Venture Capital and the Diffusion of Knowledge

Juanita González-Uribe*
Columbia GSB
JOB MARKET PAPER

This version: February 6, 2013

Abstract

This paper estimates the effect of Venture Capital (VC) on the diffusion of knowledge. I compare citations to patents invented in VC-backed companies to those of comparable patents invented elsewhere. To isolate the causal effect, I exploit time variation in the assets of state pension funds that allocate capital to VC. This variation provides a valid instrument if the effect of changes in innovation opportunities within a state is uniform across local patents in the same technology-class and vintage-year. I find that after VC financing annual citations to a given patent increase 19% relative to the citations of comparable patents. Additional results are consistent with two mechanisms: VCs certify the value of patents to the general public and facilitate communication among companies in their portfolios.

*Columbia Graduate School of Business, e-mail: jgonzalez14@gsb.columbia.edu. I am indebted to Morten Sørensen for his invaluable guidance and encouragement. I am deeply grateful to Bruce Kogut, William Greene, Daniel Paravisini, and Daniel Wolfenzon for their suggestions. I thank Camilo García-Jimeno, Victoria Ivashina, Andrés Liberman, Daniel Osorio, Amit Seru, Pian Shu, Scott Stern, José Tesada, Arvid Ziedonis, and Rosemary Ziedonis for insightful comments. I am grateful for helpful comments from seminar participants at the FED/NYU Stern Conference on the Role of Private Equity in the U.S Economy, the 5th Annual Searle Center Conference on Innovation and Entrepreneurship, the 9th Annual Conference of Corporate Finance-Olin Business School University of Washington at St. Louis, Columbia University, Universidad de los Andes Colombia, Universidad Católica de Chile, Duke University, University of North Carolina at Chapel Hill, University of Hawaii at Manoa, and Harvard Business School. Funding for this work was provided by the Ewing Marion Kauffman Foundation.

Does the diffusion of knowledge depend on the environment in which ideas are developed? This paper explores this question by examining how the diffusion of an idea is affected by Venture Capital (VC) financing of the company that patented the idea. Venture Capitalists (VCs) invest in privately held innovative business. In addition to providing capital, they are generally believed to contribute value in other ways (e.g. Hellman and Puri, 2000; 2002). In this paper, I show that VC financing has a positive, causal effect on the diffusion of patented knowledge. The empirical evidence points to two mechanisms: VCs facilitate communication among companies in their portfolios, and more broadly, VC financing appears to certify the value of innovations to the general public.

I use patent citations to measure knowledge diffusion (Jaffe, 1986; Hall et al., 2001; Jaffe and Trajtenberg, 2002). Legislation requires inventors to cite all previous patents that their inventions build upon. Subject to caveats, discussed below, these citations are an indirect measure of knowledge linkages between innovations (Hall, et al. 2003). To distinguish the effect of VC financing on knowledge diffusion from its effect on knowledge production, I study a sample of patents invented in companies before they are VC financed. I compare subsequent increases in citations to these patents to the citations of comparable patents in the same technology-class and vintage-year, and not invented in VC-backed companies. The comparison focuses on knowledge diffusion outside company boundaries, and only includes citations from inventors outside the patenting company. My first finding is that after VC financing citations to a given patent increase by 19% relative to the citations of comparable patents.

The first finding suggests that the diffusion of already existing, disclosed ideas increases with VC ownership. While this result is interesting, one concern is the endogeneity of VC investments. For instance, VCs may anticipate which existing innovations will be dominant and cited in the future. Alternatively, VC financing may increase awareness of innovations and affect future citations. To isolate the causal effect, I use time-series variation in the assets of state public pension funds as an instrumental variable (IV) (Mollica and Zingales, 2007). This IV approach relies on the home-bias of state pension funds in their VC investments (Hochberg and Rauh, 2012), and on the exclusion restriction that changes in pension assets are independent of the innovation opportunities facing the companies. One potential concern with this exclusion restriction is that unobserved economic activity at the state level may affect both the size of state pension funds and the innovative opportunities of local companies. Since the analysis compares citations to patents filed by VC-backed companies to those of comparable patents, the exclusion restriction is satisfied, as long as the effect of unobserved economic activity on innovation opportunities within a state is uniform across local patents in the same technology-class and vintage-year.[1] As a robustness check, I relax this identification assumption by eliminating citations directly linked to local innovation opportunities and only counting citations from inventors in states other than the home-state of the patent. Using this IV approach, I find evidence that the effect of VC financing on patent citations is causal.

[1] For example, it assumes that natural gas shale discoveries affect citations to all Hydraulic Fracture patents filed in 1995 and developed in California in a similar manner.

The second part of this paper explores some mechanisms driving the effect of VC financing on patent citations. One potential mechanism is that VC financing increases awareness of companies' innovations, possibly certifies their value, and spurs follow-on innovation by other inventors. In addition, VCs may also facilitate communication among companies in their portfolio. To test these mechanisms, I distinguish between two types of citations: those from inventors in companies financed by the same VC, portfolio-linked, and those from all other unrelated inventors, non-portfolio-linked. Consistent with the first mechanism, I find a causal increase in non-portfolio-linked citations. Consistent with the second mechanism, I find that the increase in portfolio-linked citations is four times stronger than the increase in non-portfolio-linked citations. I also analyze inventor mobility and patent sales around the financing event as potential channels behind the VC financing effect on patent citations. Inventors may choose to move to other companies after VC financing, for example, if the presence of VC investors implies a transition from creative freedom to a commercial focus (a la Aghion et al., 2008). This inventor mobility can facilitate knowledge flows between inventors' new and old employers. Also, companies may sell patent outside their core areas after VC financing and directly transfer knowledge to buyers. My findings suggest, however, that the effect of VC financing on patent citations is not driven by either of these two channels.

The last part of this paper addresses concerns about the relationship between the dependent variable in the analyses, patent citations, and what I really want to measure, knowledge diffusion. For example, patent reviewers may also become aware of a company after it is VC financed. Since citations from patent reviewers are included in the analysis, citations may increase when there is no diffusion of knowledge. I test this alternative story using a sub-sample of patents for which I can distinguish the citations added by patent reviewers and exclude those from the analysis. Results remain qualitatively similar, which minimizes concerns regarding the interpretation of patent citations as knowledge flows. I consider and test other alternative stories.

This paper contributes to the literature that relates the diffusion of innovation to the institutional environment in which new technology is developed (Mokyr, 2003; Gans, Murray and Stern, 2010; Williams, 2011; Gans and Murray, 2012). I extend this literature by focusing on the diffusion of already patented innovation and showing that conditional on disclosure the institutional setting also matters for diffusion.

This paper also relates to the literature that considers the role of VC on innovation (Kortum and Lerner, 1998; Hirukawa and Ueda, 2008; Nanda and Rhodes-Kropf, 2011). I offer a new approach to investigate this question by using micro-level data and by focusing on knowledge diffusion. A back-of-the-envelope calculation based on the findings suggests that by facilitating the diffusion of their companies' patents, VCs have contributed 2% to 10% of patent production in the U.S. This finding helps explain why researchers using industry-level data estimate that VCs contribute to 14% of patent production (Kortum and Lerner, 1998) even though less than 4% of patents have been assigned to VC-backed companies. I argue that at least part of this difference can be attributed to knowledge spillovers generated by VCs.

Finally, the paper also relates to the literature that explores non-financial services VCs provide to their companies. Previously documented mechanisms include recruiting key managers (Hellmann and Puri, 2002), implementing strong governance mechanisms (Hochberg, 2011), and facilitating strategic alliances (Lindsey, 2008). I find evidence that VCs help diffuse knowledge across companies in their portfolio. Consistent with Hellmann (2002), my findings suggests that VC portfolios change the complementary assets available to companies. Since patent citations have been shown to be associated with value (Hall, 2005), this non-financial service of VCs may have value implications for VC-backed companies.

The rest of this paper is organized as follows. Section 1 explains the data sources used to construct the sample and presents summary statistics. In Section 2, I discuss the empirical strategy used to identify the effect of VCs on knowledge diffusion and present results. Section 3 explores the mechanisms behind this effect. Section 4 discusses patent citations as a measure of knowledge flows and considers alternative interpretations. Section 5 concludes. Finally, the Appendix presents a detailed account on the procedure used in the construction of the sample.

1 Data Description and Summary Statistics

The data on VC investments are from SDC's VentureXpert. Companies headquartered in the United States (U.S.) and financed by U.S.-based VC firms from 1976 to 2008 are identified. Data on patents comes from the Harvard Business School (HBS) patent data-base, which has information on U.S. patent assignments from January 1976 through December 2008 based on the records from the U.S. Patent and Trademark Office (USPTO). I combine the two data sources by searching for each of the VC-backed company names among the patent assignees. The Appendix has a detailed account of the matching procedure and includes summary statistics for the matched sample.

To distinguish the effect of VC financing on knowledge diffusion from its effect on knowledge production, I restrict the data to patents filed by companies at the USPTO at least two years before they are financed by a VC.[2] Since my empirical strategy explores subsequent changes in citations to these patents, I only consider companies that were financed by VCs between 1977 and 2003, so I can observe at least two years of citations before VC financing, and five years afterwards. After these restrictions, the analysis sample consists of 2,336 patents filed by 752 companies.

Table 1 presents summary statistics of the analysis sample and explores its representativeness of all patents that are assigned to VC-backed companies, and of all companies financed by VCs. Panel

[2] There are two dates associated to patents that are relevant for this study: the application-year and the grant-year. The application-year corresponds to the year in which inventors file their patents at the USPTO. The grant year corresponds to the year in which the USTPO grants the patent to the inventor. The lag between these two dates is on average 2 years, and is not is not statistically different for patents invented by companies with and without VC-investors. In unreported results I restrict the sample to patents granted at least two years before they are financed by a VC. Results are robust to this change.

B shows that the analysis sample is slightly more concentrated in Massachusetts, Pennsylvania, and Texas (Panel B). Also, the sample is composed of relatively more mature (Panel C) and successful (Panel D) companies from industries that rely on patents to protect their Intellectual Property (IP), such as medical health and semiconductors (Panel E).

Using these patents, I construct a data-set at the patent-year level where the variable of interest is the annual number of citations received by patents from their application year until 2008. Since the analysis focuses on knowledge diffusion outside company boundaries, I only include citations from inventors outside the patenting company. Panel G of Table 1 shows summary statistics of annual citations. Reflecting the well-known skewness in patent citation data, mean and median annual citations are 0.92 and 0, respectively. I also classify citations by state using data on the geographical location of the citing inventors. Panel G shows summary statistics of out-of-state citations, which exclude citations from inventors in the home-state of the companies that filed the patents.

1.1 Citation baseline

Patent citation rates have been increasing over time and tend to vary according to technology-class and vintage-year (Hall et al., 2001). To control for these aggregate trends and for patent life-cycle effects in the analysis, I define a set of comparable patents as follows. For every patent in my sample I determine all U.S. patents assigned to the same USPTO technology-class, with the same application-year, and that were not filed by a VC-backed company.[3] At present, the USPTO has assigned more than 400 technology-classes, examples of which include Radio Wave Antennas and Wheel Substitutes for Land Vehicles.

Using the comparable patents, I construct an annual citation baseline as:

$$b_t = \frac{Total\ Citations_t}{Number\ of\ Comparable\ Patents}, \tag{1}$$

where $Total\ Citations_t$ corresponds to citations received by comparable patents at time t. Panel G in Table 1 reports summary statistics of the citation baseline. On average, the patents invented in VC-backed companies receive 0.32 more annual citations than comparable patents. Panel G in Table 1 also reports summary statistics of a citation baseline at the state level, in which the comparable patents are additionally restricted to have been invented in the home-state of the VC-backed company that filed the corresponding sample patent.

[3] In unreported results I use the grant-year as vintage-year, and also, both the application- and grant- year. Results remain robust to these alternative definitions. However, following Hall et al (2001), I use application-year to avoid noise from the review process at the USPTO.

1.2 Restricted Sample

I collect information on financial assets held by state and local public pension funds from the State and Local Government Public-Employee Retirement Systems annual survey. This survey is conducted by the Census Bureau and is available starting in 1993. The 1993-2008 period is referred to as the restricted sample throughout, and corresponds to the sample used in the IV analysis of Section 2.3.

Table 2 reports summary statistics on the restricted sample, which consists of 1,657 patents filed by 517 VC-backed companies. Panel B reports the value of the assets held by local and state public pension funds deflated by the Producer Price Index (PPI) and expressed in billions of 1982 U.S. dollars. Panels B, C, D, and E show that the restricted sample is fairly representative of the analysis sample. The main difference is that the restricted sample is slightly overrepresented in early stage and biotech companies. Finally, Panel G in Table 2 reports descriptive statistics on annual citations to patents in the restricted sample, the annual citation baseline, and the annual citation baseline at the state level. Compared to the analysis sample, average annual citations to patents increase for the restricted sample, reflecting the overall increase in citations throughout the period.

2 Empirical Analysis

2.1 Univariate Tests

Table 3 presents preliminary evidence that citations to patents increase after companies are VC financed. On average, patents are cited 0.64 times a year before VC financing. After VC financing, however, average annual citations increase by 63% to 1.04. This percentage increase is summarized by the Ratio of 1.63 reported in Table 3. The average annual citation baseline also increases, which illustrates aggregate citation trends. After controlling for these trends, the estimated percentage increase in citations after VC financing decreases from 63% to 33%. This adjusted percentage increase is summarized by the Ratio of Ratios of 1.33 reported in Table 3.

Table 3 also shows that even before VC financing annual citations to patents are on average significantly higher than the citation baseline. This difference does not invalidate the use of the citation baseline to control for aggregate trends in citations at the technology-class and vintage-year level. The key assumption is that citation trends, and not necessarily the levels, would be similar across patents and comparable patents in the absence of VC financing. I return to this assumption in the next section.

2.2 Poisson Regressions

Citation data are non-negative and discrete, thus, I use a Poisson model, which is the standard model for count data (Cameron and Trivedi, 1998).[4] I estimate the following equation:

$$E\left[Cites_{pt}|VC_{pt}, b_t\right] = \exp\left(\alpha_p + \ln(b_t) + \beta VC_{pt}\right), \tag{2}$$

where the expected number of citations received by patent p in year t, $Cites_{pt}$, is an exponential function of a dummy variable, VC_{pt}, which equals one after VC financing. I include a full set of patent fixed-effects in the estimation, α_p, which absorb all time-invariant patent heterogeneity. To control for aggregate trends in citations, I offset the citation baseline, b_t, in the estimation. This is implemented by including in the Poisson regression the logarithm of b_t with a coefficient fixed to one.[5] To understand the intuition behind this approach, note that in equation (2) the expected number of citations received by patents absent VC financing equals the citation baseline times a constant adjustment.

Table 4 reports the results from the Poisson analysis. All reported coefficients are incidence rates and reflect the proportional increase of annual citations to an increase in the explanatory variable. An incidence rate greater than one corresponds to a positive effect of the explanatory variable on annual citations to patents. An incidence rate below one corresponds to a negative effect. Correspondingly, indications of statistical significance do not reflect whether the coefficients are different from zero, as is usual, but rather whether they differ from one. Finally, since all patents filed by the same company are subject to company-specific shocks, standard errors are clustered at the company level.

Column (1) in Table 4 reports the results from a pooled Poisson regression of equation (2) excluding the citation baseline. I estimate the model using maximum likelihood (MLE). The interpretation of the coefficient for VC_{pt} is that annual citations to patents increase 62.7% after VC financing. Note the correspondence between the estimated coefficient and the Ratio reported in Table 3.[6]

Column (2) in Table 4 summarizes the results from a pooled Poisson regression of equation (2) by MLE. After controlling for aggregate trends, the estimated increase in citations declines. The interpretation of the coefficient for VC_{pt} in Column (2) is as follows: after VC financing, annual

[4] Another common model for count data is the Negative Binomial model which is a generalization of the Poisson model that addresses overdispersion by including an additional error term to capture unobserved factors. In unreported analyses I replicate the analsysis using this model. Results hold and are not statistically different across models.

[5] The baseline removes any aggregate annual variation. This technique is similar to including time-fixed effects (cross technology-class and vintage-year) in the estimation.

[6] The estimated constant in Column (1) of Table 4, corresponds to the average annual citations to patents before VC financing reported in Table 3.

citations to patents increase 34.6% in excess of the citation baseline. Note the correspondence between the estimate in Column (2) and the Ratio of Ratios reported in Table 3.[7]

Column (3) in Table 4 presents results from the fixed-effects Poisson model. One concern in using fixed-effects in non-linear models is that estimates may be inconsistent because of the incidental parameters problem.[8] I follow the literature and estimate the model using conditional quasi-maximum likelihood (QMLE) as developed by Hausman et al. (1984), which eliminates the patent fixed-effects by conditioning on $\sum Cites_{pt}$ (a sufficient statistic of α_p).[9] The interpretation of the coefficient for VC_{pt} is that annual citations to the same patent, in excess of the citation baseline, increase by 18.9% after VC financing of the issuing company.

In unreported results, I repeat the analysis of Table 4 excluding California, Massachusetts, and Texas, and restricting the sample to the pre- and post-dotcom periods. The effect is not statistically different across sub-samples. I also examine the heterogeneity of the results by patent age. I find that the increase is highest for patents younger than five years, but the effect is also positive and significant for patents between five and ten years of age. In addition, I explore the heterogeneity of results by VC skill as measured by the number of prior successful financing rounds. I find that more experienced VCs have a greater effect on citations, but this stronger effect is not statistically significant. I distinguish between the extensive and intensive margins. I find that conditional on having been cited before VC financing, the increase in citations to patents is not statistically significant. In contrast, for patents with no prior citations, the effect is large and significant. Furthermore, I experiment with different definitions of the baseline by excluding from the set of comparable patents those that are never cited throughout the sample, those that originate in large companies, universities, and the public sector [10], and by adding geographical restrictions using the citation baseline at the state level defined in Section 1.1. Results are quantitatively similar across the different versions of the baseline.[11] Finally, I also explore the effect of VCs on the dispersion of citations across technology classes (i.e., the generality measure of Hall et al., 2001), and find no significant effect.

Figure 1 explores the dynamics of the effect uncovered in Column (3) of Table 4. I estimate a fixed-effects Poisson model where the independent variables are indicators for individual years relative to the year of VC financing, and restricting observations to two years before–, and five years

[7] The constant in Column (2) corresponds to the ratio between average annual citations to patents and the average citation baseline before VC financing reported in Table 3.

[8] The fixed-effects Poisson model, however, is one of the few models for which consistency of the MLE holds despite the presence of incidental parameters (Cameron and Trivedi, 1998).

[9] The Poisson model is in the linear exponential family and the coefficient estimates remain consistent as long as the conditional mean is correctly specified (Wooldridge, 1999). In the estimation, therefore, I do not assume that the mean and the variance are equal, or arbitrary independence across observations. Instead, I compute the variance-covariance matrix using the outer product of the gradient vector and clustering the standard errors at the company level.

[10] I thank Scott Stern for this suggestion.

[11] In future versions of the paper I may report main results using more restrictive baselines as precision generally increases.

after–, the VC financing event (Event-year 0 is omitted from the estimation to avoid multicollinearity with the patent fixed-effect). Figure 1 plots the estimated coefficients (solid line) together with their 95% confidence interval (dashed lines). Before VC financing, citations to the same patent in excess of the baseline are not statistically different from those in the year of VC financing. This pattern is reflected in the estimated coefficients of the dummy variables indicating event-years pre VC financing: neither is statistically different from one. This result is reassuring, as it shows that the subsequent increase in citations is not driven by a pre-existing trend in citations. In contrast, the estimated coefficients of the dummy variables indicating event-years post VC financing are all larger than one, and significant from event-year 2 onwards. Note that although not significant, the point estimates for the dummy variables indicating event-years pre VC financing are actually negative, which suggests that the patents in the sample were relatively unknown before the issuing companies are financed by a VC. This pattern is consistent with the aforementioned results on the extensive and intensive margins.

One interpretation of the temporal patterns in Figure 1 is that since there is no pre-trend, the increase in citations reflects the causal effect of VC financing. Alternatively, VCs may be able to anticipate which innovations are more likely to be dominant in the future, and the increase in citations at least partially reflects VCs' ability to time their investments. Regardless of the interpretation, and given that citations are associated with value (Hall et al., 2005), the results help visualize why VCs can command the 2/20 compensation scheme[12]. Even if the increase in citations only reflects the skill of VCs in timing their investments, this is interesting from a financial perspective, as it means that VCs can cherry pick projects before any other agent in the market. This ability to pick projects is not easy to imitate, and consequently translates into high returns. From the point of view of policy, however, it is important to disentangle the two interpretations because the policy implications are very different.

2.3 Addressing endogeneity of VC investments

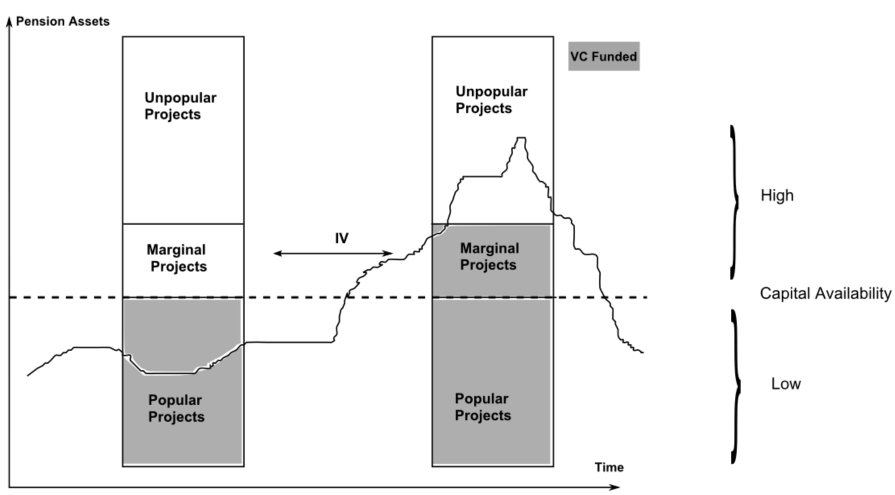
To isolate the causal effect, I use time-series variation in the assets of state public pension funds as an instrumental variable (IV) (Mollica and Zingales, 2007). This IV approach relies on the home-bias of state pension funds in their VC investments (Hochberg and Rauh, 2012), and on the exclusion restriction that changes in pension assets are independent of the innovation opportunities facing the companies. One concern with this exclusion restriction is that unobserved economic activity may affect both pension assets and companies' innovation opportunities. Since the analysis compares citations to patents filed by VC-backed companies to those of comparable patents, the main identification assumption is that the effect of unobserved economic activity on innovation opportunities within a state is uniform across local patents in the same technology-class

[12] The 2 refers to the 2% management fees the VCs charge to cover operating expenses and salaries. The 20 refers to the 20% carry fee– the percentage of the profits they make for their investors that they get to keep.

and vintage-year. In this section, I explain the IV approach in detail.

2.3.1 Intuition

The intuition behind the IV approach is best explained following the same logic as the local average treatment effect (LATE) of the linear literature. Start by assuming that VCs select which companies to finance based on the unobserved and heterogenous future popularity of their patents. Every year companies are classified into three classes: popular, marginal, and unpopular. Popular (unpopular) companies are those for which the future popularity of their patents is high (low) and will (will not) be funded irrespective of the availability of capital for VCs. Marginal companies, with average popular patents, are funded only if the availability of capital for VCs is sufficiently high. For simplicity, assume that every period the availability of capital for VCs can be either high or low. If there is high availability of capital, VCs finance their marginal companies, otherwise, marginal companies are not funded. The IV approach is equivalent to comparing the average outcome for marginal patents across periods of high and low availability of capital for VCs. The figure below illustrates the example.



This example helps clarifies common misconceptions of IV. For instance, the estimation strategy does not assume that given high capital availability VCs randomly pick the companies they finance. Instead, the identification strategy relies on two assumptions. First, that the size of state pension assets is indicative of local availability of capital for VCs. Second, that the average quality of companies (and their patents) faced by VCs within a state is comparable across periods of high and low pension assets. In the next sections, I discuss these two assumptions in detail.

2.3.2 State pension assets and local investments by domestic VC firms

The validity of the first identification assumption stems from the home-bias of state pension funds in their VC investments, and the home-bias of VCs in their financing of portfolio companies. Hochberg and Rauh (2012) document that public pension funds display a 23 percentage point home-state overweighting in VC investments. On the home-bias of VC financing, there is abundant colloquial and formal evidence (see Lerner, 1995; Chen et al., 2009; Douglas and Na, 2010). To provide additional suggestive evidence, I collect information from VentureXpert on total value of investments by VC firms (both inside and outside their home-state) and estimate the following equation:

$$Investment_{st} = \alpha + \theta_s + \gamma_t + \beta Pension_{st-1} + \eta_{st} \tag{3}$$

where $Investment_{stL}$ is the value of investments by VCs headquartered in state s at year t (deflated by PPI). $Pension_{st-1}$ corresponds to the assets of local and state public pension funds in state s, deflated by the PPI and lagged by one year[13]. I include in the estimation state fixed-effects, θ_s, which control for the time-invariant importance of VC investments in states. I also include time fixed-effects, γ_t, which control for aggregate trends.

Table 5 summarizes results from estimating equation (3) and clustering standard errors at the state level. The interpretation of the coefficient in Column (1) is as follows: an increase of $1 billion in state pension assets increases the value of VC investments by $52 million. The second (third) column uses as the dependent variable the value of investments by VCs in local (non-local) companies. The interpretation of the coefficient in the second (third) column is as follows: an increase of $1 billion in state pension assets increases the value of VC investments in local (non-local) companies by $36 ($16) million. Columns (4)-(6) replicate the analysis of Columns (1)-(3) restricting the dependent variable to VC financing of *new* companies. The results are robust to this restriction. Overall, this table shows that the size of state pension assets affects the availability of capital for VCs.

2.3.3 Local innovation opportunities and public pension assets

The second identification assumption is the exclusion restriction and cannot be tested.[14] To examine its validity, consider the three main sources of variation in the size of state pension assets: demographic conditions, pension policy, and returns to past investments. The first two are deter-

[13] The process for VCs of raising a fund and beginning to deploy capital takes about one to two years.

[14] As suggestive evidence, in unreported regressions I ran a placebo test where I test whether the correlation between pension assets and relative citations to patents in the sample is significant for all periods during the life of a patent. I find that the correlation is only significant while the VC is an investor in the company.

mined by broader socioeconomic considerations other than current innovation opportunities and are unlikely to raise any concerns. Returns to past investments, however, may reflect unobserved economic activity at the state level that can affect both the size of state pension funds and the innovative opportunities of local companies. Since for every patent filed by a VC-backed company the citation baseline includes citations to comparable patents not necessarily invented in the same state, changes in innovation opportunities within a VC-backed company's state may affect disproportionately citations received by its patents relative to the citation baseline. This disproportional effect could imply that the average quality of patents faced by VCs, as measured by citations in excess of the baseline, across periods of high and low pension assets is not comparable. This lack of comparability would raise concerns regarding the exclusion restriction.

To address this concern, I use the citation baseline at the state level defined in Section 1.1. The identification assumption is that the effect of unobserved economic activity on innovation opportunities within a state is uniform across local patents in the same technology-class and vintage-year. As an additional robustness check, I relax this identification assumption by eliminating citations directly linked to local innovation opportunities and only counting citations from inventors in states other than the home-state of the patent (out-of-state citations).[15]

2.3.4 Econometric Considerations

I now turn to a rigorous econometric treatment of the IV approach. I start by noting that to estimate the fixed-effects Poisson model using IV one may think to follow the conventional approach in the linear literature. Regress the endogenous variable VC_{pt} against the instrument and the patent fixed-effects, and use the predicted value from that regression as a regressor in the fixed-effects Poisson model instead of VC_{pt}.[16] This approach, however, is not valid because the expected value operator does not pass through non-linear functions. Instead, to estimate the non-linear IV I follow Wooldridge (1997) and Windmeijer (2000) and use a quasi-differentiation of the fixed-effects Poisson model, together with the implied exclusion restriction of the IV, to derive moment conditions, which I estimate using the generalized method of moments (GMM-IV hereafter).

The main intuition behind this approach is as follows. Recall from Section 2.1 that although the fixed-effects Poisson model does not suffer from the incidental parameters problem, the fixed-effects are still eliminated for convenience in the estimation. In this sub-section, I follow the same principle using an approach similar to the within transformation in the linear literature. The starting point is the fixed-effects Poisson equation (2). To simplify, let $x_{pt} \equiv \ln(b_t) + VC_{pt}$ and $B \equiv [\beta \; 1]$, and consider the following transformation that eliminates α_p as suggested by Wooldridge (1997):

[15] See Table 1 Panel G for summary statistics on out-of-state citations.

[16] This is an example of the so-called forbidden regressions.

$$\frac{Cites_{pt}}{\exp\left[x_{pt}\right]B} - \frac{Cites_{pt+1}}{\exp\left[x_{pt+1}\right]B} = \alpha_p\left(\varepsilon_{pt} - \varepsilon_{pt+1}\right), \tag{4}$$

Let $Pension_t$ correspond to the assets of local and state public pension funds at year t in the home-state of the company that filed patent p. Combining equation (4) with the exclusion restriction, i.e., $E\left[\varepsilon_{pt}|Pension_t, ..., Pension_1, b_t, \alpha_p\right] = 1$, gives the following moment conditions:

$$E\left[\frac{Cites_{pt}}{\exp\left[x_{pt}\right]B} - \frac{Cites_{pt+1}}{\exp\left[x_{pt+1}\right]B} \mid Pension_{t-1}, b_t\right] = 0. \tag{5}$$

Using the moment conditions (5) generally causes computation problems. For example, when the regressors include dummy variables such that the moment conditions can be made close enough to zero by choosing arbitrarily large βs. To address this concern, I follow Windmeijer (2002), and multiply through by $\exp(\mu_x\beta)$, where $\mu_x = (NT)^{-1}\sum\sum x_{pt}$. This adjustment minimizes the computational problem because the deviated variables, $x_{pt} - \overline{x}$, will always take on positive and negative values.

The modified moment conditions I estimate to calculate the GMM-IV estimator are the following:

$$E\left[\frac{Cites_{pt}}{\exp\left[x_{pt} - \mu_x\right]B} - \frac{Cites_{pt+1}}{\exp\left[x_{pt+1} - \mu_x\right]B} \mid Pension_{t-1}, b_t\right] = 0. \tag{6}$$

2.3.5 IV results

Table 6 presents the basic IV results. I begin by providing the fixed-effects Poisson estimates using the restricted sample of the GMM-IV approach in Column (1).[17] The estimated coefficient remains positive and statistically significant, and is not significantly different from the basic Poisson estimate using the analysis sample in Table 4. The GMM-IV estimator using $Pension_{t-1}$ to instrument for VC_{pt} is reported in Column (2). A comparison between Columns (1) and (2) reveals that the estimated effect increases from 17.1% to 49.5% after accounting for non-random timing of VC investments. The difference between the two estimates, however, is not statistically significant.

Table (6) also reports results using a standard linear IV approach (2SLS). As the dependent variable, I use scaled citations defined as the ratio between citations to patents and the citation baseline. As shown in Columns (5)-(6), the results are similar. The second panel in Column (6) reports results from the first-stage, where I regress the endogenous variable, VC_{pt}, on the instrument and on patent fixed-effects. The F-statistic suggests that the instrument is unlikely to

[17] Recall that the data on state pension assets is only available at the Census starting on 1993. For details, see Section 1.2.

be weak (Stock and Yogo, 2005).[18]

Tables 7 and 8 present the IV approach using the citation baseline at the state level. Table 8 summarizes results of the main robustness check where I use out-of-state citations as the dependent variable to minimize concerns that unobserved economic activity may affect both VC financing and patent citations. I report standard errors clustered at the company-level as they are more conservative than those at the state level, which is suggestive of small-cluster bias. Tables 7 and 8 show that the results continue to hold and are quantitatively similar to the basic IV results in Table 6.[19][20]

As extra robustness checks, Tables 6-8 report results using alternative instruments. In each of these tables, the third (seventh) column reports estimates using $Pension_{t-1}$ normalized by average state GDP as instrument and using the GMM-IV (2SLS) approach. Similarly, the fourth (eight) column reports estimates using $Pension_{t-1}$ demeaned at the time-level as instrument and using the GMM-IV (2SLS) approach. The results are qualitatively similar.

In additional unreported regressions, I exclude California from the sample and results are qualitatively similar. I also experiment with different versions of the baseline by excluding from the matching patents those that originate in large companies and those that are never cited throughout the sample. I also sharpen identification by defining the citation baseline at the city level.[21] Results are quantitatively similar across the different versions of the citation baseline, and significance improves in most cases.[22] In conclusion, the effect of VC financing on patent citations is always positive, and significant across most IV specifications.

2.4 Interpretation of results

An interesting finding that emerges from Tables 6-8 is that IV estimates (GMM-IV and 2SLS) always exceed their basic (Poisson and OLS) counterparts, although the difference is never significant. If one assumes on a priori grounds that the basic approach leads to upward-biased estimates of the true causal effect of VCs, the even larger IV estimates present something of a puzzle. This puzzling result is common to all papers in the VC literature that instrument VC investments using changes

[18] The correlation in-sample between the endogenous variable and the instrument is 0.18 and is statistically significant at the 1% level.

[19] However, note that the interpretation of the results changes. To illustrate, the coefficient of Column (1) in Table 8 is interpreted as follows: after companies are financed by a VC, out-state citations to the same patent increase by 24.7%, relative to other patents in the same technology-class and vintage-year and issued in the same state.

[20] Note that the difference in observations from Tables 6, 7, and 8, is due to the fact that by restricting the dependent variable to out-state citations or/and defining relative citations at the state level, there are patents for which there is not enough variation for the fixed-effects Poisson to be estimated. Consequently, comparisons across models do not have a straightforward interpretation.

[21] I thank Amit Seru for this suggestion.

[22] When I define the citation baseline at the city level, I lose a lot of observations, thus the power in these regressions is reduced. In future versions of this paper, I may use one of these alternative versions of the baseline as the preferred set of results.

in the availability of capital for VCs [see Kortum and Lerner (2000), Mollica and Zingales (2007), and Bernstein et al. (2011) for the PE case].[23] Does this mean that VCs have no skill in selecting companies and that the increase in citations corresponds to the average treatment effect (ATE)? Not necessarily, because identification using IV is representative only of the marginal patents whose treatment is affected by the instrument (as illustrated in the figure of Section 2.3.1), and which are not necessarily representative of the general population of patents. This is the standard argument of LATE in the linear literature.

An alternative interpretation is that there is underlying heterogeneity in the effect of VC financing, and changes in the capital available for VCs trigger investments on a sub-population of patents whose diffusion is particularly responsive to VC financing. For example, if the abundance of capital allows VCs to experiment more effectively and shifts the startups that they finance towards those that are riskier and more innovative. To explore this idea, I compare the average novelty of patents from companies financed across periods of high and low availability of capital as determined by pension assets. A company is defined to have been financed in a hot (cold) market, if pension assets in the home-state of the company during the year of the VC investment are within the top (bottom) 25% of the sample. As a proxy for novelty, for each patent I construct the "originality" measure of Hall et al. (2001) as one minus the Herfindahl index of the cited patents across technology-classes.[24] The intuition is that patents that combine existing knowledge from few technology-classes to create something new (and useful) probably constitute more marginal improvements relative to patents that combine more different ideas ex-ante.

Table 9 shows that patents funded in hot markets are on average more original than those funded in cold markets.[25] The difference is statistically significant, even after controlling for the average originality of matching patents. This suggestive evidence is in line with Hirukawa and Ueda (2008) and Nanda and Rhodes-Kropf (2011), and suggests that VCs play an additional role on innovation. Not only do VCs finance the innovation of their portfolio companies and facilitate the diffusion of knowledge, they also seem to use available capital to experiment, a role that is arguably needed for the commercialization and diffusion of novel technologies (Nanda and Rhodes-Kropf, 2011).

For policy evaluation, the issue that remains regards external validity. To inform policy, however, the ATE may be less relevant than the average return for the group who will be impacted by a proposed reform (Imbens, 2009). Since an important part of growth policies seek to stimulate VC

[23] These results echo the debate in the literature of the returns to schooling, particularly the papers by Card (1994; 2001).

[24] I use both standard and adjusted originality measures. The latter is based on the bias-correction described in Jaffe and Trajtenberg (2002).

[25] Graphically, Table 9 compares the shaded regions of the rectangles in the Figure of Section 2.3.1 across periods of high and low pension assets. The exercise of Table 9, therefore, provides suggestive descriptive evidence of the characteristics of marginal patents. A common misconception is that this finding contradicts the exclusion restriction. This is not the case. Note that the exclusion restriction is *not conditional* on VC financing, and compares the average quality of patents across periods of high and low pension assets. Graphically it compares the full rectangles, including both shaded and un-shaded regions of the rectangles.

financing via shocks to the available capital for VCs (Lerner, 2009), the results are informative for current policy.

2.5 Back-of-the-envelope calculation

Subject to caveats regarding the representativeness of the IV results discussed in Section 2.4 (i.e., LATE vs ATE), a back-of-the-envelope calculation based on the findings is that 1.6% to 10.24% of extra patent production can be attributed to VCs facilitating the diffusion of their targets' patented knowledge. The calculation is as follows. Average annual citations to patents pre-financing are 0.64. VC financing increases annual citations by roughly 20% (using the basic Poisson estimate). Assuming a patent life of 20 years, this implies that each VC-backed patent receives 2.6 extra citations because of increased diffusion caused by VCs.[26] On average, patents cite 6.5 other patents as relevant art in their applications. Assuming that every citation contributes with at least an equal share of the new innovation, and since 4% of patents have been assigned to VC-backed companies (see the Appendix), this implies that a range of 1.6% to 10.24% of extra patents in the U.S. can be traced-back to VCs facilitating the diffusion of knowledge. This finding helps explain why researchers using industry-level data estimate that VCs contribute to 14% of patent production (Kortum and Lerner, 1998) in spite of the small percentage of patents assigned to VC-backed companies. I suggest that at least part of this difference can be attributed to knowledge spillovers generated by VCs.[27]

3 Disentangling Mechanisms

Having shown that VC financing has a causal effect on patent citations, the second part of this paper turns to disentangling some of the mechanisms behind this effect.

3.1 Knowledge Diffusion and VC portfolios

One potential mechanism is that VC financing increases awareness of companies' innovations, possibly certifies their value, and spurs follow-on innovation by other inventors. A second potential mechanism is that VCs facilitate communication among companies in their portfolio. For example, VCs often organize summits where managers of their companies informally interact. Also, by actively participating in their company boards, VCs can detect technological complementarities across companies in their portfolio, and encourage their portfolio companies to communicate.

[26] I approximate the life of a patent with 20 years because for utility patents, protection lasts a maximum of 20 years after the application year (provided that renewal fees are paid). However, note that by law patents that have expired still need to be cited if they consitute relevant prior art.

[27] 4%*2.6=10.24% and 10.24%/6.5=1.6%.

To test these two mechanisms, I distinguish between two types of citations:

1. Portfolio-linked: those from inventors in other companies financed by the same VC.
2. Non-portfolio-linked: Otherwise.

Table 10 shows that average annual portfolio-linked citations to patents are 0.002 before VC financing, and increase by 305% after VC financing. In contrast, annual non-portfolio-linked citations are 0.64 before VC financing, and increase by 62% afterwards.[28] To control for changes in citation behavior and in the industry composition of companies over time, I use a similar approach as discussed in Section 2 and calculate a citation baseline by type of citation. By definition, comparable patents have no portfolio-linked citations because they are not invented by VC-backed companies. In order to classify their citations as portfolio-linked, thus, I use information on the VC-backed company that filed the sample patent. Table 10 shows that after controlling for the aggregate increase in citations using the portfolio-linked (non-portfolio-linked) citation baseline, the percentage increase post VC financing in portfolio-linked (non portfolio-linked) citations decreases from 305% (62%) to 47% (33%).

To formally test these mechanisms, I estimate Poisson models where I allow the VC investment to affect differently each type of citation. I start by estimating the following equation,

$$Cites_{pCt} = \exp\left(\sum_C \gamma_C D_C + \sum_C \beta_C VC_{pt} * D_C + \sum_C \ln\left(b_{Ct}\right)\right) \varepsilon_{pCt}, \tag{7}$$

where $Cites_{pCt}$ are citations at time t, to patent p, of type C, where $C \in \{NP, P\}$. NP and P stand for non-portfolio- and portfolio-linked, respectively. D_{NP} is a dummy variable that equals one when $C = NP$. D_P is defined analogously. b_{Ct} corresponds to the citation baseline specific to the type of citation C. By including b_{Ct} in the estimation, I control for aggregate changes in citations at the technology-class, vintage-year and type of citation level. [29] VC_{pt} is a dummy that equals one after the issuing company is financed by a VC and ε_{pCt} is an $i.i.d$ random variable (with mean equal to 1) that captures idiosyncratic multiplicative shocks at the patent-type of citation level. If portfolio- and non-portfolio-linked citations are equally sensitive to VC financing, the difference between β_P and β_{NP} should be zero. Hence, if the difference between these two coefficients is positive, I interpret this as evidence that the effect of VCs on citations is stronger inside VC portfolios. Panel A of Table 11 summarizes results. Panel B tests whether the β_Cs are statistically different, using a chi-squared test.

[28] The difference in magnitudes between portfolio- and non portfolio-linked citations reflects the small size of VC portfolios.On average, the companies in my sample join portfolios with 17 other VC-backed companies.

[29] This technique is similar to including type of citation cross time fixed-effects, since it removes any aggregate annual variation by type of citation.

Column (1) of Table 11 reports results from estimating equation (7) using a pooled Poisson model that excludes the citation baseline. Standard errors at clustered at the patent level. The interpretation of the coefficient for $VC_{pt} * D_{NP}$ ($VC_{pt} * D_P$) is that non-portfolio-linked citations to patents increase by 62.0% (305.2%) after VC financing.[30]. Panel B confirms that the difference between the estimated coefficients for $VC_{pt} * D_P$ and $VC_{pt} * D_{NP}$ is significantly different from zero. Column (2) of Table 11 reports results from estimating (7) using a pooled Poisson model. As expected, after controlling for aggregate trends the estimated effect decreases compared to Column (1). Column (3) in Table 11 summarizes results from estimating equation (7) including patent-cross-type of citation fixed-effects, which control for unobserved heterogeneity in patents and type of citations. The interpretation of the coefficient for $VC_{pt} * D_{NP}$ ($VC_{pt} * D_P$) is that after VC financing non-portfolio-linked (portfolio-linked) citations to a given patent increase by 21.5% (178.5%), relative to the citation baseline. Finally, Panel B confirms that the estimated percentage increase in portfolio-linked citations is statistically larger than in non-portfolio-linked citations.

Similar to Section 2, In Table 12 I address the concern of non-random timing of VC selection using an IV approach. The first four columns report coefficient estimates using portfolio-linked citations as the dependent variable and instrumenting the timing of VC financing using $Pension_{t-1}$. Although the variation of the instrument is at the state level, the standard errors are clustered at the patent level, because of potential small-cluster bias. The first four columns show that the effect of VC financing on portfolio-linked citations reported in Table 11 remains positive, but is no longer significant. This lack of significance is likely driven by the small number of observations used in these regressions. The last four columns of Table 12, report coefficient estimates using non-portfolio-linked citations as the dependent variable. The positive impact of VC financing on non-portfolio-linked citations is robust to controlling for non-random selection by VCs. In addition, the percentage increase in citations is still estimated to be larger for portfolio- than for non-portfolio-linked citations, but the difference is no longer significant.

In summary, consistent with VC financing increasing awareness of companies' innovations, possibly certifying their value, and spurring follow-on innovation by other inventors, I find a causal and strong increase in non-portfolio-linked citations. Consistent with VCs facilitating communication across companies in their portfolios, I find that the increase in portfolio-linked citations is four times stronger than the increase in non-portfolio-linked citations.

3.1.1 Robustness checks and extensions

Non-portfolio-linked citations may increase after VC financing without an increase in the general awareness of the innovations. For example, citations may be concentrated among companies inside

[30] The coefficients of 0.635 and 0.002 for D_{NP} and D_P respectively, represent average portfolio- and non portfolio-linked citations to patents before the VC investment. Note the correspondence of these numbers with the annual averages reported in Table 10.

the VC industry. To test this alternative hypothesis I exclude from the analysis citations from inventors in VC-backed companies. The estimated effect of VC financing is still larger than one and statistically significant, consistent with the more general certification effect of VC financing.

It has been suggested that syndication networks among VC firms matter for performance (Hochberg et al., 2007). One natural question is whether they also matter for knowledge diffusion. In unreported results, I classify non-portfolio-linked citations as Syndication-linked if the citing company is backed by at least one VC with whom one of the investors of the cited company has syndicated an investment in the past. Syndication-linked citations do not significantly increase after VC-financing, thus, I find no evidence that information is diffused within VC syndication networks.

3.2 Knowledge Diffusion and Inventor Mobility

Inventors may choose to move to other companies after VC financing. For example, if the presence of VC investors implies a transition from creative freedom to a commercial focus (a la Aghion et al., 2008) or if the decision making becomes more centralized with VC arrival (Seru, 2012).This inventor mobility can facilitate knowledge flows between inventors' new and old employers (Almeida and Kogut, 1999; Kim and Marschke, 2005; Agrawal and Singh, 2011; Azoulay, Graff Zivin, and Sampat, 2012).

To test this mechanism, I analyze inventor mobility around the VC financing event. This analysis is facilitated by the HBS data-set that includes a unique identifier for inventors after a detailed clean-up and analysis of the original patent records (Lai, D' Amour and Fleming, 2008). Using this identifier, I am able to trace individual mobility in my sample using changes in assignees through time. Overall, I have information on 11,627 inventors that work at VC-backed companies with patents, and their subsequent inventions in the same company or in other assignees. I distinguish between two types of citations:

1. Inventor-linked: those from inventors who invented a patent in the VC-backed company before VC-financing, or their co-workers.

2. Non-inventor-linked: Otherwise

Table 13 shows that even after excluding inventor-linked citations from the sample, both portfolio and non-portfolio-linked citations significantly increase after VC financing. This type of comparison is informative but is likely to be biased. One concern is that the propensity of inventors to change jobs is not constant over time. To address this concern and also control for aggregate trends in citations and patent life-cycle effects, I follow a similar approach to the one in Section 2. I construct a citation baseline based on average inventor-linked and non-inventor-linked citations to comparable patents. Column (6) shows that even after controlling for aggregate trends in inventor mobility and citations, the percentage increase in non-inventor-linked citations after VC financing is positive and significant. In Tables 14 and 15, I replicate the Poisson analysis from Section 2.2. using non-inventor-linked citations as the dependent variable. As expected from Table 13, the results continue to hold. This is evidence that the bulk of the increase in citations post VC financing cannot be linked to inventor mobility.

Note that one drawback from measuring mobility using data on patent assignments is that not all moves are observable. First, I only record movements of inventors; other workers can move and disseminate knowledge. Second, even if I focus on inventor mobility, the data are still necessarily incomplete. I can identify the movement of an inventor only if the individual invents in the new workplace. Some inventors may change jobs and enter executive positions in which they no longer apply for patents, but still influence the company's innovation. The findings imply, thus, that the effect of VC financing on citations is not fully explained by inventor mobility *observable* in the data.

3.3 Knowledge Diffusion and Patent Sales

Companies may sell patent outside their core areas after VC financing and directly transfer knowledge to buyers.[31] Prior research has shown an association between VC financing and patent trade. Katila and Shane (2005) find that patents are more likely to be licensed in industries where VC financing is prevalent.

To test this mechanism, I collect from the USPTO data on patent reassignments, which acknowledge the transfer of the rights, title, and interest in a patent. A typical assignment is characterized by a unique identifier, the patent number, the names of the buyer (i.e., assignee) and the seller, and the date in which the private agreement between the two parties was signed. After standardizing assignee names, I exclude all records for which the buyer matches the primary assignee of the patent. I also exclude records of administrative events such as a name change.

I combine the clean reassignment data to my sample using patent numbers. Table 16, Panel A, reports summary statistics. Of the 2,336 patents in the sample, 375 are sold by their primary assignees. The small number of matches reflects the size of the patent market, only 13.5% of granted patents in the U.S. are ever sold during their life-cycle (Serrano, 2010). Panel B shows that there is an increase in the probability that a patent is sold after VC financing. The result holds even after controlling for the likelihood that similar patents are sold. Figure 2 illustrates the sharp increase in the probability of a patent sale (solid line), relative to average sales (dashed line), after VC financing.

To test whether patent sales explain the effect of VC financing on citations, I split the sample of patents into two groups: those that are sold and those that are not sold by 2012. Panel C of Table

[31] Also note that patent sales can also be used by VCs to recoup their capital in case of a liquidation.

16 shows that citations similarly increase post financing for both groups of patents. Thus, although there is an increase in the likelihood that patents are traded after companies are VC financed, the subsequent increase in citations cannot be traced to this effect.

One drawback from the reassignment data is that it is not exhaustive of all forms in which a company's IP can be traded (e.g., licenses). My findings thus imply that the effect of VC financing on citations is not fully explained by patent trade that is *observable* in my data.

3.4 Discussion

Overall, this section points to two mechanisms. First, VC financing increases awareness of innovations, possibly certifies their value, and spurs follow-on innovation by other inventors. There is a causal increase in citations from inventors outside VC portfolios, which cannot be entirely traced back to inventor mobility or to patent sales. This increase in the general awareness of innovations can take several forms. VC financing may act as a certification of the quality of innovations or provide the necessary resources for companies to bring their products to market and increase their exposure. Disentangling between these channels is outside the scope of this paper. As suggestive evidence of increased awareness of companies after VC financing, I took the names of companies financed by VCs in 2006 as reported in VentureXpert and downloaded from Google Insights normalized[32] weekly hits for these names in Google from 2004 until 2011.[33] I standardized names by stripping them of punctuation, capitalization, and common acronyms. Consistent with increased exposure post VC financing, Figure 3 shows an increase in the number of hits after 2006 for the names of the portfolio companies (solid line) relative to the word "Gold" (dashed-line).

Second, VCs facilitate communication among companies in their portfolios. This effect cannot be entirely traced back to inventor turnover among companies financed by the same VC. One potential channel behind this finding, is that VCs encourage their companies to participate in research alliances (Lindsey, 2008), which are known to promote knowledge flows (Gomes-Casseres et al., 2006). Another possibility is that VCs recycle executives across the companies in their portfolio, and knowledge is diffused with top management. Disentangling between these channels is outside the scope of this paper.

4 Knowledge Diffusion and Patent Citations

While the analysis so far suggests a strong relationship between VC financing and patent citations, one concern remains. The increase in citations may be due to of a shift in the propensity to

[32] Weekly searches are divided by the maximum number of searches in the entire period and multiplied by 100.
[33] Data on Google Insights is only available starting on 2004.

cite patents issued by VC-backed companies that is stimulated by the VC financing process itself, but that is not associated with knowledge flows.

For example, patent reviewers may also become aware of a company after it is VC financed. Since citations from patent reviewers are included in the analysis, citations may increase when there is no diffusion of knowledge. I test this alternative story using the sub-sample of patents filed on 2001 for which I can distinguish the citations added by patent reviewers and exclude those from the analysis. Unreported results remain qualitatively similar (although power is significantly reduced), which minimizes concerns regarding the interpretation of patent citations as knowledge flows.

Another nuanced view is that potential targets may strategically cite patents issued by VC-backed companies in order to attract VC finance. To address this concern, I use investments by VCs in public companies as an informal test. Since companies that are public are subject to close monitoring and information disclosures, one should expect no extra boost on diffusion from VC financing, unless citations are used strategically by potential targets. In unreported results, I find that the coefficient estimate of VC_{pt} is close to one and is not statistically significant, which minimizes this concern.

A final alternative story is that the increase in citations is due to "litigation fear." Inventors may decide to cite the patents of a company only after the company is VC financed, because the threat of litigation before VC financing is weak. This concern is however minimized to the extent that citations represent no protection against patent infringement law suits. (For more on this topic see the Supreme Court Ruling of Microsoft Corp. versus I4I Limited Partnership, 2010). In other words, inventors may choose not to infringe patents once the inventing companies are VC financed, but the VC financing event should have little effect on their citation behavior.

5 Conclusion

This paper investigates how the diffusion of an idea is affected by VC financing of the company that patented the idea. I find a strong and causal effect on diffusion as measured by patent citations. The empirical evidence points to two mechanisms. First, VCs increase awareness of innovations, possibly by certifying their value to the general public, and influence the direction of aggregate innovative activity. Second, VCs provide a platform for interaction that facilitates communication across portfolio companies.

The main identification challenge in estimating the effect of VC financing on the diffusion of existing knowledge is the endogeneity of VC investments. I address this challenge using time series variation in the size of public pension assets as an instrumental variable. The validity of this approach relies on the home-bias of state pension funds in their VC investments (Hochberg and

Rauh, 2012), and on the exclusion restriction that changes in pension assets are independent of companies' innovation opportunities. To address concerns that unobserved economic activity affects both the size of state pension assets and companies' innovation opportunities, I compare citations to patents filed by VC-backed companies to those of comparable patents. The exclusion restriction is satisfied, thus, as long as the effect of unobserved economic activity on innovation opportunities within a state is uniform across local patents in the same technology-class and vintage-year.

This paper contributes to our understanding of how financial intermediaries affect innovation. I find evidence that VCs have a multiplier effect on innovation that goes above and beyond financing the innovation of their targets. This result is informative for policy makers that seek to spur innovation by stimulating VC activity. My findings suggest that VC financing increases diffusion of ideas both inside and outside the VC industry, which implies that venture capital not only rewards VC-backed companies, but also creates benefits that are shared by society at large and can have important distributional consequences. However, this feedback from finance to the creation of scientific knowledge does not necessarily imply that all innovation should be financed through VCs. By focusing exclusively on research with high short-term rewards more basic research may be sacrificed, which can be costly for innovation in the long-run. Assessing the general equilibrium effects of the role of VCs on innovation is a fruitful avenue for future research.

Finally, my finding that VC portfolios are conduits for information flows also deserves more attention. I show that the stronger increase in citations inside VC portfolios cannot be explained by inventor turnover among companies that share a common VC. However, it is possible that the mobility of other personnel can explain this concentration of knowledge flows inside VC portfolios. There is plenty of informal evidence that VCs recycle executives across portfolio companies. Exploring whether this evidence is systematic, and whether it is associated with knowledge spillovers, are other avenues for future research.

6 References

Aghion,P., Dewatripont, M. and J. Stein (2008) Academic Freedom, Private-Sector Focus and the Process of Innovation. In *RAND Journal of Economics* 39(3), 617-635.

Agrawal, A. and J. Singh (2011) Recruiting for Ideas: How Firms Exploit the Prior Inventions of New Hires. In *Management Science* 57(1), 129-150.

Alcacer, J. and M. Gittelman (2006) Patent Citations as a Measure of Knowledge Flows: The Influence of Examiner Citations In *Review of Economics and Statistics* 88(4), 774-779.

Almeida, P. and B. Kogut (1999) Localization of Knowledge and the Mobility of Engineers in Regional Networks. *Management Science* 45(7), 905-917.

Angrist, J.D. and J. Pischke (2008) *Mostly Harmless Econometrics*, Princeton: Princeton University Press.

Arrow, K. J. (1962) The Economic Implications of Learning by Doing. *Review of Economic Studies* 29(3),155-73

Azoulay P., Graff Zivin, J. S. and B. Sampat (2012) The Diffusion of Scientific Knowledge Across Time and Space: Evidence from Professional Transitions for the Superstars of Medicine. Chapter 2 in *The Rate & Direction of Inventive Activity Revisited* (edited by Joshua Lerner and Scott Stern),107-155, 2012. Chicago, IL: University of Chicago Press.

Audretsch D. and M. Feldman (1996) R&D Spillovers and the Geography of Innovation and Production. In *American Economic Review* 86(4), 253-273.

Bernstein, S. (2012) Does Going Public Affect Innovation? Unpublished Working Paper. Available at SSRN: http://ssrn.com/abstract=2061441

Bernstein, S., Lerner, J., Sørensen, M. and P. Stromberg (2008) Private Equity and Industry Performance. Unpublished Harvard Business School Entrepreneurial Management Working Paper No. 10-045. Available at SSRN: http://ssrn.com/abstract=1524829.

Bloom N. and J. Van Reenen (2002) Patents, Real Options and Firm Performance. In *Economic Journal* 112 (March), C97-C116.

Bresnahan, T. (2012) Generality, Recombination and Reuse. Chapter 13 in *The Rate & Direction of Inventive Activity Revisited* (edited by Joshua Lerner and Scott Stern),611-667, 2012. Chicago, IL: University of Chicago Press.

Card, D. (2001) Estimating the Return to Schooling: Progress on Some Persistent Econometric Problems. In *Econometrica* 69(5), 1127-1160.

Dahl, M. and Ø. R. Pederson (2005) Social Networks in the R&D Process: The Case of the Wireless Communication Industry Around Aalborg, Denmark. In *Journal of Engineering and Technology Management* 22(1-2), 75-92.

Duguet, E. and M. MacGarvie (2005) How Well Do Patent Citations Measure Knowledge Spillovers? Evidence from French innovation surveys. In *Economics of Innovation and New Technology* 14(5), 375-393

Florida R.L. and M. Kenney (1988) Venture Capital, High Technology and Regional Development. In *Regional Studies* 22 (1), 33-48.

Gans, J. and F. Murray (2012). Funding Scientific Knowledge. Selection, Disclosure and the Public-Private Portfolio. Chapter 1 in *The Rate & Direction of Inventive Activity Revisited* (edited by Joshua Lerner and Scott Stern),51-105, 2012. Chicago, IL: University of Chicago Press.

Gans, J., Murray, F. and S. Stern (2010) Contracting Over the Disclosure of Scientific Knowledge: Intellectual Property and Academic Publication. MIT Working Paper. Available at: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1559871

GAO, State and Local Government Pension Plans. Report to the Ranking Member, Committee on Finance, U.S. State.

Gomes-Casseres, B., Hagedoom, J. and A. B. Jaffe (2006) Do Alliances Promote Knowledge Flows? In *Journal of Financial Economics* 80 (1), 5-33.

Gourieroux, C., Monfort, A. and A. Trognon (1984) Pseudo Maximum Likelihood Methods: Application to Poisson Models. In *Econometrica* 52(3), 701-720.

Hall, B., Jaffe, A. B. and M. Trajtenberg (2001) The NBER Patent Citations Data File: Lessons, Insights and Methodological Tools. Working Paper no. 8498, National Bureau of Economic Research

Hall, B., Jaffe, A. B. and M. Trajtenberg (2005) Market Value and Patent Citations. In *RAND Journal of Economics* 36(1), 16-38

Hall, B. Mairesse J. and L. Turner (2007) Identifying Age, Cohort and Period Effects in Scientific Research Productivity - Discussion and Illustration Using Simulated and Actual Data on French Physicists. In *Economics of Innovation and New Technology* 16(2), 159-177.

Hausman, J., Hall, B. and Z. Griliches (1984) Econometric Models for Count Data with an Application to the Patents - R&D relationship. In *Econometrica* 52, 909-938.

Hellman, T. and M. Puri (2000) The Interaction between Product Market and Financing Strategy: The Role of Venture Capital. In *Review of Financial Studies* 13(4), 959-984.

Hellman, T. and M. Puri (2002) Venture Capital and the Professionalization of Start-Up Firms: Empirical Evidence. In *Journal of Finance* 57(1), 69-197.

Hirukawa, M. and M. Ueda (2008) Venture Capital and Industrial Innovation. Available at: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1242693

Hochberg, Y. (2011) Venture Capital and Corporate Governance in the Newly Public Firm. In *Review of Finance* 0, 1-52.

Hochberg, Y. Ljungqvist, A. and Y. Lu (2007) Whom you Know Matters: Venture Capital Networks and Investment Performance. In *Journal of Finance* 62(1), 251-301.

Hochberg, Y. and J. Rauh (2011) Local Overweighting and Underperformance: Evidence from Limited Partner Private Equity Investments. Unpublished Working Paper. Available at SSRN: http://ssrn.com/abstract=1798747

Imbens, G. and J. Angrist (1994) Identification and Estimation of Local Average Treatment Effects. In *Econometrica* (62)2, 467-475.

Imbens, G. (2009) Better LATE Than Nothing: Some Comments on Deaton (2009) and Heckman and Urzua (2009) In *Journal of Economic Literature* 48(2), 399-423.

Jaffe, A. B. (1986) Technological Opportunity and Spillovers of R&D: Evidence from Firms' Patents, Profits and Market Value. In *American Economic Review* 76 (5), 984-1001.

Jaffe, A. B., Henderson, R. and M. Trajtenberg (1993) Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations. In *Quarterly Journal of Economics* 108, 577-598.

Jaffe, A. B. and M. Trajtenberg (2002) *Patents, Citations and Innovations: A Window on the Knowledge Economy.* Cambridge, Mass.: MIT Press

Jones, B. (2009) The Burden of Knowledge and the "Death of the Renaissance Man": Is Innovation Getting Harder?. In *Review of Economic Studies* 76, 283-317.

Kim, J. and G. Marschke (2005) Labor Mobility of Scientists, Technological Diffusion, and the Firm's Patenting Decision. In *RAND Journal of Economics* 36(2), 298-317.

Kogut, B. and U. Zander (1993). Knowledge of the Firm and the Evolutionary Theory of the Multinational Corporation. In *Journal of International Business Studies* 24(4), 625-645.

Kogut, B. and U. Zander (1992) Knowledge of the Firm, Combinative Capabilities, and the Replication of Technology. In *Organization Science* 3(3), 383-397.

Kortum, S. and J. Lerner (1998) Assessing the Contribution of Venture Capital to Innovation. In *RAND Journal of Economics* 31(4), 674-692.

Lai, R., D'Amour A., and L. Fleming (2009) The Careers and Coauthorship Networks of U.S. Patent-Holders, since 1975, hdl:1902.1/12367 UNF:5:daJuoNgCZlcYY8RqU+/j2Q== Harvard Business School;Harvard Institute for Quantitative Social Science [Distributor]

Lemley M. and B. Sampat (2010) Examiner Experience and Patent Office Outcomes. In *Review of Economics and Statistics* 94(3), 817-827.

Lerner, J. (1995) Venture Capital and the Oversight of Private Firms. In *Journal of Finance* 50, 301-318.

Lerner, J. (2009) *Boulevard of Broken Dreams.* Princeton: Princeton University Press.

Lerner, J., Sørensen, M. and P. Stromberg (2011) Private Equity and Long-Run Investment: The Case of Innovation. In *Journal of Finance 66* (2), 445-477.

Lindsey, L. (2008) Blurring Firm Boundaries: The Role of Venture Capital in Strategic Alliances. In *Journal of Finance* 63 (3), 1137-1168.

Megginson, W.L. and K. A. Weiss (1991) Venture Capitalist Certification in Initial Public Offerings. In *Journal of Finance* 46(3), 879-903

Mokyr, J. (2003) Thinking about Technology and Institutions. In Macalester International: Vol. 13, Article 8.

Mollica, M. and L. Zingales (2007) The Impact of Venture Capital on Innovation and the Creation of New Business. Unpublished Working Paper, University of Chicago.

Nanda, R. and M. Rhodes-Kropf (2011) Investment Cycles and Startup Innovation. Unpublished HBS Working Paper No. 12-032. Available at SSRN: http://ssrn.com/abstract=1950581.

Popov, A. and P. Rosenboom (2009) Venture Capital and Patented Innovation: Evidence from Europe. In *Economic Policy* 27(71), 447-482.

Roach and Cohen (2010) Lens or Prism? Patent Citations as a Measure of Knowledge Flows from Public Research. *Forthcoming Management Science.*

Romer, P. (1986) Increasing returns and long-run growth. In *Journal of Political Economy* 94(5), 1002-37.

Romer, P. (1990) Endogenous Technological Change. In *Journal of Political Economy*, 98(5) Part 2: The Problem of Development: A Conference of the Institute for the Study of Free Enterprise Systems. S71-S102.

Sahlman, W. (1990) The Structure and Governance of Venture Capital Organizations. In *Journal of Financial Economics* 27(1990), 473-521.

Samila, S. and O. Sorenson (2011). Noncompete Covenants: Incentives to Innovate or Impediments to Growth. In *Management Science* 57(93), 425-438.

Sampat, B. (2010) When Do Applicants Search for Prior Art? In *Journal of Law and Economics* 53(2), 399-416

Sampat, B. and A. Ziedonis (2004) Patent Citations and The Economic Value of Patents. In *Handbook of Quantitative Science and Technology Research* 277-297.

Saxenian, A. L. (1990) Regional Networks and the Resurgence of Silicon Valley. In *California Management Review*. 33(1), 89-103.

Saxenian, A. L. (1994) *Regional advantage*. Cambridge: Harvard University Press.

Seru, A. (2012) Firm Boundaries Matter:Evidence from Conglomerates and R&D Activity. In *Journal of Financial Economics (forthcoming).*

Serrano, C. (2010) The Dynamics of the Transfer and Renewal of Patents. In *RAND Journal of Economics.* 41(4):686-708.

Sørensen, M. (2007) How Smart is Smart Money? A Two Sided Matching Model of Venture Capital. In *Journal of Finance* 62(6) 2725-2762.

Sorenson, O. and T. Stuart (2001) Syndication Networks and the Spatial Distribution of Venture Capital Investments. In *American Journal of Sociology.* 106(6), 1546-1588.

Sorenson, O. and A. Fleming (2004). Science and the Diffusion of Knowledge. In *Research Policy* 33, 1615-1634.

Stock, J and M. Yogo (2005) Testing for Weak Instruments in Linear IV Regression. Chapter 5 in *Identification and Inference for Econometric Models: Essays in Honor of Thomas J. Rothenberg* (edited by James Stock and Donald Andrews) 80-108, Cambridge University Press.

Trajtenberg, M. (1990). A Penny for your Quotes: Patent Citations and the Value of Innovations. In *Rand Journal of Economics* 21, 172-187.

Williams, H. (2011) Intellectual Property Rights and Innovation: Evidence from the Human Genome, 2011. NBER working paper 16213

Windmeijer, F. (2000) Moment Conditions for Fixed Effects Count Data Models with Endogenous Regressors. In *Economic Letters* 68, 21-24.

Wooldridge, J.M. (1997) Multiplicative Panel Data Models Without the Strict Exogeneity Assumption. In *Economic Theory* 13, 667-678.

Wooldridge, J.M. (1999) Distribution-Free Estimation of Some Nonlinear Panel Data Models. In *Journal of Econometrics* 90, 77-97

Wooldridge, J.M. (2010) *Econometric Analysis of Cross Section and Panel Data.* Second Edition. MIT Press.

Zander, U and B. Kogut (1995) Knowledge and the speed of the transfer and imitation of organizational capabilities: An empirical test. In *Organization Science*, 6(1), 76-92.

Figure 1- Estimated temporal trends in citations to patents

The solid lines in the plot correspond to the coefficient estimates of a fixed-effects Poisson model in which the dependent variable corresponds to annual citations to patents, and the explanatory variables are Event Year dummies. I restrict the sample to a [-2,5] year window around the financing event of the issuing company. The 95% confidence interval (corresponding to robust standard errors, clustered at the issuing company level) around these estimates is plotted with dashed lines. The reference period for interpreting the plot is the year of the VC investment (Event Year 0).

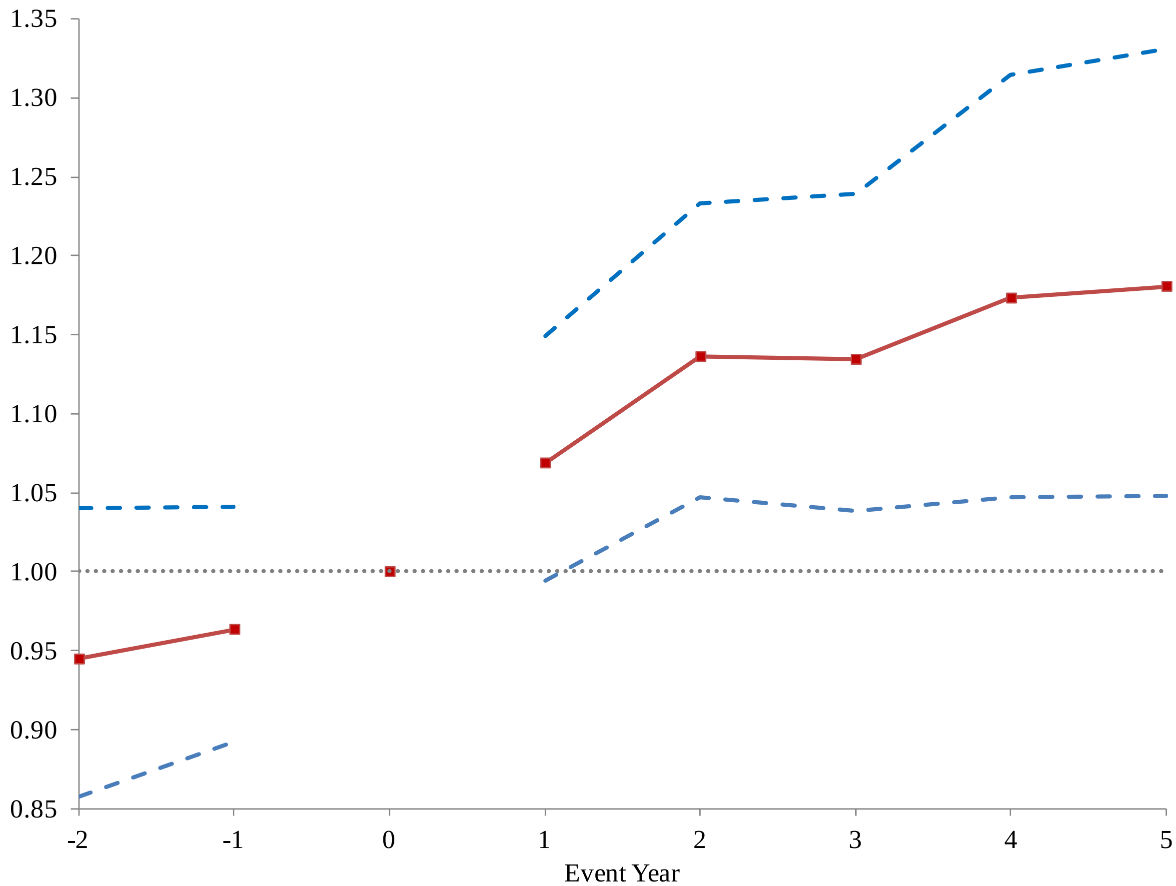


Figure 2- Patent sale likelihood

The figure presents the annual probability that a patent is sold in the two years before, and nine years after a VC invests in the issuing company. The solid line describes patents in my sample, and the dashed line corresponds to matching patents at the technology-class and application- year, and that were not financed by a VC.

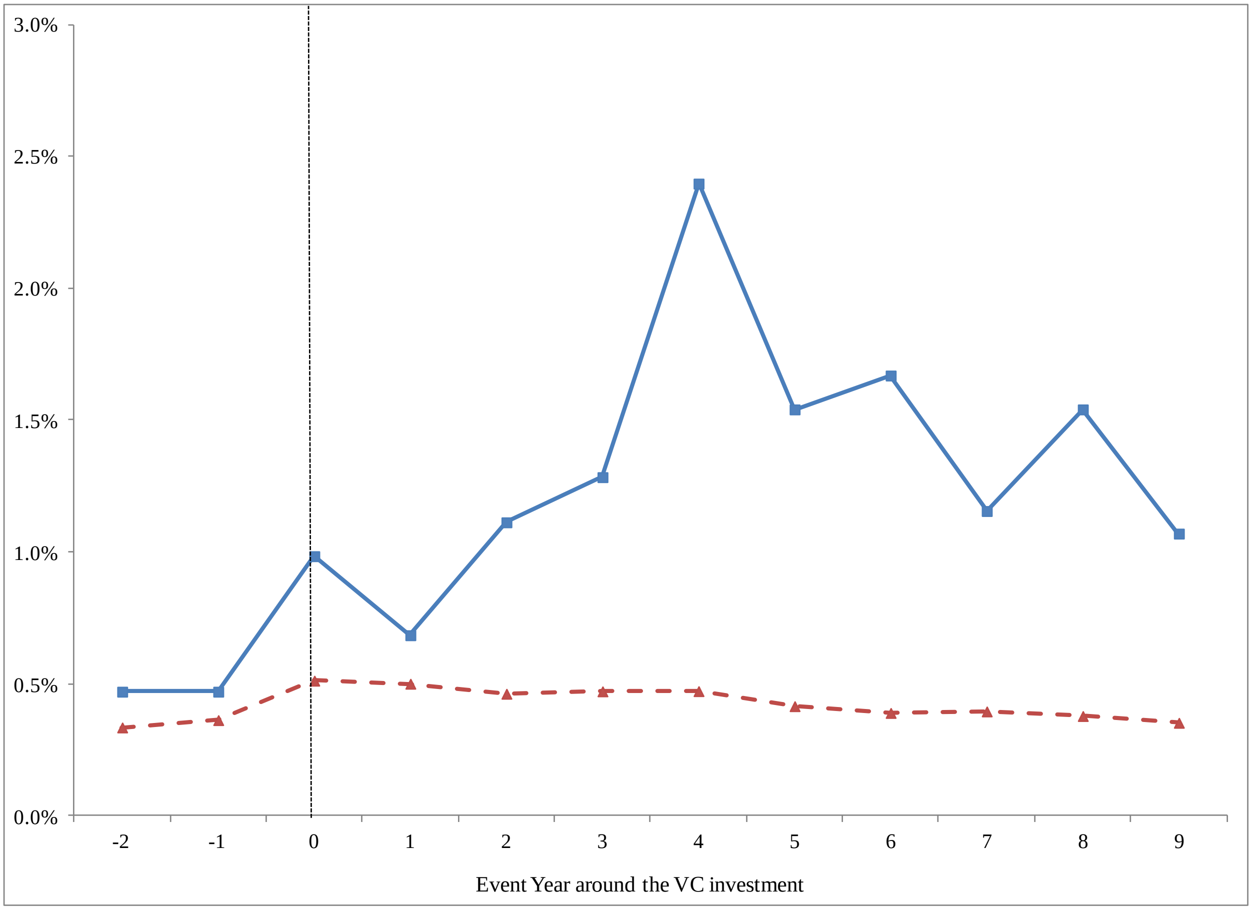


Figure 3- Exposure Effect of VC investments

The figure presents the normalized annual searches made in Google to companies that were first financed by a VC in 2006. To construct the graph, I strip company names of punctuation, capitalization and common acronyms and search for weekly hits in Google Insights since January 2004 until the end of 2011. The solid line corresponds to average annual searches to the normalized names, relative to the total number of searches done on Google over time. The dashed lines correspond to average annual searched to the word "Gold". Google Insights analyzes only a portion of Google web searches to compute how many searches have been done for the entered terms, relative to the total number of searches done on Google over time. This analysis indicates the likelihood of a random user to search for a particular search term at a certain time. Google Insights designates a certain threshold of traffic for search terms, so that those with low volume won't appear. It also eliminates repeated queries from a single user over a short period of time, so that the level of interest isn't artificially impacted by this type of queries. The information on companies that were first financed by a VC in 2006 is from SDC Thompson.

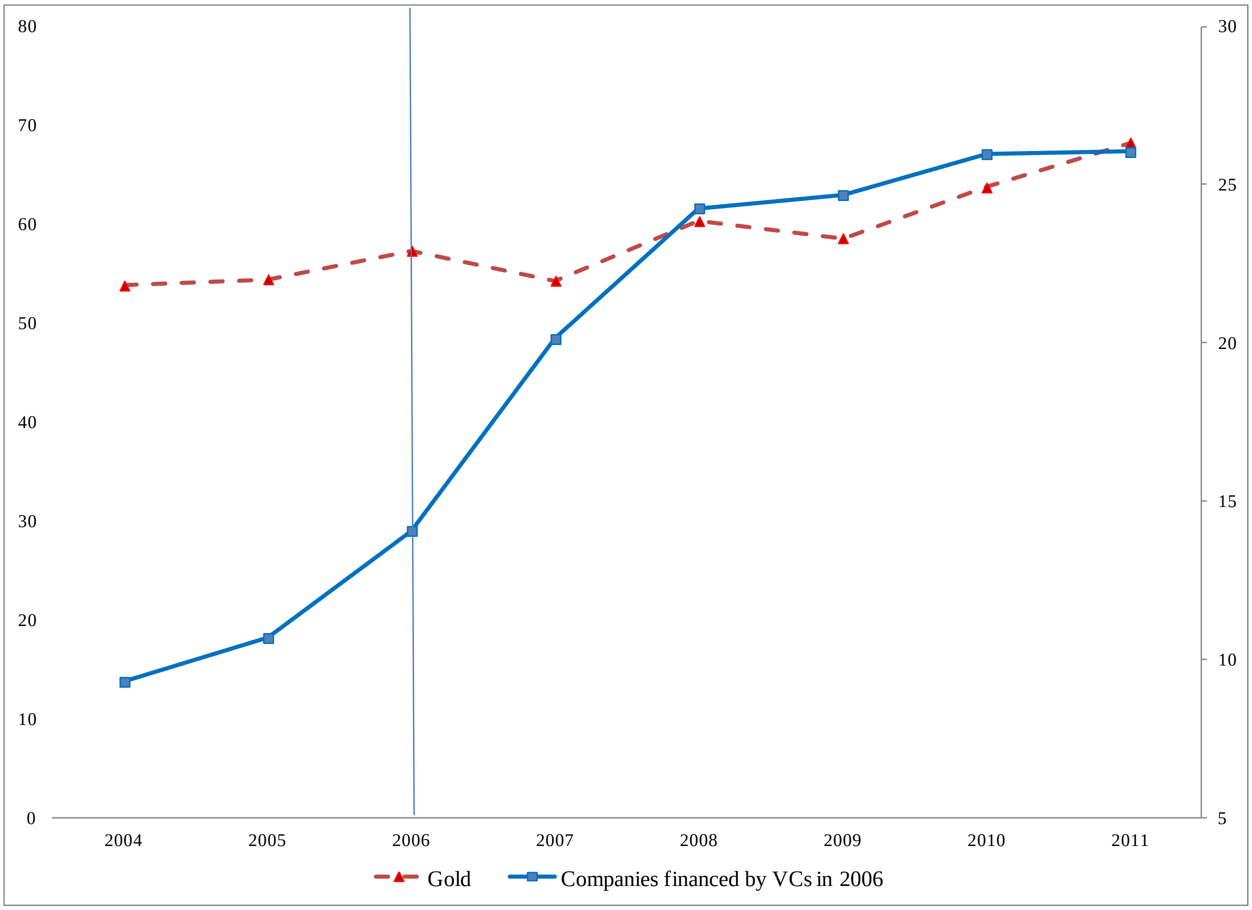


Table 1 - Summary statistics analysis sample

The sample consists of 2,336 patents filed by 752 VC-backed companies at least two years before they were first financed by a VC (347 VC firms). For Panel B I use the state of the company as reported in the VentureXpert database. For Panels B, C, D and E the percentage of companies used for comparisons consists of 5,108 VC-backed companies that patent from the full matched sample, and of 20,058 companies included in the VentureXpert database between 1976 and 2009 (See Appendix 1 for details). The industry classification used in Panel D is based on the VentureXpert files.

Panel A. Application and grant years of patents, and transaction years for the VC investments

Year	Patent Applications	Patent Grants	VC investments
1976	144	3	
1977	78	73	
1978	84	85	3
1979	69	66	6
1980	45	67	10
1981	48	73	28
1982	47	37	15
1983	46	37	14
1984	62	52	15
1985	71	52	8
1986	44	70	17
1987	56	64	20
1988	70	53	12
1989	70	77	15
1990	66	62	16
1991	74	59	16
1992	92	61	13
1993	95	71	9
1994	99	80	18
1995	139	93	24
1996	117	78	38
1997	188	85	36
1998	207	132	67
1999	117	152	56
2000	126	160	86
2001	82	148	55
2002		107	77
2003		96	78
2004		51	
2005		45	
2006		30	
2007		9	
2008		8	
Total	2,336	2,336	752

Panel B. Distribution by state of companies and associated patents: Top States in Analysis Sample

State	% of Companies: Analysis Sample	% of Companies: Full Matched Sample	% of Companies: Overall VC Population	% of Patents: Analysis Sample	% of Patents: Full Matched Sample
CA	34.97%	44.4%	38.8%	32.6%	56.5%
CO	2.39%	2.7%	2.9%	3.6%	1.2%
CT	2.66%	1.7%	1.6%	3.3%	0.8%
IL	2.53%	1.9%	2.2%	2.1%	0.6%
MA	14.10%	12.8%	10.8%	10.5%	9.0%
NJ	2.66%	2.6%	2.5%	2.3%	1.1%
NY	3.99%	2.9%	5.3%	3.4%	1.8%
PA	3.46%	3.1%	3.4%	5.0%	2.2%
TX	5.32%	4.8%	5.7%	9.7%	5.9%
WA	2.66%	2.9%	3.2%	1.9%	10.7%

Panel C. Distribution of companies by type of VC investment

	# of companies	% of Companies: Analysis Sample	% of Companies: Full matched Sample	% of Companies: Overall VC Population
Bridge Loan	21	2.8%	1.7%	2.4%
Early Stage	257	34.2%	38.2%	39.8%
Expansion	299	39.8%	25.3%	25.7%
Later Stage	91	12.1%	7.0%	5.9%
Seed	84	11.2%	27.8%	26.1%
Total	752			

Panel D. Distribution of companies by type of VC exit

	# of companies	% of Companies: Analysis Sample	% of Companies: Full matched Sample	% of Companies: Overall VC Population
Acquisition	282	37.5%	34.9%	30.8%
Active	209	27.8%	29.9%	35.8%
Bankruptcy	9	1.2%	1.3%	1.3%
Defunct	140	18.6%	14.4%	19.9%
Merger	10	1.3%	1.6%	1.6%
Other	11	1.4%	1.7%	1.6%
Went Public	91	12.1%	16.3%	9.1%
Total	752			

Panel E. Industry distribution of companies and patents

	# of companies	% of Companies: Analysis Sample	% of Companies: Full matched Sample	% of Companies: Overall VC Population	# of patents	% of Patents: Analysis Sample	% of Patents: Full matched Sample
Biotechnology	63	8.4%	9.9%	6.1%	199	8.5%	11.6%
Comm. and Media	75	10.0%	11.0%	10.3%	215	9.2%	8.9%
Computer Hardware	51	6.8%	8.9%	6.3%	104	4.5%	17.2%
Computer Software	94	12.5%	16.3%	21.3%	180	7.7%	13.6%
Consumer Related	33	4.4%	2.0%	4.8%	125	5.4%	1.1%
Industrial Energy	97	12.9%	8.0%	5.1%	377	16.1%	4.4%
Internet Specific	37	4.9%	8.5%	20.7%	50	2.1%	2.0%
Medical Health	145	19.3%	16.8%	11.6%	505	21.6%	15.1%
Other Products	30	4.0%	2.6%	6.6%	101	4.3%	0.8%
Semiconductors	127	16.9%	16.0%	7.2%	480	20.5%	25.3%
Total	752				2,336		

Panel F. Distribution of patent age the year of the VC investment

	Number of patents	Percentage of sample
2 Years	462	19.78
3 Years	643	27.53
4 Years	325	13.91
5 Years	210	8.99
Between 6 years and 10 years	411	17.59
More than 10 years	285	12.19
Total	2,336	

Panel G. Annual Citations (excluding self-citations)

	Citation type	Baseline state-level	Mean	S. D.	Med.	Min	Max	Obs.
Patents	All		0.92	2.45	0.00	0.00	60.00	43,519
Citation Baseline	All	No	0.63	0.77	0.39	0.00	13.50	43,519
Citation Baseline	All	Yes	0.61	1.14	0.25	0.00	32.00	43,519
Patents	Out-state		0.75	2.01	0.00	0.00	43.00	46,519
Citation Baseline	Out-state	Yes	0.49	0.96	0.20	0.00	29.00	46,519

Table 2 - Summary statistics restricted sample 1993-2008

Information on public state pension funds' assets from the Census Bureau is only available from 1993 to 2008. The sample restricted to VC investments during this period consists of 1,657 patents filed by 517 VC-backed companies. For Panel B, I use the state of the company as reported in the VentureXpert database. Pension Funds' Assets is the value of the assets held by local and state pension funds deflated and expressed in billions of 1982 U.S. dollars.

Panel A. Application and grant years of patents and transaction years for the VC

Year	Patent Applications	Patent Grants	VC investments
1976	29		
1977	16	14	
1978	19	20	
1979	16	13	
1980	10	18	
1981	12	14	
1982	18	8	
1983	15	8	
1984	25	21	
1985	25	17	
1986	23	30	
1987	27	26	
1988	39	21	
1989	48	38	
1990	29	42	
1991	50	29	
1992	86	36	
1993	95	45	
1994	99	70	
1995	139	88	13
1996	117	77	18
1997	188	85	25
1998	207	132	44
1999	117	151	49
2000	126	160	81
2001	82	148	53
2002		107	74
2003		96	77
2004		51	
2005		45	
2006		30	
2007		9	
2008		8	
Total	1,657	1,657	517

Panel B. Distribution of public pension funds' assets, companies and patents by state

	Pension Assets		% of Patents		% of Companies	
	Mean	Std. dev.	Restricted sample	Analysis sample	Restricted sample	Analysis sample
AL	17.17	1.45	0.60%	0.40%	0.40%	0.30%
AZ	19.87	3.48	1.80%	1.80%	1.00%	1.60%
CA	284.37	64.69	37.00%	32.60%	37.30%	35.00%
CO	21.67	5.06	3.10%	3.60%	2.30%	2.40%
CT	15.17	3.06	2.70%	3.30%	2.30%	2.70%
DC	3.26	0.81	0.40%	0.30%	0.20%	0.10%
FL	71.01	16.19	1.90%	2.00%	1.90%	2.00%
GA	35.36	8.79	1.10%	1.30%	1.90%	2.00%
ID	5.06	0.95	0.90%	0.70%	0.40%	0.40%
IL	63.77	14.25	2.10%	2.10%	2.10%	2.50%
IN	14.25	2.09	0.10%	0.20%	0.20%	0.40%
KS	6.79	1.91	0.10%	0.20%	0.20%	0.10%
LA	17.67	3.99	0.80%	0.60%	0.60%	0.40%
MA	29.68	7.12	7.80%	10.50%	12.00%	14.10%
MD	26.89	4.89	3.30%	3.40%	2.30%	2.40%
ME	4.98	1.59	0.30%	0.20%	0.20%	0.10%
MI	48.45	7.78	0.80%	1.20%	1.40%	1.60%
MN	28.79	5.21	1.20%	1.30%	2.10%	2.30%
MO	28.27	5.12	0.40%	0.70%	0.60%	0.90%
NC	39.50	6.05	1.20%	0.80%	2.10%	1.50%
NE	4.82	1.11	0.20%	0.20%	0.20%	0.10%
NH	2.81	0.84	0.80%	1.20%	1.20%	1.50%
NJ	36.51	6.14	2.70%	2.30%	2.70%	2.70%
NM	10.42	2.48	0.50%	0.30%	0.60%	0.40%
NV	9.47	3.18	0.10%	0.00%	0.20%	0.10%
NY	177.14	35.43	3.80%	3.40%	4.60%	4.00%
OH	80.36	12.57	2.30%	2.20%	1.70%	1.90%
OR	27.04	7.99	0.40%	1.90%	0.60%	0.70%
PA	58.90	10.68	2.90%	5.00%	3.10%	3.50%
RI	4.49	1.51	0.10%	0.10%	0.40%	0.30%
SC	14.78	2.91	0.10%	0.20%	0.20%	0.30%
TN	20.31	3.82	2.20%	2.00%	0.60%	0.80%
TX	90.55	21.93	12.20%	9.70%	6.40%	5.30%
UT	9.51	2.35	0.30%	0.70%	0.80%	0.80%
VA	30.84	6.57	0.80%	0.90%	1.40%	1.20%
VT	1.53	0.40	0.30%	0.30%	0.20%	0.30%
WA	32.06	5.81	2.20%	1.90%	3.10%	2.70%
WI	44.96	7.34	0.20%	0.30%	0.40%	0.40%
WY	2.85	0.81	0.30%	0.20%	0.20%	0.10%

Panel C. Distribution of companies by type of VC investment

	Number of Companies	Percentage of sample	
		Restricted Sample	Analysis Sample
Bridge Loan	16	3.1%	2.8%
Early Stage	209	40.4%	34.2%
Expansion	192	37.1%	39.8%
Later Stage	57	11.0%	12.1%
Seed	43	8.3%	11.2%
Total	517		

Panel D. Industry distribution of companies and patents

	# of companies	% of Companies		# of patents	% of Patents	
		Restricted Sample	Analysis Sample		Restricted Sample	Analysis Sample
Biotechnology	55	10.6%	8.4%	179	10.8%	8.5%
Comm. and Media	52	10.1%	10.0%	170	10.3%	9.2%
Computer Hardware	27	5.2%	6.8%	58	3.5%	4.5%
Computer Software	72	13.9%	12.5%	141	8.5%	7.7%
Consumer Related	17	3.3%	4.4%	79	4.8%	5.4%
Industrial Energy	42	8.1%	12.9%	218	13.2%	16.1%
Internet Specific	36	7.0%	4.9%	46	2.8%	2.1%
Medical Health	109	21.1%	19.3%	362	21.9%	21.6%
Other Products	21	4.1%	4.0%	67	4.0%	4.3%
Semiconductors	86	16.6%	16.9%	337	20.3%	20.5%
Total	517			1,657		

Panel E. Distribution of companies by type of VC exit

	Number of companies	% of Companies	
		Restricted Sample	Analysis Sample
Acquisition	185	35.8%	37.5%
Active	184	35.6%	27.8%
Bankruptcy	8	1.6%	1.2%
Defunct	77	14.9%	18.6%
Merger	4	0.8%	1.3%
Other	7	1.4%	1.4%
Went Public	52	10.1%	12.1%
Total	517		

Panel F. Distribution of patent age the year of the VC investment

	Number of patents	Percentage of sample
2 Years	329	28.12
3 Years	425	34.79
4 Years	220	15.9
5 Years	132	9.06
Between 6 years and 10 years	298	12.14
More than 10 years	253	
Total	1,657	

Panel G. Annual Citations (excluding self-citations)

	Citation type	Baseline state-level	Mean	S. D.	Med.	Min	Max	Obs.
Patents	All		1.21	2.97	0.00	0.00	60.00	21,757
Citation Baseline	All	No	0.82	0.92	0.56	0.00	11.47	21,757
Citation Baseline	All	Yes	0.84	1.39	0.41	0.00	28.69	21,757
Patents	Out-state		0.97	2.40	0.00	0.00	43.00	21,757
Citation Baseline	Out-state	Yes	0.66	1.15	0.31	0.00	25.67	21,757

Table 3– Univariate tests VC investments and patent citations

This table compares average annual citations to patents to the average annual citation baseline before and after the VC investment. *, **, and *** indicate statistical significance at the 10%, 5% and 1% level.

	Annual Average			
	Pre-VC	Post-VC	Difference	Ratio
Citations	0.64	1.04	0.40***	1.63
	(1.69)	(2.69)		
Citation Baseline	0.54	0.66		
	(0.61)	(0.83)		
Difference	0.10***	0.37***		
Ratio	1.19	1.60		
			Difference in Difference	Ratio of Ratios
			0.28***	1.33

Table 4 –Poisson Regressions VC investments and patent citations

The table contains Poisson regression estimates. An observation is a patent-year. The dependent variable is annual citations. VC_{pt} is an indicator variable that equals 1 after VC investment. b_t corresponds to average citations received at year *t* by matching patents in the same technology-class and with the same application-year. The Poisson model requires that annual average citations to matching patents be different from zero, which explains the difference in observations across columns (1)-(2). The fixed-effects Poisson model requires variation in the dependent variable for each patent, which explains the difference in observations across columns (2), and, (3). The reported coefficients are incidence rates. A coefficient greater than one corresponds to a positive relationship between the explanatory variable and the citation intensity. Standard errors are clustered at the issuing company level. *, **, and *** indicate statistical significance at the 10%, 5% and 1% levels, respectively.

	(1)	(2)	(3)
Model	Poisson	Poisson	Poisson
VC_{pt}	1.627***	1.328***	1.189***
	(0.106)	(0.063)	(0.045)
Constant	0.636***	1.177***	
	(0.038)	(0.050)	
Observations	43,519	41,172	38,981
Number of patents	2,336	2,336	2,183
Number of companies	752	752	723
Offset b_t	No	Yes	Yes
Patent FE	No	No	Yes

Table 5 – Investments by VCs and state pension funds' assets

The table reports the relation between investments by VC firms and pension funds' assets in their home-state. Observations are at the state-year level. The dependent variable is stated at the beginning of each column. Investments correspond to the value of investments made by VC firms (in billions 1982 U.S dollars). Local Investments correspond to the value of investments made by VC firms in local companies (in billions 1982 U.S dollars). New Investments correspond to the value of investments made by VC firms in new companies (in billions 1982 U.S dollars). Standard errors are clustered at the state level. $Pension$ corresponds to the value of assets held by local and state pension funds in 1982 billion U.S. dollars and lagged by 1 year. *, **, and *** indicate statistical significance at the 10%, 5% and 1% level.

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Variable	Investments	Local Investments	Non-local Investments	New Investments	New Local Investments	Non-local New Investments
$Pension$	0.052***	0.036**	0.016***	0.013***	0.009**	0.004***
	(0.016)	(0.014)	(0.003)	(0.004)	(0.003)	(0.001)
Constant	-1.502**	-1.126**	-0.376***	-0.401**	-0.291**	-0.109***
	(0.591)	(0.486)	(0.114)	(0.157)	(0.131)	(0.028)
Obs.	765	765	765	765	765	765
$State$ F.E.	Yes	Yes	Yes	Yes	Yes	Yes
$Year$ F.E.	Yes	Yes	Yes	Yes	Yes	Yes

Table 6– IV analysis VC investments and patent citations

This table contains Poisson and Linear regression coefficients of the IV analysis. An observation is a patent-year. The dependent variable is indicated on top of each column. The instrument used in each model is indicated on top of each column. Scaled Cites corresponds to the number of citations a patent receives divided by the average number of citations received by matching patents in the same technology-class and filed the same year as the patent. VC_{pt} is a dummy variable that equals one after the VC investment. $Pension$ corresponds to the value of assets held by local and state pension funds in the home-state of the company expressed in 1982 billion U.S. dollars, lagged by 1 year. $Pension\_Norm$ corresponds to the value of assets held by local and state pension funds in the home-state of the company expressed 1982 billion U.S. dollars, normalized by average state GDP, and lagged by 1 year. $\Delta Pension$ corresponds to the value of assets held by local and state pension funds in the home-state of the company expressed in 1982 billion U.S. dollars, lagged by 1 year, and demeaned by time. b_t corresponds to average citations received at year t by matching patents in the same technology-class and with the same application-year. The estimated coefficients for columns (1)-(4) are incidence rates, except for the estimated coefficient of the first-stage. For the incidence rates a coefficient greater than one corresponds to a positive relationship between the explanatory variable and the citation intensity. Standard errors are clustered at the state level. *, **, and *** indicate statistical significance at the 10%, 5% and 1% level.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dependent Variable	Cites	Cites	Cites	Cites	Scaled Cites	Scaled Cites	Scaled Cites	Scaled Cites
Instrument		$Pension$	$Pension\_Norm$	$\Delta Pension$		$Pension$	$Pension\_Norm$	$\Delta Pension$
Model	Poisson	GMM-IVs	GMM-IVs	GMM-IVs	OLS	2SLS	2SLS	2SLS
VC_{pt}	1.171***	1.495***	1.318	1.564***	0.225*	0.607***	0.524*	0.683***
	(0.069)	(0.199)	(0.239)	(0.179)	(0.132)	(0.218)	(0.272)	(0.187)
First Stage						0.006***	2.409**	0.006***
						(0.001)	(0.969)	(0.001)
F-test						31.13	6.187	84.46
Obs.	15,622	15,622	15,622	15,622	18,928	18,928	18,928	18,928
# patents	1,487	1,487	1,487	1,487	1,657	1,657	1,657	1,657
# companies	491	491	491	491	517	517	517	517
# states	38	38	38	38	39	39	39	39
Offset b_t	Yes	Yes	Yes	Yes				
Patent FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Table 7– First robustness check IV analysis VC investments and patent citations

This table contains Poisson and Linear regression coefficients of the IV analysis. An observation is a patent-year. The dependent variable is indicated on top of each column. The instrument used in each model is indicated on top of each column. Scaled Cites corresponds to the number of citations a patent receives divided by the average number of citations received by matching patents in the same technology-class, with the same application-year and issued in the same state as the patent. VC_{pt} is a dummy variable that equals one after the VC investment. $Pension$ corresponds to the value of assets held by local and state pension funds in the home-state of the company expressed in 1982 billion U.S. dollars, lagged by 1 year. $Pension\_Norm$ corresponds to the value of assets held by local and state pension funds in the home-state of the company expressed 1982 billion U.S. dollars, normalized by average state GDP, and lagged by 1 year. $\Delta Pension$ corresponds to the value of assets held by local and state pension funds in the home-state of the company expressed in 1982 billion U.S. dollars, lagged by 1 year, and demeaned by time. b_t corresponds to average citations received at year *t* by matching patents in the same technology-class, with the same application-year and issued in the same state. The estimated coefficients for columns (1)-(4) are incidence rates, except for the estimated coefficient of the first-stage. For the incidence rates a coefficient greater than one corresponds to a positive relationship between the explanatory variable and the citation intensity. Standard errors are clustered at the company level. Standard errors for the First Stage are clustered at the state level. *, **, and *** indicate statistical significance at the 10%, 5% and 1% level.

	(1)	(2)	(3)	(4)	(4)	(5)	(6)	(7)
Dependent Variable	Cites	Cites	Cites	Cites	Scaled Cites	Scaled Cites	Scaled Cites	Scaled Cites
Instrument		$Pension$	$Pension\_Norm$	$\Delta Pension$		$Pension$	$Pension\_Norm$	$\Delta Pension$
Model	Poisson	GMM-IVs	GMM-IVs	GMM-IVs	OLS	2SLS	2SLS	2SLS
VC_{pt}	1.258***	1.435**	1.428**	1.513**	0.277**	0.485**	0.435*	0.506**
	(0.067)	(0.219)	(0.223)	(0.263)	(0.123)	(0.215)	(0.256)	(0.230)
First Stage						0.007***	3.066	0.007***
						(0.001)	(8.818)	(0.001)
F-test						41.50	12.09	92.86
Obs.	11,726	11,726	11,726	11,726	14,820	14,810	14,810	14,810
# of patents	1,288	1,288	1,288	1,288	1,513	1,503	1,503	1,503
# companies	456	456	456	456	490	490	490	490
# of states	32	32	32	32	32	32	32	32
Offset b_t	Yes	Yes	Yes	Yes				
Patent FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Table 8– Second robustness check IV analysis VC investments and patent citations

This table contains Poisson and Linear regression coefficients of the IV analysis. An observation is a patent-year. The dependent variable is indicated on top of each column. The instrument used in each model is indicated on top of each column. Scaled Out-state Cites corresponds to the number of citations a patent receives from assignees in a different state divided by the average number of out-state citations received by matching patents in the same technology-class, with the same application-year and issued in the same state as the patent. VC_{pt} is a dummy variable that equals one after the VC investment. $Pension$ corresponds to the value of assets held by local and state pension funds in the home-state of the company expressed in 1982 billion U.S. dollars, lagged by 1 year. $Pension\_Norm$ corresponds to the value of assets held by local and state pension funds in the home-state of the company expressed in 1982 billion U.S. dollars, normalized by average state GDP, and lagged by 1 year. $\Delta Pension$ corresponds to the value of assets held by local and state pension funds in the home-state of the company expressed in 1982 billion U.S. dollars, lagged by 1 year, and demeaned by time. b_t corresponds to average citations received at year t by matching patents in the same technology-class, with the same application-year and issued in the same state. The estimated coefficients for columns (1)-(4) are incidence rates, except for the estimated coefficient of the first-stage. For the incidence rates a coefficient greater than one corresponds to a positive relationship between the explanatory variable and the citation intensity. Standard errors are clustered at the company level. Standard errors for the First Stage are clustered at the state level. *, **, and *** indicate statistical significance at the 10%, 5% and 1% level.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dependent Variable	Out-state Cites	Out-state Cites	Out-state Cites	Out-state Cites	Scaled Out-state Cites	Scaled Out-state Cites	Scaled Out-state Cites	Scaled Out-state Cites
Instrument		$Pension$	$Pension\_Norm$	$\Delta Pension$		$Pension$	$Pension\_Norm$	$\Delta Pension$
Model	Poisson	GMM-IVs	GMM-IVs	GMM-IVs	OLS	2SLS	2SLS	2SLS
VC_{pt}	1.292***	1.384**	1.451***	1.457**	0.310**	0.670**	0.614**	0.704**
	(0.065)	(0.214)	(0.205)	(0.262)	(0.123)	(0.262)	(0.276)	(0.286)
First Stage						0.006***	3.061***	0.007***
						(0.001)	(0.902)	(0.001)
F-test						43.70	11.56	101.67
Observations	11,020	11,020	11,020	11,020	14,566	14,566	14,566	14,566
# patents	1,223	1,223	1,223	1,223	1,500	1,500	1,500	1,500
# companies	447	447	447	447	490	490	490	490
# states	31	31	31	31	32	32	32	32
Offset b_t	Yes	Yes	Yes	Yes				
Patent FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Table 9- Originality

This table reports Originality and Relative Originality measures for patents that are funded in hot versus cold markets. A patent is said to have been financed in a hot market if the size of local public pension funds' assets in the home-state of the company is above the 75th percentile of the state's average the year of the VC investment. Analogously, a patent is said to have been financed in a cold market if the size of local public pension funds' assets in the home-state of the company is below the 25th percentile of the state's average the year of the VC investment. *, **, and *** indicate statistical significance at the 10%, 5% and 1% level.

	Hot Market	Cold Market	
	Top 75%	Bottom 25%	Difference
Originality	0.55	0.37	0.18***
Originality Adjusted	0.64	0.45	0.18***
Relative Originality	0.15	0.09	0.06***
Relative Originality Adjusted	0.15	0.10	0.06**

Table 10– Univariate Tests VC investments and patent citations inside and outside VC portfolios

Panel A compares average annual portfolio-linked citations to patents to the average annual portfolio-linked citation baseline before and after the VC investment. Panel B compares average annual non-portfolio-linked citations to patents to the annual average non-portfolio-linked citation baseline before and after the VC investment. *, **, and *** indicate statistical significance at the 10%, 5% and 1% level.

Panel A. Portfolio-Linked Citations

	Annual Average			
	Pre-VC	Post-VC	Difference	Ratio
Citations	0.002	0.008	0.006***	4.05
	(0.08)	(0.15)		
Citation Baseline	0.001	0.003	0.002***	2.75
	(0.61)	(0.83)		
Difference	0.001	0.005***		
Ratio	2.17	2.88		
			Difference in Difference	Ratio of Ratios
			0.004***	1.47

Panel B. Non-Portfolio-Linked Citations

	Annual Average			
	Pre-VC	Post-VC	Difference	Ratio
Citations	0.64	1.03	0.39***	1.62
	(1.68)	(2.68)		
Citation Baseline	0.54	0.66	0.12***	1.22
	(0.61)	(0.83)		
Difference	0.11***	1.37***		
Ratio	1.19	1.56		
			Difference in Difference	Ratio of Ratios
			0.27***	1.33

Table 11 – Poisson Regressions VC investments and patent citations inside and outside VC portfolios

The table presents Poisson estimates where the effect of VC financing is allowed to affect differently citations that originate inside or outside VC portfolios. An observation is at the patent, year, and type of citation level. The dependent variable is annual citations. $Portfolio-linked$ ($Non-portfolio-linked$) is a dummy that equals one if the type of citation is portfolio- linked (non portfolio- linked). VC_{pt} is a dummy that equals one after the VC investment. b_{tC} corresponds to average citations of type C, where $C=\{Portfolio-linked, Non-portfolio-linked\}$, received at year t by matching patents in the same technology-class and with the same application-year. The Poisson model requires that annual average citations to matching patents be different from zero, which explains the difference in observations across columns (1)-(2). The fixed-effects Poisson model requires variation in the dependent variable for each patent-type of citation group for estimation, which explains the difference in observations across columns (2), and, (3). The reported coefficients are incidence rates. A coefficient greater than one corresponds to a positive relationship between the explanatory variable and the citation intensity. Standard errors are clustered at the patent level and reported in parenthesis. *, **, and *** indicate statistical significance at the 10%, 5% and 1% levels, respectively.

	(1)	(2)	(3)
Model	Poisson	Poisson	Poisson
A. Estimated Coefficients			
$Non-Portfolio-linked$	0.635***	1.176***	
	(0.024)	(0.040)	
$Portfolio-linked$	0.002***	0.868	
	(0.001)	(0.497)	
$VC_{pt}*Non-Portfolio-linked$ (I)	1.620***	1.325***	1.186***
	(0.067)	(0.050)	(0.035)
$VC_{pt}*Portfolio-linked$ (II)	4.052***	2.437	2.785**
	(1.619)	(1.541)	(1.276)
B. Difference in Coefficients			
Chi2	5.37	0.93	3.44
p- value Chi2 test	(0.03)	(0.34)	(0.06)
Observations	87,038	45,064	39,299
# of patents	2,336	2,336	2,183
# of companies	752	752	726
Offset b_{tC}	No	Yes	Yes
Patent-type of citation FE	No	No	Yes

Table 12 – IV analysis VC investments and patent citations inside and outside VC portfolios

This table contains Poisson and Linear regression coefficients of the IV analysis. The dependent variable is indicated on top of each column. An observation is at the patent- year level. VC_{pt} is a dummy that equals one after the VC investment. b_{tC} corresponds to average citations of type C, where C is either portfolio- or non portfolio-linked citations, received at year t by matching patents in the same technology-class and with the same application-year. The GMM-IVs and 2SLS specifications use $Pension$, the size of local and state pension funds' real assets in the home-state of the company lagged by one year, to instrument VC_{pt}. For columns (1)-(4) standard errors are clustered at the patent level and reported in parenthesis. For columns (5)-(8) standard errors are clustered at the state level and reported in parenthesis. *, **, and *** indicate statistical significance at the 10%, 5% and 1% levels, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dependent variable	Portfolio-linked Citations	Portfolio-linked Citations	Portfolio-linked Scaled Citations	Portfolio-linked Scaled Citations	Non-portfolio-linked Citations	Non-portfolio-linked Citations	Non-portfolio-linked Scaled Citations	Non-portfolio-linked Scaled Citations
Model	Poisson	GMM-IVs	OLS	2SLS	Poisson	GMM-IVs	OLS	2SLS
VC_{pt}	2.390*	6.274	1.361	5.113	1.167***	1.463***	0.224*	0.596***
	(1.228)	(14.289)	(1.908)	(3.747)	(0.039)	(0.191)	(0.132)	(0.215)
First Stage				0.007***				0.006***
				(0.000)				(0.001)
F test				437.56				31.13
Observations	115	115	2,617	2,513	15,622	15,622	18,928	18,928
# of patents	15	15	590	486	1,487	1,487	1,657	1,657
# of companies	9	9	235	235	491	491	517	517
# of states	3	3	25	25	38	38	39	39
Offset b_{tC}	Yes	Yes			Yes	Yes		
Patent FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Table 13– Univariate Tests VC investments and inventor and non-inventor-linked citations

The table presents average annual inventor- and non-inventor- linked citations. Panel A, reports citations from all assignees. Panel B, reports portfolio-linked citations. Panel C reports non portfolio-linked citations. Column (5) presents the ratio between average annual non- inventor-linked citations post VC-investment, and average annual citations pre VC-investment for patents. Column (6) presents the "ratio of ratios" defined as the ratio between the ratio of average annual non- inventor-linked citations post and pre VC-financing for patents, to the ratio of average annual non-nventor-linked citations post and pre VC-financing for matching patents. *, **, and *** indicate statistical significance at the 10%, 5% and 1% level.

Panel A. All Citations

	Average Annual Citations				Average Annual Citation Baseline				Ratio	Ratio of Ratios
	Pre-VC		Post-VC		Pre-VC		Post-VC			
	(1)		(2)		(3)		(4)		(5)	(6)
Non- inventor-linked	0.64	[1.69]	0.97	[2.53]	0.54	[0.61]	0.58	[0.83]	1.52	1.411
Inventor-linked			0.07	[0.56]			0.08	[0.14]		

Panel B. Portfolio-linked Citations

	Average Annual Citations				Average Annual Citation Baseline				Ratio	Ratio of Ratios
	Pre-VC		Post-VC		Pre-VC		Post-VC			
	(1)		(2)		(3)		(4)		(5)	(6)
Non-inventor-linked	0.0020	[0.08]	0.0076	[0.14]	0.0009	[0.08]	0.0024	[0.02]	3.87	1.459
Inventor-linked			0.0004	[0.03]			0.0004			

Panel C. Non-Portfolio-linked Citations

	Average Annual Citations				Average Annual Citation Baseline				Ratio	Ratio of Ratios
	Pre-VC		Post-VC		Pre-VC		Post-VC			
	(1)		(2)		(3)		(4)		(5)	(6)
Non-inventor-linked	0.63	[1.68]	0.96	[2.52]	0.54	[0.61]	0.58	[0.72]	1.52	1.408
Inventor-linked			0.07	[0.55]			0.08			

Table 14- Poisson Regressions VC investments and non-inventor-linked citations inside and outside VC portfolios

The table presents Poisson regression coefficients where the effect of the VC investment is allowed to affect differently non-inventor-linked citations that originate inside or outside VC portfolios. An observation is at the patent, year, and type of citation level. The dependent variable is annual non-inventor-linked citations. $Portfolio-linked$ ($Non-portfolio-linked$) is a dummy that equals one if the type of citation is portfolio-linked (non portfolio-linked). VC_{pt} is a dummy that equals one after the VC investment. b_{tC} corresponds to average citations received at year t by matching patents of citations type C, where $C=\{Portfolio-linked, Non-portfolio-linked\}$. The Poisson model requires that annual average citations to matching patents be different from zero, which explains the difference in observations across columns (1)-(2). The fixed-effects Poisson model requires variation in the dependent variable for each patent-type of citation group for estimation, which explains the difference in observations across columns (2), and, (3). The reported coefficients are incidence rates. A coefficient greater than one corresponds to a positive relationship between the explanatory variable and the citation intensity. Standard errors are clustered at the patent level and reported in parenthesis. *, **, and *** indicate statistical significance at the 10%, 5% and 1% levels, respectively.

	(1)	(2)	(3)
Model	Poisson	Poisson	Poisson
A. Estimated Coefficients			
$Non-Portfolio-linked$	0.635***	1.176***	
	(0.024)	(0.040)	
$Portfolio-linked$	0.002***	0.868	
	(0.001)	(0.497)	
$VC_{pt}*Non-Portfolio-linked$ (I)	1.517***	1.409***	1.281***
	(0.063)	(0.053)	(0.038)
$VC_{pt}*Portfolio-linked$ (II)	3.869***	2.620	2.891**
	(1.557)	(1.667)	(1.394)
B. Difference in Coefficients II-I			
Chi2	5.51	0.96	2.83
p- value Chi2 test	0.02	0.33	0.09
Observations	87,038	44,991	39,115
Number of patents	2,336	2,336	2,170
Number of companies	752	752	726
Offset b_{tC}	No	Yes	Yes
Patent-type of citation FE	No	No	Yes

Table 15 – IV analysis VC investments and non-inventor-linked citations inside and outside VC portfolios

This table contains linear regression coefficients of the instrumental variable analysis of non-inventor-linked inside and outside VC portfolios. The dependent variable is indicated on top of each column. An observation is at the patent- year level. VC_{pt} is a dummy that equals one after the VC investment. b_{tC} corresponds to average citations of type C, where C is either (non-inventor-linked) portfolio- or non-portfolio-linked citations, received at year t by matching patents in the same technology-class and with the same application-year. The GMM-IVs and 2SLS specifications use $Pension$, the size of local and state pension funds' real assets in the home-state of the company lagged by one year, to instrument VC_{pt}. For columns (1)-(4) standard errors are clustered at the patent level and reported in parenthesis. For columns (5)-(8) standard errors are clustered at the state level and reported in parenthesis. *, **, and *** indicate statistical significance at the 10%, 5% and 1% levels, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dependent variable	Portfolio-linked Citations	Portfolio-linked Citations	Portfolio-linked Scaled Citations	Portfolio-linked Scaled Citations	Non-portfolio-linked Citations	Non-portfolio-linked Citations	Non-portfolio-linked Scaled Citations	Non-portfolio-linked Scaled Citations
Model	Poisson	GMM-IVs	OLS	2SLS	Poisson	GMM-IVs	OLS	2SLS
VC_{pt}	2.934	4.081	0.705	2.481*	1.274***	1.595***	0.325**	0.684***
	(2.021)	(14.374)	(0.581)	(1.493)	(0.076)	(0.235)	(0.140)	(0.227)
First Stage				0.007***		0.710***		0.006***
				(0.0003)		(0.13)		(0.001)
F-test				437.56		30.04		31.13
Observations	109	109	2,617	2,513	15,516	15,516	18,928	18,928
# of patents	15	15	590	486	1,475	1,475	1,657	1,657
# of companies	9	9	210	210	491	491	517	517
# of states	3	3	24	24	38	38	39	39
Offset b_{tC}	Yes	Yes			Yes	Yes		
Patent FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Table 16 – Summary Statistics Patent Sales

This table reports summary statistics of patent sales around the VC investment. Panel A reports number of patents that were sold before and after the VC investment. Panel B compares patents and matching patents and their respective likelihood of being sold at least once throughout the sample. Standard deviations are included in parenthesis. The number of observations is reported in squared brackets. Panel C compares average annual citations for patents and the citation baseline according to whether the patents were sold or not throughout the sample.

Panel A. Number of patents sold

	Number	Percentage of Total
Total patents sold during the sample	375	16%
Patents sold at least once before the VC investment	62	3%
Patents sold at least once after the VC investment	327	14%

Panel B. Annual Likelihood that a patent is traded (percentage)

Patents		Matching Patents		Difference	Diff. in Diff.
Pre-VC	Post-VC	Pre-VC	Post-VC		
0.51	0.92	0.37	0.31	0.41***	0.477***
(7.12)	(10.21)	(0.75)	(0.60)		
[12,767]	[40,096]	[12,767]	[40,096]		

Panel C. Difference in citations to patents

	Annual Average Citations		Annual Average Citation Baseline		Diff.	Ratio	Diff.-in-Diff.	Ratio of Ratios
	Pre-VC	Post-VC	Pre-VC	Post-VC				
Sold	0.75	1.20	0.54	0.68	0.45***	1.61	0.31***	1.28
	(2.11)	(3.08)	(0.65)	(0.86)				
	[1,902]	[5,257]	[1,902]	[5,257]				
Not Sold	0.62	1.00	0.54	0.66	0.39***	1.62	0.27***	1.33
	(1.60)	(2.61)	(0.60)	(0.82)				
	[10,865]	[25,495]	[10,865]	[25,495]				
Sold - Not Sold	0.13***	0.20***	0.01	0.03***	0.07	-0.02	0.05	-0.05
	[12,767]	[30,752]	[12,767]	[30,752]				

Appendix: Data Construction

My starting point is the universe of transactions registered in VentureXpert that closed between January 1976 and December 2009. I eliminate four types of investments. First, VentureXpert contains transactions by private equity groups other than independent Venture Capital firms such as angel groups, bank affiliated firms, corporate venture capital firms, endowment foundations, pension funds, government affiliated programs, incubator development programs, individuals, insurance firm affiliates and investment management firms. While these transactions are part of the financial landscape for companies, they are not the focus of this study; hence, I eliminate them from the sample. Second, the data contain transactions by VC firms that are not focused on venture capital, such as buyout funds and funds of funds, and I eliminate these deals as well. I also remove investments by VC firms in companies that were already traded in public markets before the transaction (called PIPEs), and secondary purchases. Finally, I only include investments made by U.S. VC firms in U.S. companies. After these eliminations, the data contain 116,574 investments made to 20,058 U.S. based companies.

Capturing patent data

I match the companies involved in VC transactions to their patent records based on name. To do so, I employ the Harvard Business School (HBS) patent database. The HBS data contain all electronic records of the U.S. Patent and Trademark Office (USPTO) through December 2008, which have been cleaned and consolidated by HBS.[1] The patent database also has information on all citations made and received by patents as well as information on the inventors. I restrict my sample to primary assignments of utility patents (99%) awarded to US companies from 1976 onwards. After these restrictions the sample consists of 2,881,097 patents, awarded to 1,980,696 inventors, and issued to 242,767 U.S. assignees. The total number of citations made and received by these patents is 22,511,338.

In order to combine the two databases, I strip company names from VentureXpert, and assignee names from the HBS database, of punctuation, capitalization and common acronyms. I then match the samples on the normalized company and assignee names using a fuzzy-match procedure based on the Levenshtein edit distance. The Levenshtein edit distance is a measure of the degree of proximity between two strings, and corresponds to the number of substitutions, deletions or insertions needed to transform one string into the other one (and vice versa).[2] I assign a score for each match as a function of the Levenshtein edit distance and the length of each of the normalized company names in the match. Using a random sampling procedure, I determine a score threshold such that matches with scores above the threshold are hand checked, and those below the threshold are eliminated. During the manual check of the remaining matches, I check that the two companies are in the same state. There are ambiguous situations where the names are similar, but not exactly identical, or

[1] The database is documented in Lai, D'Amour, and Fleming (2009).

[2] For more information and an application to Perl see Text::LevenshteinXS in CPAN.

where the location of the patentee differs from that given in the records of SDC. In these cases, I research the potential matches using web searches. Finally, in some cases, there are multiple names in either of the bases that appear to match a single name in the other data set. For these, I add the observations into an aggregated entity.

Matched Sample

In total, I identify 5,018 companies that are VC-backed and with at least one U.S. utility patent grant. The total number of patents awarded to these companies from January 1976 to December 2008 is 105,484 patents. The total number of inventors in the sample is 74,666 inventors, and the total number citations made and received by these patents is 1,200,190.

The small number of matches between the two data sets likely reflects two facts. First, in many instances, specially more recently, the companies that are VC-backed belong to sectors in which IP is not usually protected using patents (e.g. internet, media, and software companies), and in which there is greater reliance on trade secrets to protect it. Second, VentureXpert includes data on all companies that received VC financing, including those that were not ultimately successful, and which may not have reached a stage in which IP should be protected.

Note that the 105,484 patents assigned to VC-backed companies correspond to less that 4% of U.S. patent stock. This stands to contrast with existing estimates from the venture capital literature on the patent stock of patents attributed to venture capital funds. For example, Kortum and Lerner (2000) analyze annual data from twenty manufacturing industries. The dependent variables is U.S. patents issued to U.S. inventors by industry and date of application. The main explanatory variables are measures of venture funding collected by Venture Economics. The authors estimate a patent production function and estimate the impact on patenting that a dollar of venture capital has relative to a dollar of R&D on industrial patent production. Using a back-of-the-envelope calculation of their findings, the author estimate that venture capital funds have accounted for approximately 14% of industrial patent production. More recent studies by Hirukawa and Ueda (2011) confirm the order of magnitude of this estimate. The difference between these macro-based estimates, and my micro-based quantification, suggests that VCs generate knowledge spillovers, and that their role on innovation goes above and beyond financing the patents of their targets.

Before I present summary statistics of the matched sample, a couple of points need to be clarified regarding institutional arrangement of patent data. There are two relevant dates associated with each patent: application and grant date. The application date marks the official date in which the inventor submitted the patent application to the USPTO office. The grant date is the date in which the patent was issued to the inventor. For patents applied for before October 2000, their content was made public the first Tuesday after grant date in the USPTO's official magazine. For patents applied for after October 2000, the American Inventor Protection Act (enacted on November 29 1999) specifies they are to be disclosed 18 months after application. Nevertheless, citations to patents start as early as the application year, which can be partially explained by technical disclosures, or diffusion of new technologies via conferences or connections among agents.

Summary Statistics

Table A1 presents summary statistics of the matched sample. Panel A shows an apparent decrease in patent applications by VC-backed companies starting on 2002. The reason for this decrease is the well documented lag between the application and the grant of a patent by the USPTO office.

For patents issued after 1976 and granted to any (VC-backed) patentee by 2008, the lag is 2.30 (2.75) years. The difference in the lag between Non VC- and VC-backed assignees is not significant. Panel A also shows an apparent decrease in the number of investments by VC-backed companies. This decrease is due in part to the expansion of investments in sectors such as internet and media that do not generally rely on patent protection, and not to a real decrease in the number of total investments by VCs. Panel B exhibits the distribution of patents and VC-backed companies that patent by state. As it is common in the VC literature, the sample is concentrated in California, Massachusetts, Washington and Texas. Panel C shows the distribution of type of first time investments by VC firms on companies that patent. The types of investments include traditional VC investments such as: Bridge Loans, Early Stage, Expansion, Later Stage and Seed. Panel D shows the distribution of companies that patent by industry, according to the industry classification from SDC. The data is concentrated in Medical Health, Semiconductors and Computer Software. Finally, Panel E shows distribution of VC-backed companies that patent by type of VC exit. Approximately 50% of companies have a successful exit, either through an IPO or acquisition.

Table A2 compares patents from VC-backed companies and patents issued to Non VC-backed assignees. Panel A shows that patents assigned to VC-backed companies receive more citations three years following the grant date. This is true for both citations made by the same assignee (self-citations) and citations made by other assignees (no self citations).

The generality measure is an statistic used in the innovation literature to describe patents, and is constructed using information on the citations received by patents. A patent has a higher generality, if it is cited by subsequent patents that belong to a wide range of technology classes. Thus, a high generality score suggests the patent presumably had a widespread impact, in that it influences subsequent innovations in a variety of fields. The generality measure corresponds to one minus the Herfindahl index of the technology classes of the citing patents. Panel B, shows that patents assigned to VC-backed companies have higher generality measures three years following the grant date.

The originality measure is an additional statistic used in the innovation literature to describe patents, and is constructed using information on the citations made by patents. A patent has a higher originality, if it cites patents that belong to a wide range of technology classes. The intuition is that patents that combine existing knowledge from few technology classes to create something new (and useful), probably constitute more marginal improvements relative to patents that combine more ex-ante different ideas. This measure is constructed as one minus the Herfindahl index of the cited patents across technological classifications. Panel C shows that VC-backed patents are on average more original.

Table A1 - Summary statistics of matched sample

The matched full sample consists of 105,484 patents awarded between 1976 and December 2008 to 5,018 companies that were financed by at least one U.S. VC firm during 1976 to 2009.

Panel A. Application and grant years of patents issued by VC-backed companies and total number of VC-backed companies by year of first VC transaction

	Patents		Companies	
	Applications	Grants	Number	Percentage
1976	247	3	20	0.4
1977	243	113	24	0.48
1978	258	225	29	0.58
1979	260	182	37	0.74
1980	246	232	77	1.53
1981	340	229	139	2.77
1982	348	217	113	2.25
1983	421	251	123	2.45
1984	518	369	138	2.75
1985	570	397	111	2.21
1986	696	463	103	2.05
1987	860	671	122	2.43
1988	1,007	699	112	2.23
1989	1,162	1,009	147	2.93
1990	1,321	976	100	1.99
1991	1,581	1,057	60	1.2
1992	1,939	1,325	77	1.53
1993	2,309	1,562	91	1.81
1994	3,166	1,814	95	1.89
1995	5,130	2,104	175	3.49
1996	5,405	2,689	214	4.26
1997	7,000	3,287	247	4.92
1998	7,354	5,288	295	5.88
1999	8,208	5,767	333	6.64
2000	9,825	6,433	497	9.9
2001	10,537	6,891	308	6.14
2002	10,583	7,424	245	4.88
2003	8,133	8,236	242	4.82
2004	7,379	7,961	236	4.7
2005	5,338	7,498	180	3.59
2006	2,430	10,139	134	2.67
2007	643	9,906	102	2.03
2008	27	10,067	67	1.34
2009			25	0.5
Total	105,484	105,484	5,018	

Panel B. Distribution of Patents and VC-backed companies by state

	Patents		Companies	
	Number	Percentage	Number	Percentage
AL	309	0.29	10	0.2
AR	1	0	1	0.02
AZ	562	0.53	47	0.94
CA	59,644	56.54	2,226	44.36
CO	1,275	1.21	137	2.73
CT	796	0.75	84	1.67
DC	72	0.07	8	0.16
DE	36	0.03	2	0.04
FL	674	0.64	75	1.49
GA	469	0.44	88	1.75
HI	6	0.01	2	0.04
IA	25	0.02	9	0.18
ID	58	0.05	7	0.14
IL	671	0.64	97	1.93
IN	332	0.31	15	0.3
KS	14	0.01	8	0.16
KY	11	0.01	4	0.08
LA	29	0.03	6	0.12
MA	9,469	8.98	643	12.81
MD	939	0.89	106	2.11
ME	13	0.01	3	0.06
MI	303	0.29	43	0.86
MN	2,713	2.57	91	1.81
MO	157	0.15	23	0.46
MS	16	0.02	5	0.1
MT	5	0	1	0.02
NC	882	0.84	80	1.59
ND	6	0.01	1	0.02
NE	20	0.02	3	0.06
NH	492	0.47	49	0.98
NJ	1,198	1.14	129	2.57
NM	52	0.05	14	0.28
NV	67	0.06	9	0.18
NY	1,905	1.81	146	2.91
OH	538	0.51	62	1.24
OK	63	0.06	10	0.2
OR	523	0.5	55	1.1
PA	2,370	2.25	155	3.09
RI	54	0.05	12	0.24
SC	19	0.02	6	0.12
SD	4	0	1	0.02
TN	164	0.16	21	0.42
TX	6,206	5.88	243	4.84
UT	204	0.19	33	0.66
VA	483	0.46	69	1.38
VT	26	0.02	3	0.06
WA	11,242	10.66	144	2.87
WI	359	0.34	28	0.56
WV	2	0	2	0.04
WY	6	0.01	2	0.04
Total	105,484		5,018	

Panel C. Distribution of type of investment by VC firms in companies that patent

Type of Investment	Number of deals	Percentage of sample
Bridge Loan	85	1.69
Early Stage	1,917	38.2
Expansion	1,269	25.29
Later Stage	350	6.97
Seed	1,397	27.84
Total	5,018	

Panel D. Industry distribution of VC investments in companies that patent

	Number of companies	Percentage of sample
Biotechnology	495	9.86
Communications and Media	554	11.04
Computer Hardware	446	8.89
Computer Software	819	16.32
Consumer Related	101	2.01
Industrial Energy	400	7.97
Internet Specific	425	8.47
Medical Health	842	16.78
Other Products	131	2.61
Semiconductors	805	16.04
Total	5,018	

Panel E. Distribution of VC-backed companies with prior patents by type of VC exit

	Number of companies	Percentage of sample
Acquisition	1,722	34.32
Active	1,537	30.63
Bankruptcy - Chapter 11	23	0.46
Bankruptcy - Chapter 7	38	0.76
Defunct	726	14.47
In Registration	20	0.4
LBO	37	0.74
Merger	82	1.63
Other	20	0.4
Pending Acquisition	7	0.14
Went Public	806	16.06
Total	5,018	

Table A2 – Comparison Patents from VC-backed versus Non VC-backed patents

The full matched sample consists of 105,484 patents awarded through December 2008 to 5,018 companies that received VC backing between 1976 and 2009Panel B, presents citation counts 3 years following the grant date, and excludes from the analysis patents granted after 2005. Panel C, presents Generality measures using the USPTO technological classification and the Hall bias correction (Hall, et al. 2001). Panel D, presents Generality measures for citations 3 years following the grant date. Panel E, presents Originality measures. See Appendix 1 for a detailed definition of the variables.

Panel A. Total Citations, Self-citations and No-self citations until 3 years after grant date

	Three-year Citations			Three-year Self Citations			Three-year No Self Citations			
	Mean	S.D.	Med.	Mean	S.D.	Med.	Mean	S.D.	Med.	Obs.
VC-backed	7.4	13.79	3	1.23	4.15	0	6.17	12.22	2	95,110
Non VC-backed	3.32	6.56	1	0.53	1.91	0	2.79	5.95	1	2,652,052
p-value t-test	0.00			0.00			0.00			

Panel B. Generality, Self Generality and No-self generality until 3 years after grant date

	Three-year Generality				Three-year Self Generality				Three-year No Self Generality			
	Mean	S.D.	Med.	Obs.	Mean	S.D.	Med.	Obs.	Mean	S.D.	Med.	Obs.
VC-backed	0.40	0.28	0.48	64,946	0.12	0.22	0.00	28,540	0.39	0.29	0.47	61,648
Non VC-backed	0.28	0.28	0.28	1,734,688	0.20	0.27	0.00	587,468	0.26	0.28	0.17	1,605,061
p-value t-test	0.00				0.00				0.00			

Panel C. Originality

	Originality				Originality Adjusted			
	Mean	S. D.	Med.	Obs.	Mean	S. D.	Med.	Obs.
VC-backed	0.455	0.283	0.5	105,484	0.56	0.31	0.50	99,551
Non VC-backed	0.305	0.293	0.32	2,775,613	0.436	0.37	0.32	2,451,091
p-value t-test	0.00				0.00			