THE UNIVERSITY OF CHICAGO

ESSAYS on VENTURE CAPITAL INVESTMENT

A DISSERTATION SUBMITED TO THE FACULTY OF THE

DIVISION OF THE SOCIAL SCIENCES IN CANDIDACY FOR THE

DEGREE OF DOCTOR OF PHILOSOPHY

DEPARTMENT OF ECONOMICS

By

Marcos A. Mollica

CHICAGO, ILLINOIS

JUNE 2006

TABLE OF CONTENTS

List of Tables

List of Figures

ACKNOWLEDGEMENTS

I would like to thank my advisor, Luigi Zingales for the patient discussions and suggestions on fundamental points of this dissertation. Thanks also to Chad Syverson and Steven Kaplan for the suggestions given on my dissertation proposal. I also would like to thank all the Brazilian friends in Chicago for making the living here more pleasant. Above all, thanks to my wife Andrea for the unconditional support along this long project, without her I would not have been able to overcome this challenge.

ABSTRACT

This dissertation is composed of two parts. In the first part, the purpose of this is to explain what drives the Venture Capital investment in a particular area. The existing literature usually addresses this question by using cross country data. In contrast, I use as the geographic unit US regions. The advantage of this approach is that it holds fixed many of the determinants of VC investment identified in the cross country studies such as volume of IPOs, stock market capitalization and accounting standards. Therefore, it is possible to better measure the influence of the availability of new ideas, in particular, the role of academic research. The results indicate that the most important variable to explain the variability of VC investment in the US regions is the quality of academic production measured by the total citation of papers. The results are robust to the exclusion of the areas with largest presence of VC (Boston and San Francisco) and to the use of universities funds to R&D investment as an instrument to the quality of research variable. In the second part, the objective is to test the proposition that VC investment have a significant positive impact on innovations and on the creation of new businesses. The central problem is to disentangle the effect of the exogenous component of the VC investment on the number of patents (our proxy for innovation) and on the number of new companies using those patents. I address these problems in two ways. First, to control for omitted factors related to the supply of high-tech ideas I will use the production of academic papers in different scientific areas in different US regions. Second, I will use the time series and cross sectional difference in the quality of academic production to identify the effect of VCs on patents. Finally, I use the region and sector variability of the data to access the impact of VC investment, patent and academic production on new businesses. This is an important step because merely finding a positive effect on patents may only indicate VCs higher propensity to patent. The results suggest that VC investment has a significant positive effect on both patents and the creation of new businesses.

Chapter 1- Determinants of VC Investments

1.1- Introduction

The American venture capital industry is often cited as an effective way to promote entrepreneurship. Examples of venture capital (VC) financed companies like Google, Yahoo and Microsoft usually are used to stress the relevance of the VC. Indeed, many studies present evidence that VC industry enhance the ability of new firms to create wealth and jobs. For instance, VC backed firms produce more and more valuable patents Kortum and Lerner (2000), they are faster in developing products and introducing them in the market and they have a higher rate of executive turnover, Hellman and Puri (2000, 2001). Furthermore, the presence of VC sector seems to produce superior macroeconomic performance, Wasmer and Will (2000).

Inspired on the successes of these new companies, many other countries have tried to mimic the American structure in the hope of busting its own economic development. Yet it seems that no country has been able to develop a venture capital sector as successful as the one present in United States, the study of Bottazi and Da Rin (2001) suggests that the substantial growth of European VC in terms of volume invested did not translate into growth and employment of the portfolio companies.

Given the potential importance of the VC activity, it is crucial to understand the main determinants of VC investment. What explains the presence of VC investment in a particular region and time? Some studies try to answer this question by analyzing a cross section of countries. The main conclusion of the cross countries studies is that financial variables such as IPO's, stock market capitalization, and account standards are the most important variables to explain VC investment. While these studies are able to uncover some important determinants of the VC investment, the credibility of the statistical results may be questioned. This because the heterogeneity in a cross country study is so large that it becomes very difficult to control for all the unobservables which could potentially be correlated with the explanatory variables. Thus, there is a large chance that the regressions are producing biased estimators. Differences in institutions, level of government intervention in the economy and even the definition of what constitutes Venture Capital financing may play a role in contaminating the statistical results.

This paper investigates the determinants of VC investment using as geographic units proper defined regions in the United States. By focusing in one particular country, I hope to diminish the problem of the presence of unobservables into the statistical model. Furthermore, the financial variables stressed in the cross countries studies will have no importance to explain the cross section variability in VC investment here because they tend to be homogeneous inside the country. Thus, I can isolate the influence of other variables linked to the availability of new ideas.

The idea to analyze typically cross country problems using US regions as units is not new. For example, Barro (1992) use data on US states to test the hypothesis of convergence in the Neoclassical Growth Model. In the Venture Capital research, the only paper I'm aware of that tries to exploit region variability to explain VC investment is Lerner and Gompers (1999). That paper focuses in US states and one of the explanatory variables used is the amount expend in R&D by universities and private companies.

The present paper uses a different regional unit, the Bureau of Economic Analysis (BEA) economic areas which are composed by counties surrounding metropolitan areas. There are a total of 179 BEA areas. Given the monitoring and advisory rule of the VC, it is a well documented fact that the VC firms tend to invest in projects close to its geographical base, Gupta and Sapienza (1992) suggests that VC exhibit highly localized investment patterns. By using more restricted defined regional units, it is possible to have a more precise idea of what are the regional variables really influencing the VC investment. In addition, the large amount of geographical units increases substantially the degrees of freedom of the statistical model and the precision of the model fixed effects. Another innovation in this work is that I use a variable based on the citation of academic papers to capture the effect of the quality of research into the VC investment. Besides, the paper combines two other datasets: the NBER U.S. Patent Citations Data File and the Investment of Venture Capital firms by Venture Economics. The idea of the paper is to explore the time, regional and sector variability of patents and paper citation to explain the variability in the VC investment.

The observed amount of investment is a result of the interaction between supply and demand of VC funds. In this paper, I start by making the simplifying assumption of a perfect capital market where an infinite amount of funds flows instantaneously to capture excess returns. Basically, it is assumed that the supply of funds is infinitely elastic and the equilibrium investment is determined exclusively by demand factors. Later, I abandon this assumption by considering the volume of pension funds investment in a particular region as a supply shifter. The idea is that, for political reasons, the local and state funds tend to have more investment at the local level. Thus, the total assets of local and state pension funds could be used as an instrument helping to identify the demand for VC investment.

The demand for VC funds is directly linked to the incentives to entrepreneurship. One important factor affecting the decision of a worker to become an entrepreneur is the availability of new ideas in a particular region and time. The story of a scientist faced with an innovation with potential commercial application but with no managerial or financial resources needed to start a business is very common. Here the VCs play an import role. Not only the VC supplies the finance to the project, but also they have important monitoring and advisory contributions which increase the probability of success of the new business Kaplan and Stromberg (1999, 2001). Thus, it is relevant to understand the importance of the academic quality in attracting VC investment. I try to capture the availability of new innovations coming from academic research by using the number of times the papers were cited through 2003. The quality of research tends to positively influence the demand for entrepreneurship because it increases probability that a researcher will face a commercially viable application.

But innovations don't take place exclusively in the research laboratories. A substantial part of new ideas are born inside the companies. To capture this second source of innovation I use the patent citation data. The idea is that the regions with higher quality inventors will be likely to have more potential entrepreneurs.

I also include a measure of income growth which tries to capture the effect of overall economic condition on the investments. Finally, the capital gain taxation tends to be important in the demand side. In order to obtain a panel data of the capital gain taxation I used the Taxim simulation software. Taxim is the NBER's FORTRAN program for calculating liabilities under US Federal and State income tax laws from individual data.

I use these datasets to construct a multilevel panel from 1980 to 1999 for the 179 BEA regions and 6 sectors of VC investment. To form this last dimension I propose a link between the patent categories and paper fields on one side and the sectors of VC investment on the other. In the end, the reduced form of the statistical model has as the dependent variable the amount of VC investment by year, region and sector. As independent variable is the paper citation, patent citation, capital gain taxation, fixed effects (sector, region and time) and regional economic variables (income growth, employment, etc.…).

The results indicate that the most important variable to explain the variability of VC investment in the US regions is the quality of academic production measured by the total citation of papers. The results are robust to the exclusion of the areas with largest presence of VC (Boston and California) and to the possibility of a positive inclined supply of funds due to the investment policies of the local and state pension funds. The results for the case with an infinitely elastic supply of funds is remarkably similar to the case where the total assets of public local and state pension funds is used to explain the supply of VC investment. Finally, with the limited data available at the international level, it seems that academic production is also an important variable to explain the cross section of countries VC investment.

There is a vast literature which documents the fact that the universities generate positive externalities in the region they are located. In this literature, the process of producing knowledge in the universities creates spillovers effects to the surrounding economies. Just to cite a few examples, Anselin, Varga and Acts (2002) provides evidence of local spatial externalities between university research and high technology innovative activity in the United States both at state level and at the Metropolitan Statistical Areas level. Adams (2003) finds that academic spillovers are more localized than industrial spillovers reflecting industry-academic cooperative movements. Audretsch, Lehmann and Warning (2003), using a dataset of high tech companies from Germany, find that firms have a propensity to locate close to universities and proximity to universities reduces the time between firm formation and listing on the stock market. Florax (1992), using a Netherlands dataset, suggests that academic knowledge production have a significant impact on investment in equipments which is strongest in the neighborhood of central places. Andreson, Quigley and Wilhelmsson (2004) find evidence that Swedish policy of spatial decentralization of post-secondary education had a significant impact on regional development. Jaffe (1989) finds that university research has a significant positive effect on corporate patent and an indirect effect on local innovation by inducing R&D spending. In this context, the present paper presents additional evidence of the spillover effect from the universities: the attraction of venture capital investment.

1.2- Data Description and Methodology

BEA Regions. 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. The division of the United States in these areas is displayed in picture 1 in the appendix.

There are 179 regions in the present division used by the BEA. The construction of these regions involves three phases. First, the economic nodes are identified. These are metropolitan areas or micropolitan areas, according to the definitions of the U.S. Office of Management and Budget for Metropolitan Statistical Areas and Core Based Statistical Areas that serve as the main centers of economic activity. In the end of this part, 344 economic nodes are identified consisting of 1311 counties. Second, each of the remaining 1830 counties are assigned to these economic nodes based on county-to-county commuting flow. Finally, in the third major phase, the 344 CEAs were used as building blocks to form the new 179 BEA economic areas. The two major factors that were used in the aggregation process were commuting patterns and economic and geographic size. The aggregation is such that each economic area is a region of sufficient size to support regional statistical analyses and each economic area is a labor market that is independent of other labor markets. For a more detailed explanation of these areas see Johnson and Kort (2004).

Basically, there are three main reasons to use this region division in this work instead of the usual states division. First, from the statistical point of view the degrees of freedom are substantially increased, the BEA are composed by 179 regions as opposed to 50 states. Second, 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. In this sense, this could potentially increase the reliability of region specific effects. Finally, the Bureau of Economic analysis publishes a large number of economic data for these regions such as personal income, wages and employment which potentially could be used as controls in the empirical model. Figure 1 in the appendix depicts the US map divided by the BEA areas.

Thomson Science Indicator. I use the product University Science Indicators 1981-2003 from Thomson scientific to obtain a measured of scientific production in each region. This product is database of research performance measures on 284research universities in the United States. Based on the publication and citation statistics of ISI® from the same company, University Science Indicators contains the number of ISI-indexed papers from each university and the number of times the papers were cited through 2003. The use of citation data is especially important because it gives a measured of the quality of the scientific production. The variable used in this work is the number of times the papers were cited through 2003. In this database there are 74 fields of research. The variability in the fields is a potentially important source of identification of region specific effects.

Taxsim simulation software. In order to obtain a panel data of the capital gain taxation I used the Taxsim simulation software. Taxsim is the NBER's FORTRAN program for calculating liabilities under US Federal and State income tax laws from individual data. The tax used in this is the maximum tax rate for an additional $1000 of income on an initial $250,000 of wage income. The taxpayer is assumed to be married and filing jointly. No personal deductions are included except for the state income tax on the federal return and the federal tax on the state return (where allowed). This deductibility would be the main difference between these rates and the maximum bracket rate which might be published in a summary of state tax laws. Because the tax rates are only available at the state level, when a BEA area lies in more than one state, I use the average rate over these states.

Venture Capital Investment. The data on 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). In table 1 in the appendix I present fraction of the total investment of each sector. This paper works only with the sectors that could potentially be inked with academic research. These are what I called Tech sectors: Biotechnology, Computer (hardware, software, services), Industrial/Energy, Medical/Health, Semiconductors/Other Elect. The other sectors (non Tech) are: Consumer Related, Internet Specific, and Other Products. These are left out of the study since they don't seem to have a direct link with the academic production. Thus the paper will try to explain a fraction of about 71% of the VC investment. In table 2 I show the frequency of VC investment in particular ranges for some periods in the sample. It can be seen from there that most of the areas have positive investment in all the periods.

Patent Citation Data. The data of patent citations are from the NBER patent citation dataset. For this study it is used the citation received by US patents with application year between 1980 and 1999. These patents are divided in 2 groups: the patents assigned to US non government organizations (mostly corporations according to the NBER). In the US, a granted US patent is assigned to the inventor. If that inventor works for a company, as part of their employment, they assign the patent to their 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.

Pension Funds Data. State and Local Government Employee-Retirements Systems is an annual survey conducted by the Census Bureau. The survey coverage includes public employee retirement systems administered by state and local governments throughout the nation. State and Local Government Employee-Retirements Systems is an annual survey conducted by the Census Bureau. The survey coverage includes public employee retirement systems administered by state and local governments throughout the nation. I use in this paper the total assets of these funds deflated by the PPI. The data is available from 1993 to 1999.

Mapping from the patent and paper datasets to the VC investment dataset. Ideally, to link these datasets I would need the information on all individual projects to determine which paper and patent were relevant to its development. Unfortunately, I don't have access to this data at this point and have 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 3 in the appendix. It can be seen that the link is quite obvious in some cases as in the case of the medical sector. But in other sectors like the industrial/energy sector this is not so clear. 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.

1.3- Demand and Supply for VC Funds

The supply of VC funds is the amount of money available to invest in venture capital funds and firms. Therefore, from the point of view of the investor who puts money in these funds, VC investment is another asset he could use to portfolio diversification purposes. I start by assuming that the supply of funds to venture capital is infinitely elastic and the equilibrium investment is exclusively determined by the demand side. This assumption is related to the existence of assets which are close substitutes to VC funds. Indeed, Cochrane (2005), using a sample corrected for selection bias, shows that the arithmetic mean returns, standard deviations, and arithmetic alphas in venture capital are very similar to the ones of a portfolio composed by the smallest NASDAQ stocks. Potentially, capital gains taxation could be important to explain the expected return as well. However, since the major investors in the VC funds are pension funds which are considered tax exempt, taxation is not likely to be important factor and is not included in the supply side. As explained below, the capital gain will likely to be important to explain the demand for funds.

Later I abandon this assumption of infinitely elastic supply of funds by considering the size of local and state public pension funds as a potential shifter of the supply of funds schedule. The idea is that these funds, by political pressure, tend to invest part of its resources in local projects. In this context, the supply of funds for a particular region is positively related to the total amount of resources in region pension funds. As described on the empirical section below, the variable assets can be used to identify the demand schedule for VC investment. In this case, the supply required return to invest in VC projects in a particular region is given by:

$$R^s_t = \alpha_0 + \alpha_1 X_{t+} \alpha_2 I^s_t + \alpha_3 \text{Quality}_t + \alpha_4 \text{Assets}_t + u_t$$

Where,

I^s_t: Supply of funds to VC investment (US$).

X_t : is a set of exogenous variables (region growth, for example)

Quality_t: is a set of variables capture the quality of the projects in a particular region (citation of paper and patent are used)

Assets_t: is the total assets of local and state public pension funds in US$.

u_t: supply shocks

The demand for VC funds is directly linked to the incentives to entrepreneurship. One important factor affecting the decision of a worker to become an entrepreneur is the availability of new ideas in a particular region and time. I try to capture the availability of new ideas in a particular region with 2 variables. The first is the number of times the papers were cited through 2003. Since many high tech innovations are originated in the research laboratories of the universities, it is relevant to have a variable which to measure the research quality in a particular region. The quality of research tends to positively influence the demand for entrepreneurship because it increases probability that a researcher will face a commercially viable application.

But innovations don't take place exclusively in the research labs. A substantial part of new ideas are born inside the companies. To capture this second source of innovation I use the patent citation data. The idea is that the regions with higher quality inventors will be likely to have more potential entrepreneurs. Other factor suggested in some studies of VC determinants in across countries studies is the rigidity of labor market. If it is easy for a worker to go in and out of the labor force then the cost to become an entrepreneur is low. Finally, the capital gain taxation tend to be import in the demand side as pointed in Poterba (1989) and Lerner and Gompers (1992). The demand equation is then given by:

$$R^d_t=\alpha_0+\alpha_1 X_{t+}\alpha_2 I^d_t+\alpha_3 Quality_t+\alpha_4 T_t+\varepsilon_t$$

where,

I^d_t :Demand for VC investment in time t Past VC investments

X_t : is a set of exogenous variables (region growth, for example)

T:Capital Gain taxation

$Quality_t$:is a set of variables capture the quality of the projects in a particular region (citation of paper and patent are used)

ε_t: demand shocks.

Equilibrium when the supply is positively inclined is given by:

$R^d=R^s$

$\beta\ \alpha_0+\alpha_1 X_{t+}\alpha_2 I^d_t+\alpha_3 Quality_t+\alpha_4 T_t+\varepsilon_t$

=

$\alpha_0+\alpha_1 X_{t+}\alpha_2 I^s_t+\alpha_3 Quality_t+\alpha_4 Assets_t+u_t$

$I_t=(((\alpha_0-\beta_0))/((\beta_2-\alpha_2)))+(((\alpha_1-\beta_1))/((\beta_2-\alpha_2)))X_t+(((\alpha_3-\beta_3))/((\beta_2-\alpha_2)))Quality_t+$

$+((\alpha_4)/((\beta_2-\alpha_2)))Assets_t+((\beta_3)/((\beta_2-\alpha_2)))T_t+(((\varepsilon_t-u_t)/((\beta_2-\alpha_2)))$

These ideas motivate the empirical model explained in the next section.

1.4- Empirical Model

The empirical model to be estimated is a multilevel panel with 3 dimensions: time (1981-1999), sector (7) and region (179 BEA areas). The equation is given by:

$I\_{i,j,t} = \gamma_0+\delta_t+\eta_i+\theta_j+\gamma_1 Paper_{i,j,t}+\gamma_2 Patentc_{i,j,t}+\gamma_3 Patentnc_{i,j,t}$

$+\gamma_4 Growth_{i,j,t}+\gamma_5 Tax_{i,j,t}+\ \varepsilon_t$

This model includes time, region and sector fixed effects δ_t, η_i, θ_j. The dependent variable is the real investment of VC firms and funds (US$ investment by thousands of

BEA habitants). The explanatory variables are the total citation of papers (paper), citation of patents form corporations (Patentc), citation of patents from individuals (Patentnc), personal income growth (Growth) and capital gain taxation (Tax). Since the investment of VC in different sectors is likely to be correlated inside a particular BEA area and heteroscedastic across areas. In this case, the usual OLS standard deviation will be biased and the t-statistics will be contaminated. To correct for this problem, I used a robust method to estimate the standard deviation of the coefficients in all the estimations[1].

If the supply of funds to venture capital industry is perfectly elastic, then the parameters of this reduced form model would coincide with the parameters of the demand equation. Thus, we would not have any simultaneity problem interfering with identification of the structural model parameters. However, with a positively sloped supply of fund we have to worry about identification. In this case, I am able to estimate a restricted form of the above model where the only potentially endogenous variable is the paper citation and VC investment. I exclude the patent varibles. This does not seem to be a problem because the paper citation tends to drive out theses variable, more details in the results section. The simplified system estimated is:

---

[1] To explain the principle of robust standard deviation start with the usual multiple regression model $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$ where $E\{\boldsymbol{\varepsilon}\} = \mathbf{o}$ and $E\{\boldsymbol{\varepsilon\varepsilon}'\} = \boldsymbol{\Omega}$, a positive definite matrix. Then the covariance matrix of the OLS estimate $\mathbf{b} = (\mathbf{X'X})^{-1}\mathbf{X'Y}$ is $\boldsymbol{\sigma}^{\mathbf{2}}\{\mathbf{b}\} = (\mathbf{X'X})^{-1}\mathbf{X'\Omega X}(\mathbf{X'X})^{-1}$.When the errors are homoscedastic, $\boldsymbol{\Omega} = \sigma^2\mathbf{I}$ and the expression for $\boldsymbol{\sigma}^{\mathbf{2}}\{\mathbf{b}\}$ reduces to the usual $\boldsymbol{\sigma}^{\mathbf{2}}\{\mathbf{b}\} = \sigma^2(\mathbf{X'X})^{-1}$ OLSCM = $\mathbf{s}^{\mathbf{2}}\{\mathbf{b}\} = \mathrm{MSE}(\mathbf{X'X})^{-1}$ (where MSE = $\Sigma e_i^2/(n-p)$) OLSCM denotes the usual OLS covariance matrix of estimates. The basic idea of robust standard errors is that when the errors are heteroscedastic one can estimate the observation-specific variance σ_i^2 with the single observation on the residual as $\hat{}\boldsymbol{\Omega}_{ii} = (e_i - 0)^2/1 = e_i^2$ $\hat{}\boldsymbol{\Omega} = \mathrm{diag}\{e_i^2\}$. This leads to a ***heteroscedasticity consistent covariance matrix*** HC1 = $(n/(n-p))\,(\mathbf{X'X})^{-1}\mathbf{X'}\mathrm{diag}\{e_i^2\}\mathbf{X}(\mathbf{X'X})^{-1}$
where n/(n-p) is a degree of freedom correction factor that becomes negligible for large samples.
HC1 is called the Huber-White estimator (after Huber (1967); White (1980)) or the "sandwich" estimator because of the appearance of the formula.

$$\text{Demand: } I^d_{i,j,t}=\alpha^d_0+\delta^d_t+\eta^d_i+\theta^d_j+\alpha_1 Paper_{i,j,t}+\alpha_2 Tax_{i,j,t}+\varepsilon_{1t}$$

$$\text{Supply } I^s_{i,j,t}=\alpha^s_0+\delta^s_t+\eta^s_i+\theta^s_j+\alpha_1 Paper_{i,j,t}+\alpha_2 Assets_{i,j,t}+\varepsilon_{2t}$$

$$I^d_{i,j,t}= I^s_{i,j,t}$$

where, $\alpha^k_0+\delta^k_t+\eta^k_i+\theta^k_j$ k=s,d are supply and demand region, time and sector fixed effects

Notice that in this case demand and supply for VC investment are identified. The variable capital gain taxation is present only in the demand equation and the variable total asset of pension funds (Pension) is present only in the supply equation. The endogenous variables are Citation of Paper and VC investment by thousands of BEA habitants.

1.5- Results

1.5.1- Single Equation Estimation

In table 1.4, I present the estimation results for the basic model assuming an infinitely elastic supply of funds, that is, considering the case where the equilibrium VC investment is completely determined by the demand factors. The results indicate that personal income growth and citation of papers variables are relevant to explain the variability of the VC investment. Patent (corporation and individuals) variables are not significant at a 5% level. In order to evaluate how much of the results are driven by the regions with large volumes of VC investment, I estimated the same equations excluding the San Jose-San Francisco-Oakland and Boston-Worcester-Manchester areas. Together these regions represent almost 40% of the total VC investment in this period.

Furthermore, these are regions with great research excellence. Some of the most prominent centers of technological research such as MIT and Stanford are located there. Thus, it could be the case that the results are being heavily influenced by these outliers. It turns out that this does not seem to be the case. The results in table 1.5 show that the overall picture is the same as in table 1.4. The only relevant variable is still the citation of papers and personal income growth.

In table 1.6, I estimate the investment equation for each sector separately. This estimation gives us two informations. First, it shows that the coefficient of the citation of papers variable varies a lot across sector from 4.15 in the computer sector to 0.02 in medical sector. Second, it shows that the mapping I constructed to link the different datasets may bring noise to the relationship between VC investment and paper citation. The sectors for which the links are more obvious (computer, biotechnology and medical) the relation is quite strong. On the other hand, for the remaining sectors (communication, industrial and semiconductors) the relation is not statistically significant.

1.5.2- Simultaneous equation estimation

In tables 1.9 and 1.10, I relax the assumption of infinitely elastic supply of funds by considering the system of supply and demand presented in the previous section. In this system, the size of pension funds in a particular region is as a supply shifter. The idea is that, for political reasons, the local and state funds tend to have more investment at the local level. I used the total assets of public and local pension funds as an instrument to the paper citation variable. To identify the supply equation I used the variable capital gain taxation as a demand shifter. As I argued before, taxation is likely to affect mostly the demand for VC funds because a large fraction of the investors in these sectors (pension funds) are tax exempt. Given that the previous results suggest that the patent variables, taxation and income growth are not important to explain VC investment, I didn't include these variables here.

Evidence that pension funds shift the local supply of funds is the existence of the so called Economically Targeted Funds (ETI).The most comprehensive listing of ETI activity I found was in a 1993 Boice Dunham Group study, which was commissioned by Goldman Sachs. This analysis defined an ETI as "an investment by a public pension fund which, in addition to offering financial returns in proportion to financial risk, also offers collateral local economic benefit (e.g., job creation, home ownership) (p. 1)." Using this definition, Boice Dunham concluded that ETIs accounted for $17.5 billion or only 2 percent of the $887.3 billion of public plan assets covered in their survey. ETI activity fell into three categories: fixed income, real estate, and venture capital, with the majority going to residential mortgages see Table 1.7.

Based on this numbers I did estimation of the role of ETI in the VC industry assuming that the share of ETI on the pension funds assets is homogeneous across states. Taking the stock of investment in the VC sector by state in 1993 (by accumulating annual investment from 1980 to 1993) and then calculating the fraction of ETI in VC investment by assuming that the numbers from the table above is valid for all states. This gives that a fraction of 0.18% of the pension funds is on VC. In table 1.8 I show the proportion of VC funds by states. We see that for states with large VC funds this fraction is quite small. For example, this fraction is 1.96% in California and 0.56% in Massachusetts. But it is quite significant in other states with smaller VC funds such as Wisconsin (32.38%) and Illinois (10.55%). Of course, these calculations are approximate at best and maybe very misleading if the proportion of ETI varies a lot across states. But, at least, they give some plausibility to use the pension funds assets as a supply shifter.

In tables 1.9, 1.10 and 1.11, I present the estimation results for the system of equations suggested in the previous section. In these regressions, all the variables are expressed in per capita terms. The results show that the paper citation is still statistically significant both in demand and supply equations. The variable capital gain taxation is marginally statistically significant with a t-statistic of 1.8. However, the positive sign of this variable is not what we would expect. On the supply side, it is important to notice that the pension funds are statistically significant. Thus we cannot reject the hypothesis that regions with large pension funds will have more investment in the VC sector. When I exclude the Boston and San Francisco areas the results are qualitatively the same, only the coefficient on paper citation decreases from 0.03 to 0.02 and the capital gain taxation has the expected sign but still not significant. In terms of economic significance these estimates implies an elasticity of the demand for VC investment roughly equal to the elasticity of supply of about 0.26. The first stage regressions are presented in table 1.7 we see that pension funds variable is statistically significant and have positive effect in VC investment, capital taxation is not significant. Finally, these results are also robust to exclusion of regions with no VC investment in the period as can be seen in table 1.11.

1.5.3- Using Federal Funds to Universities to Instrument paper quality

This section I make a robustness test on the results presented in the previous sections. In particular, I'm interested on controlling for the possibility of endogenous effects and reverse causality on the empirical model. It could be the case that the volume of VC investment in a region has a positive effect on the quality of the academic production of this area. For instance, the presence of VC in biotech in Boston may attract the best scientists to work on that area. In order to control for this effect, I use the National Science Foundation (NSF) R&D Expenditures at Universities and Colleges as an instrument for paper citation. The idea is to use the R&D expenditures prior to 1980 as an instrument for the paper citation after 1980. It seems plausible to assume that these R&D expenditures are not affected by the VC investment since the great boom in the VC industry occurred after 1979 with the amendment in the Prudent Man Rule which increased the investment of pension funds in VC funds[2]. On the other hand, these expenditures is likely to have a positive impact on the quality of academic research in subsequent years.

Unfortunately, the data of NSF data on academic expenditures prior to 1980 are only available for 3 sectors: biological sciences, computer and medical. I present the results in table 12. From the first stage regression we see, as expected, that R&D expenditures from 1973-1980 have a positive effect on academic production after 1980 in all sectors. Moreover, the second stage shows that the exogenous component of paper citation is still important to explain the VC investment in biotech and medical sectors and it is not in the computer sector. Again, the problem with the computer sector may be

[2] Prior to 1979, the Employee Retirement Income Security Act (ERISA) limited pension funds from investing substantial amount of money into venture capital. The 1979 Labor Department clarification of the rule explicitly allowed pension funds managers to invest in venture capital. According, to Kortum and Lerner (2000), in 1978, from the total US 424 millions invested in VC funds, individuals accounted for the largest fraction (32%) while pension funds accounted for only 15%. Eight years later, the total amount jumped to US$ 4 billion with pension funds accounting for more than 50%.

related to the link across database that I proposed. This link is much tends to be much more precise in the medical and biotech sectors than in computers. Overall, it seems that the paper citation have positive impact on the VC investment.

1.5.4- International Comparison

In this section I perform regressions using cross country data. In Jeng and Wells (2000) the variables identified as the main determinants of VC investment variables linked to the financial market development such as volume of IPOs, stock market capitalization and accounting standards. The idea here is to test if the academic production, a variable not considered in these cross country studies, is also relevant to explain VC investment. It should be stressed that the quality of data on this level is much poorer. For example, I don't have data of VC investment divided by sector and the academic production variable I got was the total number of papers produced in each country not the citation on this paper which would give a quality adjusted measure. This last point should not be a big concern since the paper citation is highly correlated with the total number of papers produced. Most importantly, I don't have a time series of the number of paper across countries. I have only the total number of papers produced in the period 1990-1999.

One interesting feature of the international venture capital investment is that prior to the nineties almost all investment was made in the United States. This can be seen in figure 1.1, from 1980 to 1989 the share of US in VC investment has averaged more than 97%. Thus it's would be very difficult to identify the determinants of VC investment with this data simply because there is very little information outside the United States. It's likely that any variable with distinctive values for the EUA, as for example stock market capitalization, would be picked as an important determinant. This is one advantage to the approach in this paper of using US regions instead of country data. After 1990, the VC investment starts a process of internationalization evidenced by a declining share of US investment. Unfortunately, I do not the time series dimension in my variables to explain this phenomenon. What I show in table 1.13 is that in the period in which the VC investment picks up in the rest of the world, after 1990, the total number of papers is a statistically significant variable to explain the cross section of VC investment. This result is even more pronounced when I exclude the United States from the sample, in table 1.14 the t-statistics for the total number of paper estimated coefficient is higher in this case.

1.6- Concluding Remarks

The purpose of this paper was to explain what drives the Venture Capital investment in a particular area and time. The results indicate that the most important variable to explain the variability of VC investment in the US regions is the quality of academic production measured by the total citation of papers. The results are robust to the exclusion of the areas with largest presence of VC (Boston and California) and to the possibility of a positively inclined supply of funds due to the investment policies of the local and state pension funds. The results for the case with an infinitely elastic supply of funds is remarkably similar to the case where the total assets of public local and state pension funds is used to explain the supply side of VC models. Finally, with the limited data available at the international level, it seems that academic production is also an important variable to explain the cross section of countries VC investment.

1.7- 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.

Barro, Robert J. and Xavier Sala-i-Martin (1992). "Convergence." The Journal of Political Economy,Vol 100, No.2, 223-251.

Bottazi, Laura and Marco Da Rin (2001). " Venture Capital in Europe: Euro. nm and the Financing of European Innovative Firms." 33rd Economic Policy Panel Meting, Stockholm.

Cochrane, John (2005). "The Risk and Return of Venture Capital." Journal of Financial Economics 75, 3-52.

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.

Jaffe, A. (1989), .Real Effects of Academic Research. American Economic Review, 79:957-970.

Gupta, Anil K. and Harry J. Sapienza (1992). “Determinants of Capital Firms Preferences Regarding the Industry Diversity and Geography Scope of their Investments.” Journal of business Venture 7, 347-62.

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.

Huber, P. J. 1967. “The behavior of maximum likelihood estimates under nonstandard conditions.” In *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*. Berkeley, CA: University of California Press, vol. 1, 221–223.

Jeng, Leslie A. and Philippe C. Wells (2000). "The Determinats of Venture Capital Funding." Journal of Corporate Finance 6, 241-289.

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

Kaplan, Steve and Per Stromberg (1999). “Financial Contracting Theory Meets the Real World: An Empirical Analysis of Venture Capital Contracting.” Review of Economic Studies 70(2), 281-315.

Kaplan, Steve and Per Stromberg (2001). Venture Capitalists as Principals: Contracting, Screening an Monitoring." American Economic Review (Papers and Proceedings) 91, 426-430.

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

Lerner, Josh and Paul A. Gompers (1999). "What Drives Venture Capital Fund Raising?" NBER working paper 6906.

Poterba (1989) : Poterba, James M. (1989) "Venture Capital and Capital Gains Taxation." In Lawrence Summers editor. Tax Policy and the Economy. Cambridge: MIT Press.

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

White, H. (1980), "A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity," *Econometrica*, 48, 817-838.

Chapter 2- The Impact of Venture Capital on Innovations and on the Creation of New Businesses[3]

2.1- Introduction

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. Some examples are: VC backed firms produce more and more valuable patents Kortum and Lerner (2000), they are faster in developing products and introducing them in the market and they have a higher rate of executive turnover, Hellman and Puri (2000, 2001). Furthermore, the presence of VC sector seems to produce superior macroeconomic performance, Wasmer and Will (2000). The present project is intended to fill this gap in the literature.

The crucial question is to determine whether VC naturally goes where is 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.

Thus, the central problem of the project is to disentangle the effect of the exogenous component of the VC investment on the number of patents (our proxy for

[3] Prof. Luigi Zingales was my co-author on this paper

innovation) and on the number of new companies using those patents. Clearly, the positive association of VC investment and number patents may be driven by the association of these two variables with a third one. For example, the variation on the availability of opportunities of innovations to entrepreneurs in a particular state. Also, we may have a reverse causality problems, that is, the rise in patents inducing more VC investment.

We plan to address these problems in two ways. First, to control for omitted factors related to the supply of high-tech ideas we will use the production of academic papers in different scientific areas in different US regions. Second, we will use the time series and cross sectional difference 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.

The present paper uses as regional unit, the Bureau of Economic Analysis (BEA) economic areas which are composed by counties surrounding metropolitan areas. There are a total of 179 BEA areas. The large amount of geographical units increases substantially the degrees of freedom of the statistical model. In addition, given the monitoring and advisory rule of VC capitalist, it is a well documented fact that the VC firms tend to invest in projects close to its geographical base. By using more restricted defined regional units, it is possible to have a more precise idea of the regional impact of the VC investment on the number of patents. Also, it increases the precision of the model fixed effects.

As a robustness test of our result we also implement a version of the model where we control for the possibility of endogenous effects by using 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 investment which would help to 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 an OLS regression of VC investment on pension funds assets and time, region and sector fixed effects. The second stage is the Poisson model with VC investment predicted from the first stage plus other controls.

Finally, we address the question of determining if the presence of VCs in a particular region is really a factor that promotes the creation of new business. Basically, we are interested to know if the positive impact of VCs on patents translates to the effective start of new business. It could be the case that the positive association between patents and VC investment is present only because VCs tend to patent more, for example, for guarantying inventions intellectual properties.

This topic has important policy implications. It would shed some light on whether there is any social benefit in subsidizing the creation of a local VC industry and on ways to improve the transfer of scientific production from universities to commercial applications.

2.2- Related Literature

The apparent success of the Venture Capital industry in the US has raised the question of what is the effective contribution of the VC to the economy. In Kortum and Lerner (1998), this question is addressed by testing the impact of VC investment on innovations. The authors use as a measure of innovations the number of patents applications on 20 industries from 1965 to 1992. One important point of the analysis is how to address the causality problem. It may be the case that the association between VC investment and patenting is the result of the influence of a third non-observable variable, the opportunities to investment, on these two variables. The paper deals with this issue by exploring the regulatory changes in the VC industry of 1979 as an instrument to VC investment. These changes produced an exogenous shift in the resources to VC activity and are not likely to be correlated with patenting. After controlling for these causalities problems, the paper shows that VC investment has as strong positive effect on patenting. The result, however, could be criticized on the grounds that VC investment causes an increase in patenting with no impact on innovations. Venture backed firms may patent more their innovations just because they want to attract potential investors or to prevent the copy of their ideas. The paper try to test this hypothesis by using a sample of 530 venture backed and non venture backed companies based in Middlesex County, Massac huts. They show that VC backed firms not only patent more, but also their patent are more often cited in other patent applications and are more aggressively litigated. In conclusion, the paper supports the thesis that VC investment has an important impact on innovations.

In Ueda et all (2003) the question of the impact of the VC investment on innovations are addressed from a different perspective. The author's main concern is that the positive correlation between innovation and VC investment documented in the literature may be contaminated by a reverse causality problem. That is, the arrival of innovations may spur start ups which in turn increase the demand by VC investment. Thus, if this is true, innovations are causing VC investment and not the opposite. The measure of innovation used in the paper is the total factor productivity (TFP) in the US manufacturing industry. In a dynamic panel data set up, they implement a Granger causality test between TFP growth and VC investment. The test shows that TFP growth significantly and positively leads VC investment and VC investment are significantly and negatively associated with one year lagged TFP growth. They also study the time series relation between TFP growth and VC investments for each industry in their sample. The results of their time series analysis show that the causality runs in both directions. In computer and communication industries, however, the relation between TFP growth and VC investments are negative in drug and scientific instrument industries. The conclusion of the paper is that the impact of VC on innovations is not as clear cut as showed in Kortum and Lerner (1998), the hypothesis that innovations drives VC investment cannot be rejected in some cases.

In Darby and Zuker (2003) the focus is the causes of biotechnology start-ups. The authors address this question by studying the determinants of entry in the nanotechnology sector, a sub sector of the biotechnology. They use as a geographic unit the US Bureau of Economic Analysis areas, 172 regions delimited by commuting and shopping patterns. The sample is from 1981 to 1999. To define measures of the science base by region and year, they employ the top 112 universities and high-impact (highly cited) papers. They separate the papers in two groups one with authors filiated to a nano firm and the remaining. A firm entry in a region is counted in the year of the first publication in the firm-article dataset. Because their dependent variable is a counting integer, the number of firms entering in the nanotechnology sector, the econometric model utilized is the Poisson regression. The independent variables they constructed are science base, frequency and size of Venture Capital deals and other regional variables. They find that firms are entering where and when scientists are making major scientific breakthroughs as well as where the quality of the labor force is high. The venture capital, on the other hand, tends to follow technological opportunities rather then drive them.

This brief literature review suggests that the economic impact of VC investment is controversial. While Kortum and Lerner (1998) suggest that VC does have an important impact on innovations, other papers suggest that it is the other way around, that is, in some cases VC follow TFP growth in investment as in Ueda et all (2003) and in others VC investment tend to follow technological opportunities in the nanotechnology sector as in Zucker, Darby (2003). The present paper hopes to contribute to this debate.

The present paper is also related to the literature that presents evidence of spillover effects from the production of knowledge in the universities to the regional economies where they are located. For example, Anselin, Varga and Acts (2002) provides evidence of local spatial externalities between university research and high technology innovative activity in the United States both at state level and at the Metropolitan Statistical Areas level. Adams (2003) finds that academic spillovers are more localized than industrial spillovers reflecting industry-academic cooperative movements. Audretsch, Lehmann and Warning (2003), using a dataset of high tech companies from Germany, find that firms have a propensity to locate close to universities and proximity to universities reduces the time between firm formation and listing on the stock market. Florax (1992), using a Netherlands dataset, suggests that academic knowledge production have a significant impact on investment in equipments which is strongest in the neighborhood of central places. Andreson, Quigley and Wilhelmsson (2004) find evidence that Swedish policy of spatial decentralization of post-secondary education had a significant impact on regional development. Jaffe (1989) finds that university research has a significant positive effect on corporate patent and an indirect effect on local innovation by inducing R&D spending.

2.3- Data Description and Methodology

BEA Regions. 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. The division of the United States in these areas is displayed in picture 1 in the appendix.

There are 179 regions in the present division used by the BEA. The construction of these regions involves three phases. First, the economic nodes are identified. These are metropolitan areas or micropolitan areas, according to the definitions of the U.S. Office of Management and Budget for Metropolitan Statistical Areas and Core Based Statistical Areas that serve as the main centers of economic activity. In the end of this part, 344 economic nodes are identified consisting of 1311 counties. Second, each of the remaining 1830 counties are assigned to these economic nodes based on county-to-county commuting flow. Finally, in the third major phase, the 344 CEAs were used as building blocks to form the new 179 BEA economic areas. The two major factors that were used in the aggregation process were commuting patterns and economic and geographic size. The aggregation is such that each economic area is a region of sufficient size to support regional statistical analyses and each economic area is a labor market that is independent of other labor markets. For a more detailed explanation of these areas see Johnson and Kort (2004).

Basically, there are three main reasons to use this region division in this work instead of the usual states division. 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, from the statistical point of view the degrees of freedom are substantially increased, the BEA is composed by 179 regions as opposed to 50 states. In this sense, this could potentially increase the reliability of region specific effects. Furthermore, the Bureau of Economic analysis publishes a large number of economic data for these regions such as personal income, wages and employment which potentially could be used as controls in the empirical model. Figure 1 in the appendix depicts the US map divided by the BEA areas.

Thomson Science Indicator. We use the product University Science Indicators 1981-2003 from Thomson scientific to obtain a measured of scientific production in each region. This product is database of research performance measures on 284research universities in the United States. Based on the publication and citation statistics of ISI® from the same company, University Science Indicators contains the number of ISI-indexed papers from each university and the number of times the papers were cited through 2003. The use of citation data is especially important because it gives a measured of the quality of the scientific production. The variable used in this work is the number of times the papers were cited through 2003. In this database there are 74 fields of research. The variability in the fields is a potentially important source of identification of region specific effects.

Venture Capital Investment. The data on 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). In table 1 in the appendix I present fraction of the total investment of each sector. This paper works only with the sectors that could potentially be linked with academic research. These are what we called Tech sectors: Biotechnology, Computer (hardware, software, services), Industrial/Energy, Medical/Health, Semiconductors/Other Elect. The other sectors (non Tech) are: Consumer Related, Internet Specific, Other Products. These are left out of the study since they don't seem to have a direct link with the academic production. Thus the paper will try to explain a fraction of about 66% of the VC investment. In table 2 we show the frequency of VC investment in particular ranges for some periods in the sample. It can be seen from there that most of the areas have positive investment in all the periods.

Patent Citation Data. The data of patent citations are from the NBER patent citation dataset. For this study it is used the citation received by US patents with application year between 1980 and 1999. These patents are divided in 2 groups: assigned and unassigned patents. The patents assigned to US non government organizations (mostly corporations according to the NBER). In the US, a granted US patent is assigned to the inventor. If that inventor works for a company, as part of their employment, they assign the patent to their 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. For a detailed exposition of the methodology employed in the construction of the NBER Patent Dataset see Hall, Jaffe and Trajtenberg (2000). For this study we used patents unassigned and patents assigned to companies with no COMPUSTAT code. The objective is to exclude patents assigned to established companies since these would not be linked to VC finance and start up companies.

Pension Funds Data. State and Local Government Employee-Retirements Systems is an annual survey conducted by the Census Bureau. The survey coverage includes public employee retirement systems administered by state and local governments throughout the nation. State and Local Government Employee-Retirements Systems is an annual survey conducted by the Census Bureau. The survey coverage includes public employee retirement systems administered by state and local governments throughout the nation. I use in this paper the total assets of these funds deflated by the PPI. The data is available from 1993 to 1999.

New Business. Ideally, we would like to have a measure of the total measure of new businesses in a particular region (our BEA economic areas) and time. Unfortunately, we don't have this exact measure. Consequently, we have to use a proxy which we hope is a close approximation. Here, we use the change in the total number of establishments in a BEA area from one year to another. An establishment is a single physical location at which business is conducted or services or industrial operations are performed. It is not necessarily identical with a company or enterprise, which may consist of one or more establishments. When two or more activities are carried on at a single location under a single ownership, all activities generally are grouped together as a single establishment. The entire establishment is classified on the basis of its major activity and all data are included in that classification. This proxy has two limitations. First, it is the net change in the number of establishments so it includes establishments closed in the period. Second, it is derived from the total number of establishments not the total number of companies. Thus, it will show a expansion not only when a new business is started but also when a company is expanding its number of establishments. We hope that this is not much a problem in the VC financed companies which, especially the technological sector, usually don't have a large number of units.

Mapping from the patent, paper and new business datasets to the VC investment dataset. Ideally, to link these datasets I would need the information on all individual projects to determine which paper and patent were relevant to its development. Unfortunately, I don't have access to this data at this point and have 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 3 in the appendix. It can be seen that the link is quite obvious in some cases as in the case of the medical sector. But is not so clear in others like the industrial/energy sector. 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. 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.

2.4- Statistical Models

The dataset we use to test the impact of VC investment on patents have two important characteristics: First, the dependent variable, number of citation on patent, is a count of the total number of citations received by a patent. 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.

We implement 3 different model specifications:

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

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

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

Where, n_{ijt} is the number of citation of papers in region i time t, X_{ijt} is the vector of explanatory variables and μ's are fixed effects explained below. The explanatory variables are Citation of papers, past VC investment. Notice that from the properties of the Poisson distribution we have:

$E\ n_{ijt} = \lambda_{ijt}$

These models explore different variability of the data to identify the role of VC investment and paper citation on patenting. Model 1 use as a group a sector/region pair, for example, biotech in California is one group and biotech in Boston is another. Thus, in this case we use the time variability to identify the parameters. In model 2, the group is defined as a time/region pair and we use then the sector variability for identification. Finally, in model 3, we use a 3 level panel where we employ time, region and sector fixed effects separately.

The explanatory variables are Citation of papers, past VC investment. As a robustness test of our result, we also implement a version of the model where we control for the possibility of endogenous effects by using 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 investment which would help to 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.

As for the model used to test the impact of VC and patents on the creation of new business we have a different problem. The data for new business is the change in the number of establishments from the US Census Bureau *County Business Pattern* dataset. The limitation with this dataset is that we have a short time series (9 years).The problem emerges because VC investment is likely to affect the creation of new business with lags 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. The solution we give is to collapse the time dimension by taking the time average of the observations and then using 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, may take negative values.

2.5- Results

In table 2.4 in the appendix we present the results for the 3 models present in the previous section. From there we can see that both citation of papers and VC investment is important to explain the variability in patent citation. So VC investment has a direct impact on innovation that persists after we control for the availability of new ideas proxied here by the paper citation variable. This result is in contrast with the findings of Darby, M. and Lynne Zucker(2003). In terms of economic magnitude, the impact of VC on patents is significant. By the coefficients estimates from table 2.4, a one standard deviations increase on the VC investment per capita will produce an increase from 4 to 15% on the number of patents depending on the model used.

In table 2.5, we perform the same regressions but excluding Boston and San Francisco areas, the two regions with largest volume of VC investment and also prominent academic centers. 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 2.4.

In table 2.6, we perform the same regression but using the total number of papers instead of paper citation as explanatory variable. The reason to do this is a concern with reverse causality from patents to paper, that is, papers that produce patents tend to be more cited. The regression suggests that there is no strong evidence in this direction; the number of papers is significant in models 2 and 3 with even higher coefficients.

In table 2.7 we perform the robustness test of controlling for the possibility of endogenous effects by using the total assets of local and state pension funds as an instrument for VC investment. In the first column we see that the instrument is statistically significant in the first stage regression and, in column 2, we show that the results are still present, both citation of papers and VC investment are still statistically significant.

In table 2.8 we present the results for the regression with new business variable as the dependent variable. Here new business is the change in the number of establishments from one year to another. As we mentioned before, we exploited only the sector and region variability of the data. We collapse the time dimension by averaging the observation with the objective the problem of determining the lag in such a short time series. We see in table 8 that all the variables VC investment, number of patents and paper citation are individually statistically significant to explain the variation on the number of new business. Notice that the coefficient on the patent variables decreases when the other two variables are included into the regression. Also, in the last column (6) which includes all the variables, the VC investment is the one with highest t-statistic suggesting 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.

2.6- Concluding Remarks

This paper tries to access the contribution of Venture Capital investment on the promotion of innovations and new business. 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 as an instrument for VC investment the total assets of pension funds. 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 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 deviations increase on the VC investment per capita will produce an increase which varies from 4 to 15% on the number of patents depending on the model used.

2.7- 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 (2000) "Assessing the Contribution of Venture Capital to Innovation." RAND Journal of Economics, vol 31 p.674-692.

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.

Appendix A, Tables

Table 1.1- Frequency of Investment in BEA Areas

Investment Range (US$)	Period 1981 - 1985	1986 - 1990	1991 - 1995	1996 - 1999	1981-2004
0	74	64	63	41	25
0-50000	79	75	74	64	64
50000-100000	5	13	13	13	20
100000-200000	11	7	8	16	15
200000-500000	6	10	5	15	21
500000-1000000	1	4	11	8	6
1000000-5000000	3	6	4	17	18
>5000000	0	0	1	5	10
Total	179	179	179	179	179

This table gives the number of BEA areas with investment in the given range in each of the periods.

Table 1.2- Investment of VC Funds and Firms by Sectors (% of period total)

Sector	Period 1980-1984	1985-1989	1990-1994	1995-1999	2000-2004	Grand Total
Tech Sectors:	**80.03%**	**52.30%**	**60.37%**	**71.07%**	**74.05%**	**71.19%**
Biotechnology	3.93%	3.97%	5.94%	3.45%	4.22%	4.07%
Communications and Media	12.86%	13.08%	15.75%	14.94%	13.46%	14.00%
Computer Hardware	25.65%	8.48%	3.62%	2.21%	2.29%	3.15%
Computer Software and Services	8.94%	6.12%	8.72%	11.82%	13.44%	12.17%
Industrial/Energy	11.15%	6.46%	9.80%	6.71%	6.40%	6.78%
Internet Specific	0.19%	0.40%	0.98%	19.37%	19.96%	17.16%
Medical/Health	7.29%	7.97%	11.92%	8.70%	7.25%	8.00%
Semiconductors/Other Elect.	10.01%	5.82%	3.63%	3.87%	7.03%	5.87%
Non Tech Sectors:	**19.97%**	**47.70%**	**39.63%**	**28.93%**	**25.95%**	**28.81%**
Consumer Related	10.94%	24.64%	16.90%	12.69%	7.64%	10.76%
Other Products	9.03%	23.06%	22.73%	16.24%	18.30%	18.04%
Grand Total	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%

Table 1.3- Fields Links across datasets

This is the mapping from the patent and paper dataset to the VC investment dataset

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
Computer Hardware, Software and Services	Computer Hardware & Software Computer Peripherials Information Storage	Computer Science & Engineering
Consumer Related	**Amusement Devices** **Apparel & Textile** **Furniture,House Fixtures**	**No Link**
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

Table 1.3 - Continued
FIELDS Links Across Datasets

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 1.4 - Estimation Results for the sample 1981-1999 including all 179 BEA regions and 7 tech sectors.

Dependent Variable: Investment by VC firms and funds (by BEA population)			
Independent Variables			
	(1)	**(2)**	**(5)**
Citation of Papers (by BEA population)		3.28	3.39
		(0.78)***	(0.81)***
Citation of Patents from corporations (by BEA population)			-0.23
			(0.14)*
Citation of Patents from to individuals (by BEA population)			-0.69
			(0.55)
Capital Gain Taxation			-0.10
			(0.15)
Per Capita Real Income Growth			25.92
			(7.73)***
Interaction Paper/Sector		Yes	Yes
Sector Fixed Effects	Yes	Yes	Yes
Region Fixed Effects	Yes	Yes	Yes
Sector Fixed Effects	No	Yes	Yes
Number obs.	20406	20406	20406
R-sq	0.10	0.11	0.11

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).
Citations of Patents from corporation and individuals are from the NBER patent citation dataset.
For this study it is used the citation received by US patents with application year between 1980 and 1999. These patents are divided in 2 groups: the patents assigned to US non government organizations (mostly corporations according to the NBER). In the US, a granted US patent is assigned to the inventor. If that inventor works for a company, as part of their employment, they assign the patent to their 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.
Citations of papers is from the Thomson Scientific Science Indicator. This product is database of research performance measures on 284 research universities in the United States. The variable used in the above regression is the number of times the papers were cited through 2003. In order to mitigate the truncation problem of the citation dataset the total number
Capital gain taxation are obtained from the NBER Taxim software a program for calculating liabilities under US Federal and S
These data are combined to form a panel in which the time dimension is from 1981 to 1999 and the group dimension is com
The statistical model also include fixed effects for time (δt), region (ηi) and sector (θj).
*Statistical significant at 10%, **Statistical significant at 5% ***Statistical significant at 1%

Table 1.5 - Estimation Results for the sample 1981-1999 excluding Boston and California areas

Robust Standart Deviation clustering by BEA areas			
Dependent Variable: Real Investment by VC firms and funds (US$/PPI)			
Independent Variables			
	(1)	**(2)**	**(5)**
Citation of Papers (Paper)		2.09	2.17
		(0.67)***	(0.69)***
Citation of Patents from corporations (Patentc)			-0.23
			(0.13)*
Citation of Patents from to individuals (Patentnc)			-0.03
			(0.29)
Capital Gain Taxation (Tax)			-0.22
			(012)*
Per Capita Real Income Growth (Growth)			14.57
			(4.19)***
Interaction Paper/Sector		Yes	Yes
Sector Fixed Effects	Yes	Yes	Yes
Region Fixed Effects	Yes	Yes	Yes
Sector Fixed Effects	No	Yes	Yes
Number obs.	20178	20178	20178
R-sq	0.08	0.09	0.09

Boston and San Francisco areas together represent more than 35% of the VC investment in the US. In this table I check how much of the results are driven by these regions with large VC investment.
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).
Citations of Patents from corporation and individuals are from the NBER patent citation dataset.
For this study it is used the citation received by US patents with application year between 1980 and 1999. These patents are divided in 2 groups: the patents assigned to US non government organizations (mostly corporations according to the NBER). In the US, a granted US patent is assigned to the inventor. If that inventor works for a company, as part of their employment, they assign the patent to their 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.
Citations of papers are from the Thomson Scientific Science Indicator. This product is database of research performance measures (
Capital gain taxation are obtained from the NBER Taxim software a program for calculating liabilities under US Federal and State inco
These data are combined to form a panel in which the time dimension is from 1981 to 1999 and the group dimension is composed by
The statistical model also include fixed effects for time (δt), region (ηi) and sector (θj).
*Statistical significant at 10%, **Statistical significant at 5% ***Statistical significant at 1%

Table 1.6 - Estimation Results by Sectors for the sample 1981-1999 including all 179 BEA regions and 7 tech Sectors

Robust Standart Deviation

Dependent Variable: Investment by VC firms and funds divided by BEA Population

Independent Variables	**Biotechnology**	**Communications**	**Computer**	**Industrial/Energy**	**Medical/Health**	**Semiconductors**
Citation of Papers (by BEA population)	0.024	0.95	4.21	0.015	0.02	0.05
	(0.0049)***	(0.82)	(0.97)***	(0.013)	(0.004)***	(0.08)
Sector Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Region Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Number obs.	3401	3401	3401	3401	3401	3401
R-sq	0.47	0.17	0.08	0.108	0.26	0.32
R-sq fixed effects only	0.44	0.17	0.27	0.11	0.25	0.32

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).

Citations of papers is from the Thomson Scientific Science Indicator. This product is database of research performance measures on 284 research universities in the United States. The variable used in the above regression is the number of times the papers were cited through 2003. In order to mitigate the truncation problem of the citation dataset the total number of citation is divided by the time the patent or paper could have received citations.

These data are combined to form a panel in which the time dimension is from 1981 to 1999 and the group dimension is composed by the 179 Bureau of Economic Analysis Economic areas. In addition there is a third dimension given by the sector of investment of the VC investment and the respective link to the patent and paper dataset.

The statistical model also include fixed effects for time, region

*Statistical significant at 10%, **Statistical significant at 5% ***Statistical significant at 1%

Table 1.7 - State and Local Pension Plans' Economically Targeted Investments, 1993

ETI Activity Percent of Total		
Fixed Income		21.6
Loans to Small Businesses	4.5	
Private Placements	17	
Real Estate		69.3
Construction Loans	4.5	
Residential Mortgages	50.6	
Commercial Mortgages	12.5	
Equity	1.7	
Venture Capital		9.1
Total		100

Source: Boice Dunham Group (1993).

Table 1.8- Participation of Pension Funds in VC investment by state*

(Aproximation)

State	% of VC stock	State	% of VC stock
Alabama	17.26%	Montana	26.83%
Arizona	5.62%	Nebraska	15.57%
Arkansas	61.18%	Nevada	60.37%
California	1.96%	New Hampshire	0.01%
Colorado	2.10%	New Jersey	4.60%
Connecticut	2.12%	New Mexico	19.38%
Delaware	13.80%	New York	19.59%
District of Columbia	5.10%	North Carolina	15.74%
Florida	10.40%	North Dakota	15.47%
Georgia	4.14%	Ohio	26.35%
Idaho	18.98%	Oklahoma	14.74%
Illinois	10.55%	Oregon	2.55%
Indiana	10.51%	Pennsylvania	5.71%
Iowa	38.57%	Rhode Island	0.85%
Kansas	12.90%	South Carolina	31.43%
Kentucky	26.93%	Tennessee	4.49%
Louisiana	24.04%	Texas	3.29%
Maine	4.44%	Utah	6.21%
Maryland	7.45%	Vermont	2.96%
Massachusetts	0.56%	Virginia	7.07%
Michigan	13.30%	Washington	3.80%
Minnesota	6.54%	West Virginia	1.18%
Mississippi	0.67%	Wisconsin	32.38%
Missouri	17.35%		

Excluding states with less than US$5 millions investment in the period

Table 1.9 - First Stage Regressions for Supply and Demand of VC Investment Estimation Results for the sample 1993-1999 including all 179 BEA regions and 7 tech sectors

Dependent Variable (1st stage): Real Investment by VC firms and funds divided

Independent Variables	**(1)**		**(2)**	
	VC Investment	**Paper**	**VC Investment**	**Paper**
Capital Gain Taxation	-2.42	-2.88	1.42	0.24
	(2.17)	(0.61)	(0.57)***	(1.63)
Total Real Assets of Local	1.93	0.37	-0.02	-2.07
and State Pension Funds	(0.76)***	(1.65)	(1.66)	(4.69)
Time Fixed Effects	Yes	Yes	Yes	Yes
Region Fixed Effects	Yes	Yes	Yes	Yes
Sector Fixed Effects	Yes	Yes	Yes	Yes
Number obs.	7518	7518	7392	7392

(1) Estimation Results for the sample 1993-1999 including all 179 BEA regions and 7 tech sectors
(2) Estimation Results for the sample 1993-1999 excluding Boston and California areas

Here I abandon the assumption that the supply of funds to VC investment is infinitely elastic by considering the volume of pension funds investment in a particular region as a supply shifter. The idea is that, for political reasons, the local and state funds tend to have more investment at the local level.

Boston and California areas together represent more than 57% of the VC investment in the US. In column 2, I check how much of the results are driven by these regions with large VC investment.

Citation of papers is from the Thomson Scientific Science Indicator. This product is database of research performance measures on 284 research universities in the United States. The variable used in the above regression is the number of times the papers were cited through 2003. In order to mitigate the truncation problem of the citation dataset the total number of citation is divided by the time the patent or paper could have received citations.

Total Real Assets of Local and State Pension Funds is from the US Census Bureau. The variable used in the regression is the total assets of local and state pension funds deflated by the Producer price index. Because the pension fund assets are only availa
The sample is from 1993-1999.

The 2SLS regressions include sector, region and time fixed effects and the residuals are clusted by BEA regions.

*Statistical significant at 10%, **Statistical significant at 5% ***Statistical significant at 1%

Table 1.10 - 2SLS estimation for Supply and Demand of VC Investment Estimation Results for the sample 1993-1999 including all 179 BEA regions and 7 tech sectors

Dependent Variable (2nd stage): Real Investment by VC firms and funds divided by BEA Population

Independent Variables	**(1)**		**(2)**	
	Supply	**Demand**	**Supply**	**Demand**
Citation of Papers (by BEA population)	0.032	0.034	0.02	0.02
	(0.006)***	(0.006)***	(0.005)***	(0.005)***
Capital Gain Taxation		0.32		-0.19
		(0.17)*		(0.13)
Total Real Assets of Local and State Pension Funds (by BEA population)	1.92		1.53	
	(0.19)***		(0.14)***	
Time Fixed Effects	Yes	Yes	Yes	Yes
Region Fixed Effects	Yes	Yes	Yes	Yes
Sector Fixed Effects	Yes	Yes	Yes	Yes
Number obs.	7518	7518	7392	7392

(1) Estimation Results for the sample 1993-1999 including all 179 BEA regions and 7 tech sectors
(2) Estimation Results for the sample 1993-1999 excluding Boston and San Francisco areas

Here I abandon the assumption that the supply of funds to VC investment is infinitely elastic by considering the volume of pension funds investment in a particular region as a supply shifter. The idea is that, for political reasons, the local and state funds tend to have more investment at the local level.
Boston and California areas together represent more than 57% of the VC investment in the US. In column 2, I check how much of the results are driven by these regions with large VC investment.
Citation of papers is from the Thomson Scientific Science Indicator. This product is database of research performance measures on 284 research universities in the United States. The variable used in the above regression is the number of times the papers were cited through 2003. In order to mitigate the truncation problem of the citation dataset the total number of citation is divided by the time the patent or paper could have received citations.
Total Real Assets of Local and State Pension Funds is from the US Census Bureau. The variable used in the regression is the total assets of local and state pension funds deflated by the Producer price index. Because the pension fund assets are only availa
The sample is from 1993-1999.
The 2SLS regressions include sector, region and time fixed effects and the residuals are clusted by BEA regions.
*Statistical significant at 10%, **Statistical significant at 5% ***Statistical significant at 1%

Table 1.11- 2SLS estimation for Supply and Demand of VC Investment. Estimation Results for the sample 1993-1999 including all 179 BEA regions and 7 tech sectors Sample Excluding Regions with no VC investment in the period

Dependent Variable (2nd stage): Real Investment by VC firms and funds divided by BEA Population

Independent Variables	**First Stage**		**Second Stage**	
	VC Investment	**Paper**	**Supply**	**Demand**
Citation of Papers (by BEA population)			0.028	0.030
			(0.007)***	(0.007)***
Capital Gain Taxation	-2.160	-3.290		0.36
	(2.52)	-5.260		(0.19)*
Total Real Assets of Local and State Pension Funds (by BEA population)	1.9	0.330	2.03	
	(0.86)**	(1.79)	(0.21)***	
Time Fixed Effects	Yes	Yes	Yes	Yes
Region Fixed Effects	Yes	Yes	Yes	Yes
Sector Fixed Effects	Yes	Yes	Yes	Yes
Number obs.	6468	6468	6468	6468

Here I abandon the assumption that the supply of funds to VC investment is infinitely elastic by considering the volume of pension funds investment in a particular region as a supply shifter. The idea is that, for political reasons, the local and state funds tend to have more investment at the local level. In this table I exclude regions with no VC investment in the period.

Citation of papers is from the Thomson Scientific Science Indicator. This product is database of research performance measures on 284 research universities in the United States. The variable used in the above regression is the number of times the papers were cited through 2003. In order to mitigate the truncation problem of the citation dataset the total number of citation is divided by the time the patent or paper could have received citations.

Total Real Assets of Local and State Pension Funds is from the US Census Bureau. The variable used in the regression is the total assets of local and state pension funds deflated by the Producer price index. Because the pension fund assets are only available at the state level, when a BEA area lies in more than one state, I use the sum of assets over these states.

The sample is from 1993-1999.

The 2SLS regressions include sector, region and time fixed effects and the residuals are clusted by BEA regions.

*Statistically significant at 10%, **Statistically significant at 5%, ***Statistically significant at 1%

Table 1.12 – Using Federal Funds for R&D to Universities to Instrument paper quality. Estimation Results for the sample 1981-1999 including all 179 BEA regions

1st Stage: Dependent Variable: Citation of papers			
	Biotechnology	**Computer**	**Medical/Health**
R&D Funds 1973-1980	0.29	0.019	0.397
	(0.014)***	(0.001)***	(0.018)***
Number obs.	179	179	179
R-sq	0.69	0.65	0.73
2ndt Stage: Dependent Variable:Real Investment by VC firms and funds			
	Biotechnology	**Computer**	**Medical/Health**
Ciation of Papers	0.96	329.91	1.030
	(0.41)**	(339.52)	(0.515)**
Number obs.	179	179	179
R-sq	0.12	0.05	0.08

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).

Citations of Patents from corporation and individuals are from the NBER patent citation datase. R&D Expenditures are from the National Science Foundation (NSF) R&D Expenditures at Universities and Colleges Survey

Table 1.13 - Estimation Results: Cross Country Regression

Dependent Variable: VC firms and funds (US$) from 1990-1999 by GDP

Independent Variables (average from 1990-1999)

	(1)	(2)	(3)	(5)	(6)
Number of Papers	793.2				647.31
	(248.4)***				(316.04)*
Accounting Standards		0.007			-0.0001
		(0.005)			(0.006)
Market Capitalization			0.002		0.00
			(0.002)		(0.002)
Real GDP Growth				0.047	0.21
				(0.079)	(0.62)
Number obs.	21	21	21	21	21
R-sq	0.05	0.03	0.07	0.00	

Countries included: Australia, Austria, Belgium, Canada, Denmark, Finland, France, Germany, Ireland, Israel, Italy, Japan, Netherlands, New Zealand, Norway, Portugal, Spain, Sweden, Switzerland, United Kingdom, United States. The dependent variable is the U$ investment of Venture Capital firms by GDP from the Venture Economics Dataset. Accounting Standards is an index created by examining and rating companies' 1990 annual reports on their inclusion or omission of 90 items in balance sheets and income statements and published by the Center for International Financial Analysis & Research, Inc. the maximum is 90, the minimum 0. Publication counts are from a 1985 set of journals classified and covered by the Institute of Scientific Information's Science and Social Science Citation Indexes. Number of papers is article counts based on fractional assignments; for example, an article with two authors from different countries is counted as one-half of an article for each country. The sources are: Institute for Scientific Information, Science Citation Index and Social Science Citation Index: CHI Research, Inc., Science Indicators database; and National

Market capitalization is Stock market capitalization as share of GDP from the World Bank World Development Indicators

Table 1.14 - Estimation Results: Cross Country Regression Sample Excluding United States

Dependent Variable: VC firms and funds (US$) from 1990-1999 by GDP

Independent Variables (average from 1990-1999)

	(1)	**(2)**	**(3)**	**(4)**	**(6)**
Number of Papers	843.21				883.7894
	(153.86)***				(145.18)***
Accounting Standards		0.004			-0.0020
		(0.003)			(0.002)
Market Capitalization			0.0004		0.0005
			(0.0005)		(0.79)
Real GDP Growth				0.019	0.0535577
				(0.05)	(0.05)
Number obs.					
R-sq					

Countries included: Australia, Austria, Belgium, Canada, Denmark, Finland, France, Germany, Ireland, Israel, Italy, Japan, Netherlands, New Zealand, Norway, Portugal, Spain, Sweden, Switzerland, United Kingdom. The dependent variable is the U$ investment of Venture Capital firms by GDP from the Venture Economics Dataset. Accounting Standards is an index created by examining and rating companies' 1990 annual reports on their inclusion or omission of 90 items in balance sheets and income statements and published by the Center for International Financial Analysis & Research, Inc. the maximum is 90, the minimum 0. Publication counts are from a 1985 set of journals classified and covered by the Institute of Scientific Information's Science and Social Science Citation Indexes. Number of papers is article counts based on fractional assignments; for example, an article with two authors from different countries is counted as one-half of an article for each country. The sources are: Institute for Scientific Information, Science Citation Index and Social Science Citation Index: CHI Research, Inc., Science Indicators database; and Natior

Market capitalization is Stock market capitalization as share of GDP from the World Bank World Development

Table 2.1 – Summary Statistics

Sample: 1992-1997

Sector	Statistic	mean	median	sd	min	max
Biotechnology	VC investment	4,970	0	27,797	0	428,843
	Patents	12	0	48	0	578
	Pension	5,857,309	2,284,637	11,800,000	15,947	130,000,000
	New Business	3	1	12	-17	139
	Paper Citation	1,085	81	2,768	0	23,450
Communications and Media	VC investment	11,761	0	52,830	0	955,789
	Patents	17	1	59	0	691
	Pension	5,857,309	2,284,637	11,800,000	15,947	130,000,000
	New Business	22	8	68	-170	1,104
	Paper Citation	5	0	15	0	158
Computer	VC investment	19,392	0	126,757	0	2,573,818
	Patents	23	0	104	0	1,638
	Pension	5,857,309	2,284,637	11,800,000	15,947	130,000,000
	New Business	102	18	303	-18	4,924
	Paper Citation	8	0	23	0	201
Industrial/Energy	VC investment	7,607	0	47,133	0	968,180
	Patents	77	18	187	0	1,901
	Pension	5,857,309	2,284,637	11,800,000	15,947	130,000,000
	New Business	-1	0	44	-409	553
	Paper Citation	336	15	842	0	6,337
Medical/Health	VC investment	10,691	0	42,699	0	601,081
	Patents	43	5	136	0	1,363
	Pension	5,857,309	2,284,637	11,800,000	15,947	130,000,000
	New Business	113	47	271	-875	3,376
	Paper Citation	1,769	23	4,515	0	38,647
Semiconductors/ Other Elect.	VC investment	3,643	0	28,423	0	663,318
	Patents	21	2	71	0	733
	Pension	5,857,309	2,284,637	11,800,000	15,947	130,000,000
	New Business	17	6	55	-431	590
	Paper Citation	15	0	40	0	323
Total	VC investment	9,677	0	64,062	0	2,573,818
	Patents	32	2	114	0	1,901
	Pension	5,857,309	2,284,637	11,800,000	15,947	130,000,000
	New Business	43	6	177	-875	4,924
	Paper Citation	536	1	2,289	0	38,647

Table 2.2 - VC Investment on the Top 15 Regions in 1998

Areas \ Sectors	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 2.3 - Number of Papers on the Top 15 Regions in 1998

Areas \ Sectors	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 2.4 - Estimation Results Sample 1981-1997 including all 179 BEA regions and 7 tech sectors. Poisson Regressions (with bootstraped standard deviations)

Dependent Variable: Patents Unassigned or Assingned to Companies with no COMPUSTAT code			
Independent Variables	**Model 1**	**Model 2**	**Model 3**
Citation of Papers	2.43	7.80	1.87
	(0.57)***	(1.39)***	(0.19)***
VC Investments	2.83	1.75	5.05
	(1.27)**	(0.23)***	(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.	17085	17436	18258
Number of Groups	1005	2906	179
Log likelihood	-41459.71	-48688.72	-64018

All variables are normalized by the repective BEA population. The models are posison panels: prob(nit)=exp(λit).λit/(nit!) where, Model 1: logλit = μij + Xijtβ ; Model 2: logλit = μit + Xijtβ; Model 3: logλit = μi + γj + Xijtβ

*Statistical significant at 10%, **Statistical significant at 5% ***Statistical significant at 1% 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). Number of Patents are from the NBER patent dataset. For this study we used only patents with COMPUSTAT code in an attempt to isolate patent from non listed companies. Citations of papers is from the Thomson Scientific Science Indicator. This product is database of research performance measures on 284 research universities in the United States. The variable used in the above regression 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.

Table 2.5 - Estimation Results Sample 1981-1997 including all 179 BEA regions (except Boston and San Francisco Areas) and 7 tech sectors.
Poisson Regressions (with bootstraped standard deviations)

Dependent Variable: Patents Unassigned or Assingned to Companies with no COMPUSTAT code			
Independent Variables			
	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.	16881	17232	18054
Number of Groups	993	2872	179
Log likelihood	-38699.37	-145461	-57485.768

All variables are normalized by the repective BEA population. The models are posison panels: prob(nit)=exp(λit).λit/(nit!) where, Model 1: logλit = μij + Xijtβ ; Model 2: logλit = μit + Xijtβ; Model 3: logλit = μi + γj + Xijtβ

*Statistical significant at 10%, **Statistical significant at 5% ***Statistical significant at 1% 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). Number of Patents are from the NBER patent dataset. For this study we used only patents with COMPUSTAT code in an attempt to isolate patent from non listed companies. Citations of papers is from the Thomson Scientific Science Indicator. This product is database of research performance measures on 284 research universities in the United States. The variable used in the above regression 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.

Table 2.6 - Estimation Results Sample 1981-1997 including all 179 BEA regions and 7 tech sectors Using Number of Papers instead of Citations. Poisson Regressions (with bootstraped standard deviations)

Dependent Variable: Patents Unassigned or Assingned to Companies with no COMPUSTAT code			
Independent Variables	**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.	17085	17436	18258
Number of Groups	1005	2906	179
Log likelihood	-42048.87	-151325	-64863.959

All variables are normalized by the repective BEA population. The models are posison panels: prob(nit)=exp(λit).λit/(nit!) where, Model 1: logλit = μij + Xijtβ ; Model 2: logλit = μit + Xijtβ; Model 3: logλit = μi + γj + Xijtβ

*Statistical significant at 10%, **Statistical significant at 5% ***Statistical significant at 1% 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). Number of Patents are from the NBER patent dataset. For this study we used only patents with COMPUSTAT code in an attempt to isolate patent from non listed companies. Citations of papers is from the Thomson Scientific Science Indicator. This product is database of research performance measures on 284 research universities in the United States. The variable used in the above regression 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.

Table 2.7 - 2SLS estimation for Supply and Demand of VC Investment. Estimation Results for the sample 1993-1999 including all 179 BEA regions and 7 tech sectors. Robust Standart Deviation on the 1st and Bootstraped on the 2^{nd}.

Dependent Variable: Patents Unassigned or Assingned to Companies with no COMPUSTAT code

Independent Variables	**First Stage**	**Second Stage**
Citation of Papers (by BEA population)		1.00
		(0.44)**
Capital Gain Taxation		2.27
		(0.70)***
Total Real Assets of Local and State Pension Funds (by BEA population)	0.001	
	(0.0003)***	
Region/Sector Fixed Effects	Yes	Yes
Time Fixed Effects		
Log-Likelihood		
Number obs.	5370	4485

The 2SLS regressions: the first stage is a linear panel of VC investment on Pension funds assets with region+sector fixed effects (robust standarddeviations in parentesis). The second stage is a poisson panel of the predicted value from the first stage on the citation of papers (boostraped standard deviations in parentesis).

Here we use the Total Real Assets of Local and State Pension Funds is from the US Census Bureau as an instrument for VC investment. The idea is that, for political reasons, the local and state funds tend to have more investment at the local level. Citation of papers is from the Thomson Scientific Science Indicator. This product is database of research performance measures on 284 research universities in the United States. The variable used in the above regression is the number of times the papers were cited through 2003. In order to mitigate the truncation problem of the citation dataset the total number of citation is divided by the time the patent or paper could have received citations.Total Real Assets of Local and State Pension Funds is from the US Census Bureau. The variable used in the regression is the total assets of local and state pension funds deflated by the Producer price index. Because the pension fund assets are only available at the state level, when a BEA area lies in more than one state, I use the sum of assets over these states. The sample is from 1993-1999.
*Statistical significant at 10%, **Statistical significant at 5% ***Statistical significant at 1%

Table 2.8 - Estimation Results for New Business regressions. Sample 1981-1997 including all 179 BEA regions and 7 tech sectors.

Dependent Variable: Number of New Business						
Independent Variables	**(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

All variables are normalized by the repective BEA population. We collapse all variables on on the region and sectors dimensions by taking time series average on the sample 1988-1997. The model is a linear fixed effect panel: Number New Businessij= μi + γj + β1Patentsij + β2Papers + β2 VC investment + εij The data for new business 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).

Number of Patents from individuals and unassigned patents are from the NBER patent dataset. For this study it is used the total number of patents in US application year between 1980 and 1998. Citations of papers is from the Thomson Scientific Science Indicator. This product is database of research performance measures on 284 research universities in the United States. The variable used in the above regression 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. *Statistical significant at 10%, **Statistical significant at 5% ***Statistical significant at 1%

Appendix B, Figures

Figure1.1 - US Share of World VC Investment