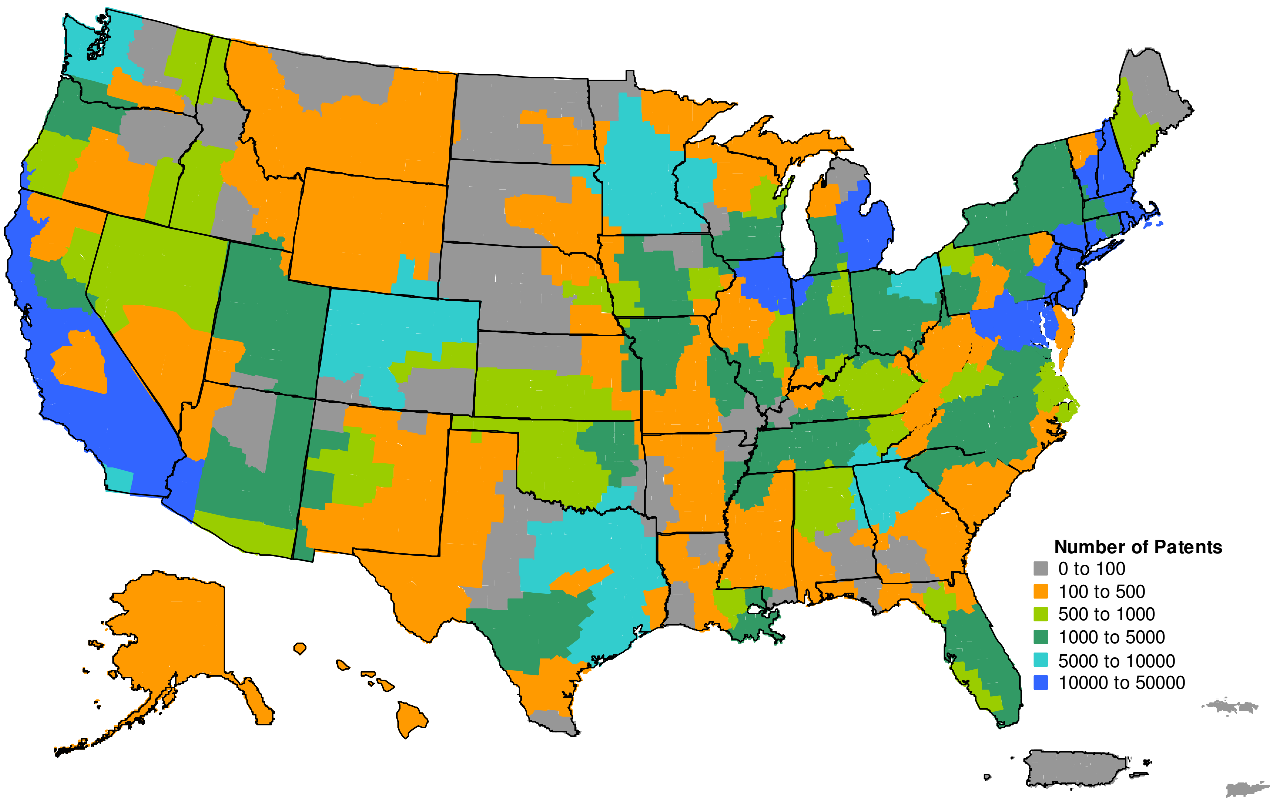

Figure 3 – Paper Citations by BEA Regions (1981-1999)