Venture Capital and the Diffusion of Knowledge

Juanita González-Uribe*

Columbia GSB

Preliminary and incomplete

This version: September 20, 2012

Abstract

This paper finds that Venture Capital (VC) increases the diffusion of knowledge. Using a data-set of patents issued by companies before they are financed by a VC, I find that citations to patents increase after VC investments, relative to a control group of patents. To isolate the causal effect, I use fluctuations in the size of state public pension funds as exogenous variation in the timing of VC investments. Additional results are consistent with VCs facilitating the transfer of knowledge inside their portfolios, and certifying the quality of their targets' patents.

*Columbia Graduate School of Business, e-mail: jgonzalez14@gsb.columbia.edu. I am indebted to my advisor Morten Sorensen 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 Arvid Ziedonis and Rosemary Ziedonis for insightful comments. I am grateful for helpful comments from seminar participants at the Fifth Annual Searle Center Conference on Innovation and Entrepreneurship, Universidad de los Andes Colombia Business and Economics Lunch, and Universidad Católica de Chile Business and Economics Lunch. Finally, I am also grateful for funding from the Ewing Marion Kauffman Foundation.

Do Venture Capitalists (VCs) affect the diffusion of knowledge? Although a large body of research examines the determinants of knowledge diffusion, the role of VCs has not been systematically addressed before. This question is important given the critical role of knowledge diffusion in promoting economic growth (Romer 1986, 1990; Lucas, 1988; Krugman, 1991; Grossman and Helpman, 1991), and given that current estimates of the impact of VCs on innovation far exceed the share of patents assigned to their targets. The difference between the 14% of patent production attributed to VC investments (Kortum and Lerner, 2000), and the 4% of patents granted to VC-backed companies, suggests that the role of VCs on innovation is not limited to financing the innovation of their targets.[1] In this paper I show that VCs also affect innovation by facilitating the diffusion of their targets' patented knowledge, and that this effect explains at least 5% of U.S. patent production.

VCs are financial intermediaries that undertake equity-linked investments in young privately held companies. In the U.S. they are the dominant form of equity financing for privately held high-technology businesses.[2] There are at least two characteristics of their investment strategy that can bear on the diffusion of the knowledge created inside their targets. First, VCs are highly selective in their investments. By selecting among innovations those with higher prospects of commercialization, VCs communicate to other inventors the commercial value of a company's knowledge. This "certification effect" can increase awareness and interest in the company's intellectual property (IP), and affect it's diffusion.[3] Second, VCs are active investors that help their targets through advice, support and improved governance (Sahlman, 1990; Lerner, 1995; Hellman and Puri, 2002; Hochberg

[1] Other papers that estimate a large role for VCs on innovation are Hirukawa and Ueda, (2011), Popov and Roosenboom (2011) and Mollica and Zingales (2007).

[2] For the past few years, other forms of equity financing have been gaining importance. For instance, Seed and Angel financing represent 35% of all attributed investments in Web 2.0 startups since 2005 (Source, Crunchbase, www.crunchbase.com). This trend is expected to continue specially after the passing of the Jumpstart Our Business Startups Act (JOBS Act) on April 2012, that will facilitate crowdfunding in privately-held companies.

[3] Megginson and Weiss (1989), show a similar role of VCs in certifying a company's quality among underwriters and its impact on IPO underpricing.

et al. 2007; Lindsey, 2007; Sørensen, 2007). This "active investor effect" of VCs can help diffuse their targets' knowledge in many ways. For instance, by providing a platform of interaction among companies, VCs can facilitate the transfer of tacit knowledge inside their portfolios.

To show that VCs affect the diffusion of knowledge, I follow the innovation literature and use data on forward citations to patents as a measure of knowledge diffusion (Jaffe, 1986; Jaffe et al., 1993; Jaffe and Trajtenberg, 2002; Hall et al., 2001). I construct a data set of "prior" patents that were issued by companies at least two years before they are financed by a VC. My empirical strategy explores how citations to prior patents change after VCs invest in the issuing companies. My main finding is that there is a causal increase in citations to prior patents after companies are financed by a VC.

I explore three mechanisms for knowledge diffusion that can be affected by VCs: company networks, inventor mobility and IP trade. Consistent with the active investor effect of VCs, I find suggestive evidence that the increase in citations is concentrated inside VC portfolios. Additional results suggest that this concentration is associated to inventor turnover across companies financed by the same VC. Consistent with the certification effect of VCs, I find that the increase in citations is not exclusive to VC portfolios or to the VC industry, cannot be entirely traced back to inventor mobility, and that after the VC investment, companies are more likely to sell their prior patents.

There are three potential sources of bias in identifying the effect of VCs on citations to prior patents. The first is that the likelihood of receiving a citation is not constant throughout time, or throughout a patent's life. This could lead to an upward bias in the correlation between VC financing and changes in citations to prior patents. To address this concern, my methodology relies on relative measures of citations, by using a set of matching patents at the technology-class and application- year level as control group.

The second potential source of bias is unobserved patent heterogeneity. Quality differences across patents can bias cross-section estimates of the correlation between VC financing and changes in relative citations. To overcome this concern, I only exploit information on within-patent changes in citations to prior patents, relative to the control group. My empirical strategy estimates the increase in the likelihood of a citation to the same prior patent after the issuing company is first financed by a VC, and, relative to other patents in the same technology-class and application year. Using a conditional patent fixed-effects Poisson model (QMLE) based on this intuition, I find that relative to similar patents, the likelihood of a citation to a prior patent increases by 18.9% after a VC invests for the first time in the issuing company.

The third potential source of bias in identifying the causal effect of VCs on citations to prior patents, is that the timing in which companies are selected by VCs may not be random. For instance, VCs may be able to anticipate which patents will be more likely to be cited in the future, and invest based on that prediction. This could generate an upward "timing bias" in the correlation between VC financing and within-patent increases in relative citations.

To overcome this identification challenge, I follow the Private Equity (PE) literature and use variations in the size of public pension funds in the home-state of companies to instrument for the timing in which companies are first financed by a VC. Public pension funds are among the largest sponsors of the VC industry and they are home biased in their PE investments (Hochberg and Rauh, 2011). In addition, there is substantial evidence that VCs also tend to invest locally (Lerner, 1995; Sorenson and Stuart, 2001). Thus, the instrument relies on the sensitivity of domestic VC investments to variations in the size of local public pension funds.

One concern regarding the instrument is that variations in state pension funds' assets may be indicative of innovation opportunities within the state. If that is the case, then the instrument and changes in relative citations to prior patents may be correlated via a state-effect, and not exclusively through the VC financing channel. I address this potential violation of the exclusion restriction in two ways. First, I define relative citations at the state-level by restricting the matching patents to have been issued in the same state. This robustness check is useful because if in fact variations in pension funds contain information about state-level changes in innovation opportunities, such changes should affect equally all patents issued in the same state, and is therefore unlikely to affect relative measures of citations at the state level.

Second, in addition to using measures of citations at the state-level, I also restrict the dependent variable to out-state citations*:* citations to patents from patentees that are located in a different state. This second robustness-check is useful because if variations in the size of state pension funds indicate local innovation opportunities, this state-effect should specially reflect on increases in citations from other patentees inside the same state. Hence, by excluding domestic citations from the dependent variable, the concern of a potential violation of the exclusion restriction is further minimized.

Using a Generalized Method of Moments (GMM-IVs) approach to estimate a fixed-effects Poisson model with endogenous regressors, I find that results continue to hold. However, contrary to the timing bias prediction, and similar to other papers in the VC literature that instrument VC investments with shocks to the availability of funds for VCs (Kortum and Lerner, 2000; Mollica and Zingales, 2007; Bernstein et al., 2011), I report GMM-IVs estimates that exceed QMLE results. A potential explanation behind this result, is that there is underlying heterogeneity in the effect of VC financing on the diffusion of knowledge, and that my GMM-IVs estimates are recovering the effect for a particularly sensitive group of patents. Since the GMM-IVs approach estimates the effect of VC financing only on those patents that respond to the instrument, if shocks to the capital available for VCs trigger investments on a sub-population of patents with a relatively high marginal return from VC selection, it is clear that the GMM-IVs estimates can exceed the QMLE results.

One reason why companies that are financed in "hot markets" may have patents that are more sensible to the VC investment, is that the abundance of capital allows investors to experiment more effectively, and shift the type of startups they finance towards those that are more novel. Consistent with VCs changing the type of investments they make as a function of available funds, I show that patents funded during hot markets are ex-ante more original than patents funded in "cold markets". My results are in accordance with Nanda and Rhodes-Kropf (2011), who find that VCs invest in riskier and more innovative projects when there is excess availability of funding capital for VCs.

For policy evaluation purposes, the concern that remains regards external validity. In particular, if the estimated effect using the GMM-IVs approach is not the average effect on a randomly picked patent, the question of how these results can inform policy remains. However, to inform policy, the effect of VC financing in the population may be less relevant than the effect for the group who will be impacted by a proposed reform. And, since growth policies generally stimulate VC financing by increasing funding capital, my results can be informative for current policy.

Having shown that citations to prior patents causally increase after companies are first financed by a VC, I turn to analyzing the mechanisms that underlie this effect. I start by exploring whether the interactions of companies inside VC portfolios facilitate the transfer of knowledge. I classify citations to patents as portfolio-linked if they originate in a company financed by the same VC as the target, and not portfolio- linked otherwise. I then estimate Poisson regressions where I allow the two types of citations to be affected separately by the VC investment. Consistent with the certification effect of VCs, I find that the increase in citations is not exclusive to VC portfolios. Consistent with the active investor effect of VCs, I find suggestive evidence that the increase in portfolio- linked citations is larger than the increase in non portfolio- linked citations.

I then explore whether the increase in citations can be traced to mobility of inventors. I classify a citation to a patent as inventor- linked if there is at least one inventor that assigned a patent to the cited company before the VC investment, and at least one patent to the citing company afterwards, and as not inventor- linked otherwise. I then estimate Poisson regressions using as dependent variable not inventor- linked citations. Consistent with the certification effect of VCs, I find that not inventor- linked citations also causally increase after the VC investment, suggesting that the effect of VCs on citations to prior patents cannot be entirely attributed to observable inventor mobility across jobs. Consistent with the active investor effect of VCs, I find suggestive evidence that the increase in citations inside VC portfolios can be attributed to inventor turnover across companies financed by the same VC.

Finally, I explore whether patent sales increase after the VC investment, and whether patent trade explains the effect of VCs on citations to prior patents. Consistent with the certification effect of VCs, I find that the probability that a prior patent is sold increases after the VC investment, however, I show that it cannot entirely explain the increase in citations.

This paper contributes to the literature that considers the role of financial intermediaries in innovation (Kortum and Lerner, 1998; Mollica and Zingales, 2007; Popov and Roosenboom, 2011; Hirukawa and Ueda, 2011; Lerner et al. 2011, Bernstein, 2011, Nanda and Rhodes-Kropf, 2011; Bernstein, 2012, Seru, 2012). I extend this work by exploring one crucial aspect of the innovation process, the transfer of knowledge across agents. A back of the envelope calculation of my findings, estimates that for every patent issued by a VC-backed company, 1.13 additional patents in the economy exploit the same knowledge to generate a new product. After including this effect, the share of patents that can be attributed to VCs increases from 4% to 9%, and decreases the gap between micro and macro level estimates of the effect of VCs on innovation.

This paper is also related to the literature that explores the value-added role of VCs. Previously documented mechanisms include VCs helping their companies recruit key managers (Hellman and Puri, 2002), implementing governance mechanisms (Hochberg, 2004) and facilitating strategic alliances (Lindsey, 2007). I find evidence that VCs facilitate the transfer of knowledge inside their portfolios. This can add value for their targets by changing the complementary assets available to support companies (Hellman, 2002). I also find suggestive evidence that part of this diffusion can be traced to inventor turnover among companies financed by the same VC. Exploring the prevalence of internal labor markets in VC firms is an interesting avenue for future research.

The rest of this paper is organized as follows. Section 1, explains the data sources used to construct my sample, and presents summary statistics. In Section 2, I discuss the empirical strategy used in this paper to identify the effect of VCs on knowledge diffusion. Section 3 explores the mechanisms behind the effect. Finally, Section 4 concludes.

1 Data Description and Summary Statistics

The sample is constructed using three main data sources: investment-level data on VC investments, assignment-level data on patent applications, and state-level data on the public pension asset pool. In this section, I give a brief overview of the steps I used to construct the sample and provide summary statistics. In Appendix 1 I include a detailed account of the procedure.

1.1 Analysis Sample

I collect data on VC investments from SDC's VentureXpert, and identify all U.S. companies that are financed by a U.S. VC firm from 1976 to 2008. Data on patent assignments to these companies comes from the Harvard Business School (HBS) patent database, which has data on U.S. patent assignments through December 2008. I combine the two data sources by searching for each of the VC-backed company names among the patent assignees. Appendix 1 has a detailed account of the matching procedure and includes summary statistics for the matched sample.

For the analysis in this paper, I restrict the data to companies with one or more successful patent applications at least two years before their first VC investment.[4] My empirical strategy studies changes in citations to these patents following the VC financing event. I consider only investments by VCs until 2003, in order to observe at least five years of citations to patents post VC financing. The analysis sample consists of 2,336 patents that were filed by 752 companies. I refer to these patents as prior patents throughout the paper.

In Table 1, I document summary statistics of the analysis sample, and check its representativeness relative to all other patents that are assigned to VC-backed companies, as well as relative to all other companies that are VC-backed and do not patent.

In Panel A, I first summarize the VC investments by year. The transactions are concentrated in the second half of 1990s until 2003. This reflects the increasing volume of transactions during these years. The absence of VC transactions before 1977 and after 2003 reflects the construction of the sample. The last reported investments of VCs in the companies with prior patents are also reported. The final investments lag the initial transactions by several years. On average VCs invest during 2.73 years in each of their

[4] In unreported results I restrict the sample to patents granted at least 2 years before they are financed by a VC. Results are robust to this change.

targets.

Panel A also displays the timing of applications and awards of the prior patents. Each patent is associated with two dates: the application date and the grant date. The application dates extend from 1976 (2 years before the first VC investment) to 2001 (2 years before the last VC investment). Patent grants lag applications by several years (on average 2.35 years in my sample), which reflects the patent application process in the USPTO office.

Panel B shows the distribution of companies and patents across the top 10 U.S. states in my sample. Compared to the full matched sample and the overall population of VC-backed companies, the analysis sample is underrepresented in California and Washington and overrepresented in Massachusetts, Pennsylvania and Texas.

Panel C, shows the distribution across types of investments. Investments in companies at Expansion Stage are the most common, followed by Early-Stage, Later-Stage and Seed. Compared to the full matched sample and the population of VC-backed companies, my sample is overrepresented in Expansion Stage companies and underrepresented in Seed Stage, which is consistent with more mature companies having a higher propensity to file for patents.

Panel D, shows the industry composition of companies, and Panel E the industry composition of patents. Patents are assigned to the primary industry of the company, as reported by VentureXpert. In later analyses, I use the patent-specific industry classification of the USPTO. Compared to the population of VC-backed companies, the analysis sample is overrepresented in Industrial Energy, Medical Health and Semiconductors, all of which are industries that are known to rely on patents as protection for their IP. The sample is underrepresented in the Computer Software and Internet Specific sectors which do not tend to use formal IP protection and rely on other mechanisms such as trade secrets.

Panel F, shows the distribution of VC-backed companies with prior patents by type of VC exit. The sample is overrepresented in companies that are acquired or go public. Finally, Panel G, shows the distribution of prior patents by age at the time of the VC investment. Approximately 80% of the patents in the sample have less than 10 years of age, and most are younger than 5 years.

Panel G has descriptive statistics on annual citations to prior patents which is the main variable in the analysis. On average, prior patents receive 0.92 citations by year. The distribution is highly skewed; median annual citations to prior patents are 0, and a few patents receive the bulk of annual citations.

1.2 Restricted Sample

The third source of data in this paper is the State and Local Government Public-Employee Retirement Systems annual survey conducted by the Census Bureau. The survey coverage includes public employee retirement systems administered by state and local governments throughout the nation. From this survey I collect information on financial assets (cash and financial holdings) held by these funds deflated by PPI. These data is only available from 1993 to 2008. I refer to the analysis sample restricted to the 1993-2008 period as the restricted sample throughout the paper.

Table 2 reports summary statistics on the restricted sample, which consists of 1,170 prior patents awarded to 434 VC-backed companies. Panel A, shows the distribution of application and grant years of prior patents, as well as the distribution of VC deals. The first two columns of Panel B reports the value of the assets held by local and state public pension funds deflated by the Produce Price Index and expressed in 2008 U.S. millions. The last four columns of Panel B reports the distribution of the restricted sample across states, and shows that the distribution across Top States mirrors the analysis sample.

Panel C, D and E, show that compared to the complete analysis sample, the restricted sample is overrepresented in Early Stage and Biotech companies. Consistent with the life cycle of companies, active (defunct) investments are also overrepresented (underrepresented) in the restricted sample. Panel F shows the distribution of prior patents by age at the time of the VC investment. By construction, since no patents filed before 1993 are included in the restricted sample, there are no patents older than 10 years of age.

Finally, Panel G reports descriptive statistics on annual citations to prior patents in the restricted sample. Compared to the analysis sample, average annual citations to prior patents increase, reflecting the overall increase in citations throughout the period.

2 Statistical Models of Knowledge Diffusion

In this section I develop a statistical framework for summarizing the evidence of VC finance on the diffusion of knowledge. I begin by specifying more precisely my measure of knowledge diffusion. I then describe the statistical model used to estimate the causal effect of VCs on the diffusion of their targets' knowledge and summarize results. Next, I interpret my findings in the context of the VC literature. Finally, I discuss the validity of my proxy for knowledge diffusion.

2.1 Measuring knowledge diffusion

Following the innovation literature, I use data on forward citations to patents as a measure of knowledge diffusion (Jaffe, 1986; Jaffe et al., 1993; Jaffe and Trajtenberg, 2002; Hall et al., 2001). When inventors file their patents at the USPTO, by law, they have to include a list of references to other patents upon which their invention builds on.[5]

[5] Citations serve a useful legal purpose; they help determine the scope of the property rights awarded to the patent. If patent B cites patent A, it implies that patent A represents a piece of previously existing

In this paper, I exploit this information, and study the diffusion of ideas generated inside companies, by tracing across time, the level and distribution of forward citations to their patents. Since my interest is in the diffusion of knowledge outside company boundaries, I consider only citations made to patents by entities other than the primary assignee.

I define the impact of VCs on the diffusion of knowledge as the percentage increase in citations to prior patents that is caused after VCs invest in the issuing companies. My identification strategy addresses three potential sources of bias in identifying the causal effect of VCs: trends in aggregate citations, patent heterogeneity, and non randomness in the timing of VC investment.

2.2 Statistical Models

In Table 3, Column (6), I present the first estimate of the percentage increase in citations to prior patents after companies are financed by a VC. I report the Incidence Rate Ratio (IRR) on forward citations to prior patents, defined as the ratio between annual average citations post VC-financing, and annual average citations pre VC-financing, which are reported in Column (2) and Column (1), respectively. The interpretation of the IRR is that the likelihood of a citation to a prior patent increases 63% after the company is financed by a VC.

The simple ratio of means is informative but is likely to be biased. One concern is that the likelihood of a forward citation to a prior patent is not constant throughout the sample. As the pace of patenting accelerates worldwide, the frequency of patent citations has increased (Lerner et al., 2011) resulting in a positive trend in forward citations. In addition, the probability of a forward citation to a patent is not constant throughout a

knowledge upon which patent B builds, and over which B cannot have a claim.

patent's life. Most citations occur before 6 years after application date (Hall et al., 2001). Hence, if VCs tend to invest in companies with relatively young patents, this patent life-cycle effect could lead to an upward bias in the reported IRR.

To address these concerns, I follow the innovation literature (Hall et al. , 2001), and define a set of matching patents to control for aggregate trends, and for patent life-cycle effects on citations. In detail, for every prior patent in my sample I determine all U.S. patents filed the same year and assigned to the same USPTO technology class that are not financed by a VC.[6][7] I calculate the average number of forward citations to matching patents every year since application date as follows,

$$b_t^s = \frac{Total\ Citations_t^s}{Number\ of\ Matching\ Patents} \tag{1}$$

where $Total\ Citations_t^s$ corresponds to the total number of citations received by all matching patents at time t. I use this average citation rate to control for aggregate changes in the likelihood of citations to prior patents, at the technology-class and application year level.

Table 3 illustrates this procedure. Columns (3) and (4) report average annual citations to matching patents before and after the VC financing event of the corresponding prior patent. The difference between these columns illustrates the aforementioned aggregate trend in citations. Column (8), reports the Relative Incidence Rate Ratio (RIRRs), defined as the ratio between the IRR of prior patents and the IRR of matching patents. The RIRR corresponds to my definition of the impact of VCs on knowledge diffusion, controlling for aggregate trends in citations. The interpretation of the RIRR of 1.33 is that the likelihood

[6] In unreported results I use the grant year to define the matching patents, and also, both the application and grant years. Results remain robust to these alternative definitions. However, following Hall et al (2001), I use application years to avoid noise from the review process in the Patent Office.

[7] Note that technology classes as defined by the USPTO are very narrowly defined, specially compared to traditional industry-classifications. To illustrate, at present the USPTO office has assigned more than 400 technology-classes.

of a citation to a prior patent increases by 33% after a VC finances the issuing company, relative to the aggregate increase in the likelihood of a citation to a matching patent.[8]

In Table 4 I turn to a regression analysis. Because annual citations to patents are non negative, and the mean is very close to zero, the ordinary least squares (OLS) estimate in the linear regression model is likely to be inconsistent (Cameron and Trivedi, 1998). This is particularly true when estimating percentage increases by using the logarithm of the annual citations in the estimation (King, 1989). To overcome this concern, I use a Poisson model, which is the standard model for count data and preserves and exploits the nonnegative and integer-valued aspect of the outcome.[9] I estimate a Poisson regression of annual forward citations to prior patents, Y^s_{pt}, on a dummy that equals one after the company is first financed by a VC, VC_{pt}. Since all patents issued by the same company are subject to company-specific shocks, I cluster standard errors at the company level. Finally, note that 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.

Table 4, Column (1), contains results from a pooled Poisson model. The coefficient for VC_{pt} in the first column shows that the likelihood of a citation to a prior patent significantly increases by 62.7% after the VC financing event of the issuing company. Note the similarity between the IRR in Column (6) of Table 3, and the estimated coefficient for VC_{pt} using the pooled Poisson model in Column (1) of Table 4. Similarly, note that the

[8] Column (7) shows that the significant increase in citations to prior patents following the financing event of the companies holds, even after subtracting the average trend in citations to matching patents.

[9] Another common model for count data is the Negative Binomial Model which is a generalization of the Poisson model that addresses overdispersion (variance larger than the mean) by including an additional error term to capture unobserved factors. In unreported analyses I replicate Columns (1) and (2) from Table 4 using this model. Results hold and are not statistically different across models.

estimated constant in Column (1) of Table 4, corresponds to average annual citations to prior patents pre VC-financing, reported in Column (1) of Table 3.

Table 4, Column (2), reports results from a pooled Poisson model that controls for the aggregate trends in citations, and patent life-cycle effects, by offsetting average citations to matching patents, b_t^s, in the estimation. This is implemented by including in the Poisson regression the logarithm of b_t^s with a coefficient fixed to one. The resulting estimates reflect the effect of the VC investment on citation counts to prior patents, relative to patents in the same technology-class and filed the same year. Note that by offsetting average citations to matching patents in the estimation, I remove any aggregate annual variation. In that sense,this technique is similar to including time-fixed effects in the estimation, though it is more stringent, because the time fixed effects are at the technology-class and application-year level. Another advantage of this technique, is that it solves the identification problem of cohort, age and period effects in the number of citations received by patents (Hall et. al. 2007; Lerner et al., 2011).

The interpretation of the coefficient for VC_{pt} in Column (2) of 1.346 is that citations to prior patents increase 34.6% after the issuing company is first financed by a VC, relative to patents in the same technology-class and application year. Note the correspondence between the RIRR of Column (8) in Table 3, and the Poisson results using the offset in Column (2) of Table 4. Similarly, note that the constant in Column (2) corresponds to the ratio between average annual citations to prior patents and average annual citations to matching patents, before the VC financing event (the ratio between Column (1) and Column (3) in Table 3).

The second source of bias in identifying the causal effect of VCs on citations to prior patents is unobserved patent heterogeneity. Most patents receive few annual citations if any, and only some patents are consistently cited throughout the sample. This unobserved patent heterogeneity can lead to an upward bias in the pooled Poisson estimate of Column

(2).

To address this concern, I turn to a patent fixed-effects regression that exploits only within-patent variation in citations after the company is financed by a VC. I estimate the following equation,

$$Y_{pt}^{s} = \exp\left(\ln(b_t^s) + \alpha_p + \beta VC_{pt}\right)\varepsilon_{pt}^{s} \tag{2}$$

where I include in the estimation a full set of patent fixed-effects, α_p, that account for cross-sectional unobserved heterogeneity and is consistent with my strategy of employing only within-patent variation. As in Column (2) of Table 4, I control for the time variation over the sample period by offsetting b_t^s in the estimation. Finally, ε_{pt}^{s} correspond to idiosyncratic multiplicative shocks.

I estimate (2) by conditional quasi-maximum likelihood (QMLE) based on the fixed-effects Poisson model developed by Hausman et al. (1984).[10] Note that although the QMLE approach is based on the fixed-effects Poisson model, in the estimation 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 (Wooldridge, 1997), and cluster standard errors at the company level. This is valid because consistency of the QMLE approach only requires correct specification of the conditional mean (equation (2)) (Wooldridge, 1999).

Column (3) of Table 4 reports the QMLE results. The coefficient of 1.189 for VC_{pt} shows that citations to the same prior patent, relative to patents in the same technology-class and applied for the same year, increase by 18.9% after the issuing company is first financed by a VC. As expected, the coefficient for VC_{pt} decreases,

[10] Hausman et al. (1984) condition the likelihood on $\sum Y_{pt}^{s}$, which is a sufficient statistic of α_p.

compared to estimates from the pooled Poisson regression in Column (2).

As robustness checks, in unreported results, I repeat the Poisson analyses reported in Table 4 for the sample restricted to states different from California, Massachusetts and Texas. I also restrict the sample to the pre - and post - dot com periods. The effect remains unchanged, and is not statistically different across sub-samples. I also examine heterogeneity of the results by patent age. I find that the increase is highest for patents younger than 5 years, but the effect is also positive and significant for patents between 5 and 10 years of age.

I also explore the dynamics in within-patent changes in citations to prior patents, relative to matching patents, around the VC financing event. I estimate a QMLE model where the independent variables are indicators for individual years relative to the year of the VC investment, and omit event year 0 from the estimation (to avoid multicollinearity with the α_p). Figure 1 plots the estimated IRRs (solid line) together with their 95% confidence interval and confidence intervals from this specification, after restricting the observations to a [-2,5] window around the financing event.

Figure 1 should be read as follows. For the years preceding the financing event, the estimated RIRR is not statistically different from the RIRR of event year 0. That is, in event year -1 and event year -2, the ratio between the likelihood of a citation to a prior patent and the likelihood of a citation to a matching patent, is the same as in event year 0. This is reflected in the estimated RIRRs for the indicator variables of the years preceding the VC financing event. Neither is statistically different from one (recall that to test the statistical significance of RIRRs one compares the estimates RIRR to one, not to zero). This result is reassuring, as it shows that the increase in the likelihood of a citation is not driven by a pre-existing trend in citations to prior patents, relative to matching patents.

In contrast, after the VC financing event, the ratio between the likelihood of a citation to a prior patent and a matching patent significantly increases relative to event year 0. This is reflected in the estimated RIRRs for the indicator variables of the years following the VC deal. The RIRRs are all higher than one, and significant, after event year 2.

There are at least two ways to interpret the temporal citation patterns, around the VC investment, illustrated in Figure 1. One interpretation is that because there is no pre-trend in the ratio between the probability of a citation to a prior patent and a random matching patent, then absent the VC investment, this ratio would have remained constant for the rest of the period. Accordingly, the differences in the ratio that emerge after the financing event of the companies and that persist in time, should be interpreted as the causal effect of VCs on citations to prior patents.

Alternatively, the divergence in the ratio of likelihoods that emerges post financing, reflects the skill of VCs in selecting within very narrow technology classes exactly those patents which will be highly cited in the future. Under this interpretation, some or all of the gap in citations to prior and matching patents would have existed even if there had been no VC investment, and cannot be attributed to a causal effect of VCs. The fact that the effect does not take place until some time after the VC investment (as reflected in the lack of significance for the estimate of event year 1), makes this interpretation less likely, as it implies VCs would have to anticipate citations far in advance. Nevertheless, potential non-randomness in the timing of VC investments constitutes a potential source of bias.

To address this third potential source of bias, and disentangle between the two interpretations of Figure 1, in the next section I implement an instrumental variables strategy.

2.2.1 Addressing non-random timing in the selection of companies by VCs

The main identification assumption of the QMLE approach is that VC_{pt} is strictly exogenous, that is, there is no correlation between the idiosyncratic shocks in citations to prior patents, and the timing in which VCs invest in the issuing companies, i.e., $E\left[VC_{pt}\varepsilon^s_{pt-k}\right] = 0, \forall k.$ [11] However, if VCs do not randomly pick the timing of their investments this condition is unlikely to hold. For example as argued above, if VCs invest in companies precisely when the companies' patents are in the technological domain that will receive a lot of citations in the future, the QMLE estimate of Column (3) in Table 4 could have an upward "timing-bias".

To address this concern, I start with equation (2) and now assume that VC_{pt} is endogenous. I use Wooldridge's quasi-differencing transformation to remove the fixed effects (see Wooldridge, 1997 and Windmeijer, 2000) which leads to the following moment conditions:

$$E\left[\frac{Y^s_{pt}}{\exp\left[x^s_{pt}-\mu_x\right]B} - \frac{Y^s_{pt+1}}{\exp\left[x^s_{pt+1}-\mu_x\right]B} \mid z_{pt}, ..., z_{p1}\right] = 0 \tag{3}$$

where $x^s_{pt} = VC_{pt} + \delta^s_t$, $\mu_x = (NT)^{-1}\sum\sum x^s_{pt}$, $B = [\beta\ 1]$, and z_{pt} is an exogenous determinant of the timing of VC investments.

As instrumental variable, z_{pt}, I use variations at the state and year level of the value of financial assets, deflated by PPI, that are held by local public pension funds in a company's home-state.[12] This instrument is in the same spirit as other papers in the

[11] There are three identification assumptions of the QMLE model. First, relative changes in citation rates over time for patents within the same technology-class and applied for the same year, are comparable. Second, the conditional mean is correctly specified in equation (2). Finally, VC_{pt} is strictly exogenous.

[12] Recall that specification (2) does not include time fixed effects because by offsetting in the estimation average annual citations to matching patents removes any aggregate year variation. Also, specification (2) does not include state-fixed effects, as those are absorbed by the patent fixed-effects. Hence, in order to control for time and state effects in the instrumental variable estimation, I use directly variations in pension

literature (Mollica and Zingales, 2007; Bernstein et al., 2011) and relies on the sensitivity of VC investments to the availability of capital from local public pension funds. The basic idea is that in states and periods where pension pools are larger, domestic VC firms are more likely to raise capital and invest it locally. Because the process of raising and beginning to deploy capital takes about 1 to 2 years, in the estimation I use a 1 year lag of the variations in the pension pool size.[13]

A valid instrument has to satisfy two restrictions. First, it must be correlated with the endogenous variable. To show that the size of state public pension funds is correlated to the timing in which local companies are selected by VCs, I run an OLS regression at the state-time level, where the dependent variable is the total value of investments in new companies by VC firms, and the explanatory variable is the size of local public pension funds, deflated by the PPI and lagged by 1 year. I include time-fixed effects and state-fixed effects in the estimation, and compute robust standard errors by clustering at the state level. Table 5 summarizes results and shows a positive and significant relation between the value of VC investments in new companies, and the size of local public pension funds. As a robustness check, in Column (2) I use as dependent variable the number of new investments and run a QMLE model. I also find a positive correlation between the total number of new investments in companies and the size of local public pension funds.

The second condition for a valid instrument is the exclusion restriction, which requires changes in public state pension funds to be conditionally independent from the unobserved time varying-heterogeneity in specification (2), i.e. $E\left[\varepsilon^s_{pt}|z_{pT},...,z_{p1},\alpha_p\right] = 1$. This condition cannot be empirically tested. However, since pension funds primarily change as a result of pension reforms, and because these reforms are normally driven by broader socioeconomic considerations rather than the innovative activity of the local VC industry, it is likely that the exclusion restriction is satisfied.

pools.

[13] In unreported results I use a lag of two years and results are quantitatively similar.

Nonetheless, one potential concern regarding the instrument is that variations in state pension funds' assets are somehow indicative of innovation opportunities within the state. If this is the case, then the pension size instrument and changes in relative citations to prior patents may be correlated via a state effect, and not exclusively through the VC financing channel. I address this potential violation of the exclusion restriction in two ways. First, I define relative citations at the state-level. Second, I exclude from the dependent variable citations originating in assignees that are located in the same state, and use only out-state citations. The exclusion restriction is unlikely to be violated in these settings because if the size of local and state pension funds is correlated to changes in innovation within a state, such a change should affect equally all patents issued in the same state, and is therefore unlikely to affect relative measures of citations at the state level, and particularly, state-level relative measures of out-state citations .

Econometric Considerations Three considerations are worthy of note. First, because the source of variation I exploit in my GMM-IVs estimation is at the state level, I cluster standard errors at the state level.

Second, in estimating a fixed-effects Poisson model using an instrumental strategy by exploiting the moment conditions (3), there is no "first stage" as in standard linear Two-Stage Least Squares (2SLS) instrumental variable techniques. In Tables (6) - (8) I document the analogue first stage of a 2SLS methodology for the sake of exposition, and to use the F test as suggestive evidence that the instrument is not weak. This is so, because I am unaware of a test for weak identification in non linear GMM models.

Finally, note that the sample of the GMM-IVs approach is different from the original analysis sample on two accounts . First, recall that the data on the size of the assets of public state pension funds is only available for the 1993-2008 period (see Table 2 for summary statistics on this restricted sample). Second, observations for the last period are eliminated from my sample because the moment conditions used in the GMM approach differentiate out the fixed effects. Hence, to facilitate comparisons between the QMLE and GMM-IVs estimates, I replicate the QMLE results for the GMM-IVs sub-sample and report them in the tables below.

GMM-IVs main results Table 6 shows that results from Table 4 are robust to controlling for non random timing in the investment of VCs. Column (1) presents the analogue first-stage of a linear instrumental variables approach, where I regress the endogenous variable, VC_{pt}, on the one-year lag of variations in the size of state pension funds in the home-state of the issuing company, and on patent fixed-effects. The F -statistic of 328.30 suggests that the instrument is unlikely to be weak (Stock and Yogo, 2005).[14]

To facilitate the comparison between the QMLE and GMM-IVs estimates, in Column (2) of Table 6, I provide the QMLE estimates using the restricted sample of the GMM-IVs approach. The coefficient remains positive and statistically significant for the restricted sample, and is not significantly different from the QMLE estimate using the entire sample in Table 4.

Column (3) presents the analogue reduced form results of a linear instrumental variables approach, obtained by substituting the endogenous variable with the instrument. As expected, there is a positive correlation between variations in state pension funds' assets and relative citations to prior patents.

Finally, the GMM-IVs estimate of β is 1.805, and is reported in Column (4). Relative to the QMLE estimator, the estimated effect increases from 21.4% to 80.5% after accounting for non-random selection by VCs. However, the difference between the two estimated effects is not statistically significant.

[14]The F statistic of the regression of the endogenous variable on the instrument and the rest of covariates is the standard test for weak instruments in linear IV. I report this test as suggestive that the instrument is not weak, as I am unaware of a test for weak identification in non linear GMM.

In unreported results I exclude California from the sample and results are quantitatively similar. I also ran a linear model with relative citations (i.e. ratio between citations of prior patents and citations of matching patents) as dependent variable, and VC_{pt} as main explanatory variable, by OLS (analogue to QMLE), and variations in the size of state pension funds in the home-state of the issuing company as instrument, by two-stage least squares (analogue to GMM-IVs). Results are quantitatively similar. Finally, as an additional robustness check, I also use as instrument variations at the state level of the value of financial assets held by local pension funds and normalized by state GDP. Results are also quantitatively similar.

Robustness Checks Tables 7 and 8 summarize results from the robustness checks used to address concerns of potential violations of the exclusion restriction. Table 7 replicates the GMM-IVs approach using relative citations at the state-level. Table 7 shows that results continue to hold and are quantitatively similar to Table 6. However, note that because relative citations in Table 7 are at the state-level, the interpretation of the estimated coefficients changes. To illustrate, the coefficient of Column (4) in Table 7 is interpreted as follows. After companies are financed by a VC for the first time, the likelihood of a citation to the same prior patent causally increases by 83.7%, relative to other patents in the same technology-class, filed the same year, and issued in the same state. In unreported results I exclude California from the specifications, results remain qualitatively similar.

Table 8 replicates the GMM-IVs approach using relative citations at the state level, and, using as dependent variable out-state citations, results continue to hold and are quantitatively similar to results reported in Table 6 and Table 7. However, note again that the interpretation of the coefficients changes. To illustrate, the coefficient of Column (4) in Table 8 is interpreted as follows. After companies are financed by a VC for the first time, the likelihood of an out-state citation causally increases by 85.1%, relative to other patents in the same technology-class, filed the same year and issued in the same state. In unreported results I exclude California from the specifications, results remain qualitatively similar.

In conclusion, the effect of VC financing on relative citations is significant for all GMM-IVs specifications. [15]

2.3 Interpretation of results

An interest finding that emerges from Tables 6, 7 and 8, is that the GMM-IVs estimates of the effect of VCs on citations to their targets' patents, exceed the corresponding QMLE estimates (although the difference is only significant for some specifications). If one assumes on a priori grounds that the QMLE approach leads to upward-biased estimates of the true causal effect of VCs, the even larger GMM-IVs estimates I document present something of a puzzle. Interestingly, this puzzling result is common to all other papers in the VC literature that instrument VC investments using shocks to the availability of funds for VCs (Kortum and Lerner, 1998; Mollica and Zingales, 2007; Bernstein et al.;. 2011).[16]

A potential explanation behind this result, is that there is underlying heterogeneity in the effect of VCs on the diffusion of knowledge, and that my GMM-IVs estimates are recovering the effect for a particularly sensitive group of patents. Since the GMM-IVs approach estimates the effect of VC investment only on those patents that respond to the instrument, if shocks to the capital available for VCs trigger investments on a sub-population of patents with a relatively high marginal return from VC selection, it is clear that the GMM-IVs estimates can exceed the QMLE results. As long as the main

[15] 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 QMLE to be estimated. Consequently, comparisons across models do not have a straightforward interpretation.

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

reason why these projects have a low chance of being selected in normal times is because of higher than average costs to investing, rather than because of lower than average returns to VC financing, then a logic similar to the "local average treatment effect" in the linear literature (Imbens and Angrist, 1994), suggests that a GMM-IVs approach based on shocks to the availability of capital for VCs will yield estimates above the marginal effect in the population of patents, and potentially above QMLE estimates.

There are at least two reasons why patents that are selected when there are relatively more funds available for VCs may be more sensitive to the VC investment. The first is that shocks to the supply of funding capital for VCs are most likely to affect selection of lower quality projects who otherwise would not have been selected. This can occur if VCs tend to go down the "quality ladder "and invest in "money chasing deals" (Gompers and Lerner, 2000) when there is excess supply of capital, or if shocks to funds are more likely to affect capital-constrained VCs, which may only have access to lower quality projects (Sørensen, 2007).

A second possibility is that an abundance of capital allows investors to experiment more effectively, shifting the type of startups that investors finance towards those that are more risky and innovative, rather than just worse. This second interpretation is supported by Nanda and Rhodes-Kropf (2011) who show that in "hot markets" VCs fund risky and innovative startups, rather than just worse projects. The authors argue that excess availability of capital for VCs facilitates the experimentation that is needed for the commercialization and diffusion of radical new technologies.

To provide suggestive evidence that VCs experiment when there is excess availability of funds, I compare the average novelty of patents funded in hot versus cold markets. A company is defined to have been financed in a hot (cold) market, if the variation in local public pension funds in the home-state of the company during the year of the VC deal is within the top (bottom) 25 percent of the sample. As a proxy for novelty, for each prior patent I construct the "originality" measure of Hall et al. 2001. This measure is constructed as one minus the Herfindahl index of the cited patents across technological classifications. Thus a higher measure of originality means that the patent is drawing on more diverse array of awards. The intuition for this measure is the idea that patents that cite awards in a broader array of technology classes are likely to be more novel as they draw from a more diverse array of awards to create something new, as opposed to patents that draw knowledge from few technology classes and are likely to constitute marginal improvements.

Table 9 compares the average originality for prior patents funded in hot versus cold markets. The first row shows that prior patents funded in hot markets have higher originality measures than those funded in cold markets. The difference is statistically significant, even after controlling for the average originality of matching patents as shown in row 2. Thus, Table 9 provides suggestive evidence that VCs invest in patents that are ex-ante more novel when there is excess availability of funds as measured by the size of state public pension funds. This result is line with Nanda and Rhodes-Kropf (2011) and also with Hirukawa and Ueda (2011), who find that high availability of funding capital is not necessarily invested to support less innovative businesses.

For policy evaluation purposes, the concern that remains regards external validity. In particular, if the estimated effect using the GMM-IVs approach is not the "average" effect on a randomly picked patent, the question of how these results can inform policy remains. Note however, that to inform policy the marginal return to VC financing in the population may be less relevant than the average return for the group who will be impacted by a proposed reform. In such cases the best available evidence may be IV estimates of the effect of VC financing based on similar reforms. In other words, in order for a study to have external validity it must be relevant for the populations the treatment may be extended to (Imbens, 2009). And since an important part of growth policies seek to stimulate VC financing via shocks to the available capital for VCs, my paper is informative for current policy.

In conclusion, the pattern in Tables 6-8 suggest that the QMLE estimate, even if upward biased as an estimate of the average causal effect of VC financing, may be a relatively conservative estimate of the causal effect for groups typically affected by supply side policies that stimulate VC activity. Under that interpretation, a conservative back of the envelope calculation of my findings is that roughly 5% of patent production can be attributed to VCs facilitating the diffusion of their targets' patented knowledge. Average annual citations to prior patents before VC financing are 0.64. VC financing increases annual citations by roughly 20% (using the QMLE estimate). Assuming a patent life of 10 years, this implies that each VC-backed patent receives 1.3 extra citations because of increased diffusion following the VC financing event. Since 4% of patents have been assigned to VC-backed companies, this implies that 5% of patents in the U.S. can be attributed to VCs facilitating the diffusion of knowledge.[17] Including this effect, the estimated share of patents ascribed to VCs increases from 4% to 9%, which is closer to the macro level estimate of Kortum and Lerner (2000).

2.4 Knowledge Diffusion and Patent Citations

In this section I discuss potential concerns of using citations counts as a proxy for knowledge diffusion.

An extensive literature on the economics of technological change has demonstrated that patent citations are a reasonable measure of the transfer of knowledge between two parties. Although citations are not a perfect measure of knowledge flows, for example, many are added by patent examiners rather than by the inventors themselves, prior research finds

[17] 4%*1.3=5%

they correlate well with actual knowledge flow (Jaffe et al., 2002; Duguet and MacGarvie, 2005; Roach and Cohen, 2010). Thus, the consensus in the literature is that citations are informative of links between patented innovations, and can be interpreted as “paper trail” evidence of knowledge diffusion

Nevertheless, it might be thought that inclusion of citations from reviewers may bias my estimates. The idea is that since patent examiners are to make sure all relevant prior art is cited, even if the inventor was unaware of it, there may be citations where there is no diffusion of knowledge (Sampat, 2010; Lemley and Sampat, 2010). However, note that since my empirical strategy is based on relative measures of citations, this concern is minimized. There is no reason why patent examiners would include patents issued by VC-backed companies more often in their examinations, than any other patent in the same technology-class and application year. However, as a robustness check, in unreported results I use information about the source of the citation, which is available starting on 2001, and exclude citations of patent reviewers from the analysis. Results remain qualitatively similar, although the resulting sample is very small and I don’t have sufficient statistical power.

A more nuanced view is that only citations to prior patents, relative to matching patents, may increase post-financing, without reflecting any difference in diffusion patterns between these types of patents. For instance, potential targets may strategically use patent citations to attract potential investors. In this scenario citations to prior patents may increase relative to matching patents, without any real effect on the relative diffusion of prior patents. To address this concern, in unreported results 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. The estimate of VC_{pt} is close to one and is not statistically significant, which minimizes concerns that the effect is due to the use of patent citations to attract a VC's attention, as opposed to real knowledge diffusion.

Finally, another nuanced view is that the increase in citations is due to "litigation fear". This concern is however unfounded to the extent that citations represent no protection against patent infringement law suits. Patent infringement cases are fought even if a formal citation to the patent supposedly infringed is included in the patent being sued, which minimizes concerns that the increase in citations stems from litigation fear as opposed to knowledge diffusion (For more on this topic see the Supreme Court Ruling of Microsoft Corp. v.s. I4I Limited Partnership, 2010).

3 Disentangling mechanisms

Having shown that citations to prior patents causally increase after a VC invests in the issuing companies, I turn to disentangling the mechanisms behind this effect. I focus on three market mediated mechanisms of knowledge diffusion: exchange of ideas through company networks, mobility of individual inventors, and Intellectual Property (IP) trade.

3.1 Knowledge Diffusion and VC-backed company networks

One of the mechanisms for knowledge diffusion that has been studied in the innovation literature is social networks. The idea is that for technologies characterized by high levels of tacitness and complexity, and that cannot be completely codified into blueprints, repeated face-to-face contact and personal interactions are useful for knowledge transfer (Audretsch and Feldman, 1996; Dahl and Pederson, 2005).

VCs can affect their new targets' networks by providing a platform for interaction with other companies in their portfolios. These interactions among companies financed by the same VC can serve as conduits for information exchange about technological developments and emerging market opportunities (Saxenian, 1994; Stuart and Sorenson, 2003). For example, VCs often organize summits where executives of their targets informally meet, and which can lead to future transfer of tacit technical knowledge across organizational boundaries. Also, by actively participating in their company boards, VCs can detect technological complementarities among companies in their portfolio, and encourage their targets to exploit them. Consistent with this idea, Lindsey (2008) shows that research alliances are more frequent among companies that share a common VC.

In this section, I explore whether the increase in citations to prior patents after companies are financed by a VC, is consistent with VCs facilitating the diffusion of knowledge across companies in their investment portfolios.

3.1.1 Data

The patent database provides the unique opportunity to study how the knowledge from the same patent diffuses to different types of agents, by exploiting information about the assignees of the citing patents. In this section I employ information about the VC investors of the citing assignees to classify citations into two types as follows:

1. Portfolio-linked (P)- If the citing assignee and the cited assignee share a common VC at the time of the citation
2. Non Portfolio-linked (NP)- otherwise.

Table 10, Panel A, shows the average size of the VC portfolios for companies with prior patents. Panel B, shows the distribution of the portfolio size by year in which the companies are first financed by a VC. On average, VC-backed companies with prior patents join portfolios with 16.6 other VC-backed companies.

Table 10, Panel C, presents average annual citations to prior patents by the issuing company's home-state, classified by type of citation. Because the size of VC portfolios is small relative to the universe of potential citers, average annual non portfolio- linked citations exceed average annual portfolio- linked citations.

3.1.2 Analysis

I start by reporting the percentage increase in citations to prior patents, after companies are financed by a VC, by type of citation. Table 11 summarizes results. Column (1) and (2), report annual average citations to prior patents by type of citation, before and after, a VC invests in the issuing company. The first row reports the analysis for portfolio-linked citations, and the second row, for non portfolio-linked citations.

Analogue to Table 3, Column (6) in Table 11 reports the percentage increase for both types of citations, as summarized by the IRRs. Column (7) shows that the difference between the IRR for portfolio- and non portfolio- linked citations is positive and significant. This means that after companies are financed by a VC, the increase in the likelihood of a portfolio- linked citation, is larger than the increase in the likelihood of a non portfolio- linked citation.

To control for aggregate trends in citations and the evolution of the VC industry, I construct average annual citations to matching patents also by type of citation. In detail, for every prior patent and every time period t, I calculate the average number of citations received by matching patents by type of citation C as:

$$b^s_{tC} = \frac{Total\ Citations\ of\ type\ C_t}{Number\ of\ Matching\ Patents} \tag{4}$$

where the numerator is the total number of citations of type C received by all matching patents at time t and $C \in \{NP, P\}$. To classify citations to matching patents as portfolio-linked, I use information on the companies that are financed by the same VC as the company of the prior patent. For example, b^{s}_{tP} corresponds to the logarithm of the average number of citations to matching patents at t from the members of the portfolios of the VCs that back the companies of the prior patents. A comparison of columns (3) and (4) shows that citations to matching patents from both, portfolio- and non portfolio-connections, increase after the VC investment.

Column (9) reports the estimated percentage increase in citations to prior patents after the VC investment, controlling for aggregate trends in citations by type of citation. The column reports RIRRs (i.e. the ratio between the IRR for prior patents to the IRR for matching patents), and shows that percentage increase in citations for both types of citations is statistically significant. Column (10) reports the difference in RIRRs by type of citation, and shows that the likelihood of a portfolio- linked citation increases significantly more than the likelihood of a non portfolio- linked citation, even after controlling for the aggregate increase in citations at the technology-class, application-year and type of citation level.

To control for potential overdispersion, clustering of observations, biases from time varying heterogeneity at the patent-and-type of citation level, and from aggregate changes in citations, I estimate Poisson models where I allow the VC investment to affect differently both types of citations. I estimate different versions of equation (5) below, by increasingly including regressors to address the different types of aforementioned concerns. Table 12 summarizes results, and shows that the significant increase for both portfolio- and non portfolio-linked citations is robust to these concerns. The concentration inside VC portfolios is apparent in all models, but is less robust across specifications.

The general model is

$$E\left[Y_{pCt}^{s}\right]=\exp\left(\sum_{C}\alpha_{pC}+\ln(b_{tC}^{s})+\sum_{C}\gamma_{C}D_{C}+\sum_{C}\beta_{C}VC_{pt}*D_{C}+\varepsilon_{pCt}\right) \tag{5}$$

where Y_{pCt} are forward 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 that equals one when $C = NP$. D_P is defined analogously. α_{pC} are patent-type of citation fixed-effects that absorb the time-invariant heterogeneity at the patent-type of citation level, and b_{tC}^{s} corresponds to average citations of type C, at time t, to patents in the same technology-class and with the same application-year as patent p. By including the different types of average citations in the estimation with a coefficient fixed to one, I control for aggregate changes over time in the likelihood of forward citations at the technology-class and application-year level by type of citation. This technique is similar to including type of citation- cross- time fixed- effects, since it removes any aggregate annual variation by type of citation. Finally, 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 zero that captures idiosyncratic multiplicative shocks at the patent-type of citation level.

The coefficients of interest are the β_C's which can be interpreted as the percentage increase in citations of type C after the VC investment. Standard errors are clustered at the patent level for columns (1) through (4) and at the state level for columns (5) and (6). Panel A of Table 12 presents the estimated coefficients. Panel B tests whether the β_C's are statistically different, using a chi-squared test.

Column (1), summarizes results from estimating equation (5) using a pooled Poisson model, excluding α_{pC} and δ_{tC}^{s} from the estimation. The coefficients of 0.635 and 0.002 for D_{NP} and D_P respectively, represent average portfolio- and non portfolio- linked citations to prior patents before the VC investment. Note the correspondence of these numbers with the averages reported in Column (1) of Table 10.

The interpretation of the estimate of 1.620 for $VC_{pt} * D_{NP}$ is that the likelihood of a non portfolio- linked citation to a prior patent increases by 62.0% after a VC finances the issuing company. In contrast, the coefficient of 4.052 for $VC_{pt} * D_P$ means that portfolio-linked citations increase 305.2% after the target company joins the portfolio, a much higher increase than for non portfolio-linked citations. Panel B confirms that the estimated IRR for portfolio- linked citations (the coefficient of $VC_{pt} * D_P$) is significantly larger than the estimated IRR for non portfolio-linked citations (the coefficient of $VC_{pt} * D_{NP}$). Finally, note the close correspondence between the estimates for $VC_{pt} * D_P$ and $VC_{pt} * D_{NP}$ in Column (1) of Table 12, and the estimated IRRs in the first and second rows of Column (6) in Table 11, respectively.

Column (2) reports the coefficient estimates of equation (5) using a pooled Poisson model, and offsetting b^s_{tC} in the estimation. The coefficient of 1.176 for D_{NP} corresponds to average non portfolio- linked citations to prior patents, relative to matching patents, before the VC financing event (roughly equal to the ratio between average citations pre-financing to prior patents, and average citations pre-financing to matching patents in Table 11).[18] Compared to Column (1) the coefficient decreases after controlling for aggregate changes in non portfolio- linked citations at the technology-class and application-year level. Similarly, the coefficient of 2.437 for $VC_{pt} * D_P$ means that the likelihood of a portfolio- linked citation to a prior patent increases by 143.7%, relative to matching patents. Note that the estimate is no longer statistically significant, although it is much higher than 1, which could be due to the loss of observations in estimating this model relative to the model in Column (1). This is because annual average portfolio- linked citations to matching patents are often zero, and consequently those observations are dropped from the estimation.[19]

[18] The coefficient of 0.868 for D_P corresponds to the ratio between average citations to prior patents pre financing, to average citations to matching patents pre financing, restricting the sample to observations for which $\ln(b^s_{tP})$ is not missing. For that sample the ratio is 0.868=0.01/0.012.

[19] This is because by osseting b^s_{tC} in the estimation, the coefficient of $\ln(b^s_{tC})$ is set to one, and since the

Finally, note the close correspondence between the estimates for $VC_{pt} * D_{NP}$ and $VC_{pt} * D_P$ in Column (2) of Table 12, and the estimated RIRRs in the first and second rows of Column (9) in Table 11, respectively.

Column (3) (Column (4)) presents results from a QMLE estimation of equation (5) with the α_{pC}s and excluding (including) δ^s_{tC} in the estimation. Interpretation of the estimated effect now excludes patent-type of citation heterogeneity (i.e. some patents are more compatible to the type of research made by other companies in the VCs' portfolios and may be driving at least some of the effect in Columns (1) and (2)). The coefficients of 1.464 and 2.898 in Column (3), for $VC_{pt} * D_{NP}$ and $VC_{pt} * D_P$ respectively, mean that the likelihood of a non portfolio- and a portfolio- linked citation to the same prior patent, increase by 46.4% and 189.8%, respectively. The coefficients of 1.186 and 2.785 in Column (4), for $VC_{pt} * D_{NP}$ and $VC_{pt} * D_P$ respectively, mean that the likelihood of a non portfolio- and a portfolio-linked citation to the same prior patent, increase by 21.5% and 178.5%, relative to matching patents, respectively. Panel B shows that for both specifications using patent-type of citation fixed-effects, the estimated percentage increase in portfolio-linked citations is statistically larger than for non portfolio-citations.

Similar to the previous section, the main identification assumption behind Columns (1) through (4) in Table 12, is that ε_{pCt} is strictly exogenous. I relax this assumption and address non randomness in the timing of investment by VCs, using the GMM-IVs approach in Column (6). As instruments for $VC_{pt} * D_{NP}$ and $VC_{pt} * D_P$, I use the interaction between the two-year lag in variations in the assets of public pension funds in the home-state of the company, and the type -of -citation dummies.[20] The estimated RIRRs using the GMM-IVs approach are positive, although they are less precise. Plus, the estimated percentage increase in citations is estimated to be larger inside, than outside, VC

logarithm of 0 is undefined, the observations where $b^s_{tC} = 0$ are effectively dropped.

[20] I use as instrument the two-year lag of variations in the size of local and state pension funds as the approach using the one-year lag couldn't find an improvement by the 100th iteration.

portfolios. However, as reported in Panel B, the difference in the estimates is no longer significant. Note however, that the statistical power in this estimation is significantly reduced for two reasons. First, the sub-sample that is available to estimate the GMM-IVs approach is much smaller. This is also consistent with the lack of significance in the difference in RIRRs by type of citation in Panel B Column (5), which reports QMLE estimates for the restricted GMM sample. In addition, in order to identify the differential increase by type of citations using the GMM-IVs approach, ideally I would have an additional instrumental variable for type of citation. Since I don't have an instrument for the type of citation, in the estimation I use the interactions between variations in the pension fund size and the type- of- citation dummies as substitute instruments, which also decreases statistical power.

Consistent with the previous section, the GMM-IVs estimate for the percentage increase in non portfolio- linked citations, exceeds the estimate using the QMLE approach. As mentioned before, one possible explanation behind this result is that there is underlying heterogeneity in the effect of VCs, and the GMM-IVs approach recovers the effect for a particularly sensitive sub-group of patents. The higher availability of capital for VCs not only affects the timing of VC investments, but also, the types of projects that are financed by VCs. With a higher availability of capital, VCs can experiment and invest in more novel projects. Those types of projects are likely to be more sensitive to the VC financing event, and see a higher impact on their diffusion following the VC investment.

Interestingly, in contrast to non- portfolio linked citations, the QMLE estimate for the percentage increase in portfolio- linked citations, exceeds the estimate using the GMM-IVs approach. This result is consistent with a model in which VCs select targets not only based on their individual characteristics, but according to whether they are complements to their existing investments, which biases upward the QMLE estimate. My finding suggests that the upward bias in the QMLE estimate due to selection on technological complementarity of potential targets and existing portfolios, surpasses the downward bias stemming from changes in the investment strategy of VCs during hot markets.

3.1.3 Interpretation

Overall, this section has two main findings. First, consistent with the active investor effect of VCs, I find suggestive evidence that the presence of a common VC as an investor facilitates the transfer of knowledge across companies. The difference in the effect of VCs on portfolio- linked and non portfolio- linked citations is positive across all specifications, and significant for most of them. Note that the GMM-IVs methodology controls for non random timing in which companies are financed by a VC, but cannot address non randomness in which VC selects the company. That is, I cannot claim that the effect I report is not at least partially driven by VCs selecting companies which are technological complements of their existing portfolios. And not do I seek to claim that. If part of the value-added role of VCs in knowledge diffusion is in selecting companies that have technological complementarities, this is an interesting result, that has implications for innovation policy and for the theory of VC investment. One potential mechanism behind this finding is the prevalence of inter-company alliances inside VC portfolios. Lindsey (2008) shows that companies that share a common VC are more likely to enter in a research alliance, and Gomes-Casseres et. al. (2006) show that inter-firm alliances are a mechanism for sharing technological knowledge. In the next section I explore an alternative channel: inventor turnover inside VC portfolios.

Second, consistent with VCs certifying the commercial value of a company's IP, I find that the increase in citations is not limited to within VC portfolios. I show a causal increase in non portfolio- linked citations after the VC investment.

3.1.4 Robustness checks and extensions

In this section I summarize results of mostly unreported analyses that test the robustness of the findings and explore alternative explanations.

I start by considering more carefully non portfolio- linked citations. An alternative reason why non portfolio- linked citations may increase post VC financing, is that they are concentrated in companies inside the VC industry, which would not be consistent with a generalized effect of VCs certifying the commercial value of their target's knowledge. To test this hypothesis I further classify non portfolio- linked citations into two types: Non VC-backed, which includes citations from private and public companies outside the VC industry as well as citations from universities and independent inventors. And VC- backed, which includes citations from companies that are also VC- backed but do not share a common VC investor with the target. I then test whether the increase in citations to prior patents outside VC portfolios is limited within the VC industry, or whether it diffuses more generally in the economy.

Table 11 presents preliminary evidence that citations from Non VC-backed assignees significantly increase post financing. Table 13 presents a formal analysis using pooled Poisson and QMLE regressions allowing the effect of VC investment to affect the three types of citations differently. The last two columns of Table 13 present QMLE and GMM-IVs estimates of the effect of VCs on citations from Non VC-backed companies, using the specification at the patent-year level and the dependent variable restricted to Non VC-backed citations. The effect of VC financing on Non VC-backed citations to prior patents is larger than one and statistically significant, across all specifications. Table 13 is consistent with VCs certifying the value of their targets' IP inside and outside the VC industry.

Second, I examine whether the positive effect on non portfolio- linked citations inside the VC industry masks the evolution of VC portfolios. Recall that a citation is defined as portfolio- linked if the citing assignee shares a common VC investor with the target when the target is selected by a VC. In other words, I do not include citations from future members of the portfolio, as for those, VC_{pt} is not properly defined. To address this issue, in unreported estimations I classify citations from new arrivals to the VC portfolio as portfolio- linked citations. The result that citations inside the VC industry, and outside portfolios, increase is robust to this change.

Third, syndication networks among VC firms have been shown to matter for VC fund performance (Hochberg et al. 2007). One natural question is whether they also matter for knowledge diffusion. In unreported results, I reclassify non portfolio citations from VC-backed companies into two groups: 1. Syndication- linked: if the citing company is backed by at least one VC with whom at least one of the investors of the cited company has syndicated an investment in the past. 2: Non related VC-backed, otherwise. I then test whether citations increase inside syndication networks. I don't find evidence that information is diffused within VC syndication networks.

3.2 Knowledge Diffusion and Labor Mobility

Building on Arrow's (1962) seminal work on the link between labor mobility and knowledge diffusion, many papers in the innovation literature have explored whether the technological know-how that is embedded in scientists' human capital is exploited by new employers when inventors switch jobs (Almeida and Kogut,1999; Kim and Marschke, 2005; Agrawal and Singh, 2011; Azoulay, Graff Zivin and Sampat, 2012). For instance, Almeida and Kogut (1999) show that the patents that semiconductor companies cite in their applications reflect the employment histories of their engineers, suggesting that ideas in the semiconductor industry are spread by the movement of key engineers among companies, especially within a geographical region.

There are at least three ways in which VCs can affect mobility of inventors. First, through their exit. Second, by changing the incentives faced by inventors from creative freedom to commercial focus. Finally, by encouraging inventor turnover inside their portfolios. In this section I examine each of these channels.

3.2.1 VC exit

One way through which VCs can affect the location of innovators is by exiting their investments. VCs specialize in taking temporal capital positions in companies, and exit their investments in order to close their funds. After a VC exits its investment, inventors can be forced to move. For instance, if VCs exit an investment in a company through an acquisition, the acquiring company absorbs the inventor team of the target and in the process all innovator-embedded knowledge is transferred. If the VC exits the company through an IPO, innovators may be more likely to leave the target (Bernstein, 2012). Finally, if a VC exits a company through a forced liquidation, inventors are forced to seek new jobs, presumably as inventors in other entities, which can also result in the diffusion of the target's knowledge.

To test whether the increase in citations is due to the mobility of innovators following the VC exit, I restrict the sample to a window around the financing event for which I am certain that the VC has not exited the investment. This is implemented by dropping observations that occur after the last known investment of the VC in the target. In unreported results, I show that results are robust to this restriction, and therefore, that the increase in citations to prior patents is not entirely due to the effect of VC exit on mobility of inventors.

3.2.2 Creative Freedom and Commercial Focus

Another way through which the presence of a VC as an investor in a company can induce inventor mobility, is if the VC's arrival implies a transition from creative freedom to commercial focus (a la Aghion et al., 2008). In this setting, inventors that value creative freedom may chose to leave the company towards entities that offer more creative flexibility. Differences of opinion can also result from leadership changes in a company, which can affect mobility choices by inventors. For instance, Brittain and Freeman (1986) find that semiconductor firms that hire a CEO from outside the firm or are acquired by non-semiconductor firms have higher spin-off rates.

To study whether the increase in citations to prior patents can be traced to mobility of inventors (while the VC remains as an investor in the company), I start by identifying which citations that occur post-financing can be associated to an inventor from the original target. This is implemented by inferring inventor mobility using information on inventors from the patent applications, and analyzing changes in assignees through time. The analysis of inventor mobility is facilitated by the HBS data-set, which 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 mobility of individual inventors in my sample. Overall, I have information of 11,627 inventors that work at the companies that issued the prior patents, and their subsequent inventions in the same company or other assignees. Using this information, I classify a forward citation to a prior patent as "inventor- linked", if at least one inventor assigns a patent to the cited company before the VC deal, and also assigns a patent to the citing assignee after the VC deal; and as "not inventor- linked" otherwise.

Table 14 reports summary statistics on inventor- linked and non inventor- linked citations. Panel A, Columns (1) and (2) report average annual citations to prior patents from all citing assignees, before and after the VC investment. Column (5) shows that even after restricting citations to not inventor- linked, the IRR is positive and significant. Columns (3) and (4) report annual average annual citations to matching patents from all citing assignees, before and after the VC investment of the prior patent. Column (6) shows that even after controlling for aggregate trends in inventor mobility, the percentage increase in not- inventor linked citations to prior patents is positive and significant.

In unreported results, I extend the analysis of Panel A, and I formally test whether the increase in citations can be attributed to inventor mobility by excluding inventor- linked citations from the dependent variable and replicating the analysis using the Poisson regressions of Section 2. Results continue to hold, which implies that the effect of VCs on knowledge cannot be entirely explained by inventors leaving the target companies after the VC deal, and directly transferring information to their new work places.

3.2.3 Inventor Turnover within VC portfolios

Table 14, Panel B, reports summary statistics on portfolio-citations classifying them into inventor- and not inventor-linked. Columns (1) and (2) report average annual portfolio- linked citations to prior patents from all citing assignees, before and after the VC investment. Column (5) shows that even after restricting portfolio- linked citations to those that are not inventor linked, the IRR is positive and significant. Columns (3) and (4) report annual average annual portfolio- linked citations to matching patents, before and after the VC investment of the prior patent. Column (6) shows that even after controlling for aggregate trends in inventor mobility, the percentage increase in citations within VC portfolios (and that are not inventor linked) is positive and significant, as reflected in the RIRR. Panel C, reports similar results for citations outside the VC portfolio.

In Table 15 I extend the analysis of Panels B and C in Table 14, and formally test whether the concentrated increase in citations inside VC portfolios can be attributed to inventor mobility across companies financed by the same VC. To that end I repeat the Poisson analysis from Section 3.1. but using as dependent variable not inventor- linked citations. There are two main results from this table. First, even after excluding inventor-linked citations, non portfolio- linked citations significantly increase. This is true even after controlling for non random timing in VC investments using the GMM-IVs procedure.

The evidence for portfolio- linked citations is less clear. For most specifications citations from portfolio- linked citations increase, and the increase is significantly larger that the increase in citations from non portfolio- linked citations. However, once I address concerns of non random timing in VC selection, the increase in citations inside VC portfolios is no longer significant, and the estimated coefficient is significantly smaller than the estimated increase in non portfolio- linked citations (see Column (6) in Table (15)). I interpret these findings as suggestive that knowledge transfer inside VC portfolios is associated to inventor turnover across companies financed by the same VCs.

Discussion

There are two main findings from the analysis of inventor-linked and not inventor-linked citations. First, consistent with VCs having a certification effect on the quality of their targets' IP, the increase in citations to prior patents outside VC portfolios cannot be entirely attributed to inventors leaving the targets and diffusing their companies' ideas. Second, consistent with the active investor role for VCs, there is suggestive evidence that the increase in citations inside VC portfolios is associated to inventor mobility across companies financed by the same VC.

One drawback from measuring the mobility of workers using data on patent assignments is that not all moves are observable. First, I only record movements of inventors. However, other types of workers can also change jobs, and facilitate the diffusion of a company's knowledge. Second, even if I focus on inventor mobility, my data is still necessarily incomplete. To see this, note that I can identify the movement of an inventor only if the inventor decides to invent in the new workplace. Some inventors may change jobs and join new companies in executive positions in which they no longer apply for patents, but can still influence the innovation efforts of the company. Thus, what my results imply is that the effect of VCs on citations to prior patents is not fully explained by the mobility of inventors that is observable in my data.

3.3 Knowledge Diffusion and Trade of Intellectual Property

The final mechanism I explore through which VCs can affect the diffusion of knowledge is patent trade. The sale of a patent constitutes the transfer of the right to exploit an invention, an can affect its diffusion outside the issuing company.

There is ample evidence that VCs help their companies become more professional. For example, Hellman and Puri (2000) find that having a VC as an investor is associated with a significant reduction in the time to bring a product to market. Similarly, Hellman and Puri (2002) find that VC-backed companies are more likely, and quicker, to professionalize in adopting stock option plans, hiring a vice president of sales, and by bringing in CEOs from outside the firm. One other way in which VCs may professionalize their targets is by encouraging them to optimize their patent portfolios. Many patents held by companies are never exploited by the producers, and if sold, could constitute a good source of revenue. Consistent with a correlation between VCs and intellectual property trade, Katila and Shane (2005) find that patents are more likely to be licensed if the issuing companies are in industries with high VC investment.

To test this mechanism, I use data on patent reassignments from the USPTO. The

USPTO registers the transfer of patents in the form of reassignments, which acknowledge the transfer of the rights , title, and interest in a patent. A typical assignment is characterized by a unique identifier (i.e. reel frame), the patent number, the names of the buyer (i.e. assignee) and the seller (i.e. assignor), and the date in which the private agreement between the two parties was signed (i.e. execution or acknowledged date). I obtained data on daily reassignments beginning in 1980 until 2012. Because the main interest in these data ultimately lies in the reallocation of the ownership of patents for technological purposes, I exclude from my data assignments recorded as administrative events, such as a name change, a security interest, a correction etc. (Serrano, 2010). Additionally, because many recorded assignments represent transactions between inventors and their employers as of the grant date of the patent, I exclude all reassignments where the assignee corresponds to the primary assignee registered at the patent office, and only include subsequent assignments (i.e. reassignments) in the data. This task is complicated by the fact that the names of the buyers of patents are not standardized by the USPTO. To overcome this concern, I standardize the names of the buyers for the reassignment data, and proceed to eliminate all the transactions for which the buyer in the transaction record matches the primary assignee of the patents as registered in the HBS data-set. To match these names I use the same fuzzy matching procedure explained in Appendix 1.

Using the clean reassignment data, I combine it to my sample of prior patents using the patent number. Table 16, Panel A reports summary statistics. Of the 2,336 prior patents in my sample, 375 are sold by their primary assignees. More interestingly, only 62 of these are sold before the VC financing event, whereas 313 are sold after the company is first financed by a VC. Thus, there is an increase in the probability that a patent is sold after the issuing company is financed by a VC. Table 16, Panel B reports this result, and shows that even after controlling for the average likelihood that similar patents are traded, the probability that a prior patent is sold increases after the financing event.

Figure 2, illustrates the annual likelihood that a patent is traded, in a window of 2 years before, and 7 after the issuing company is financed by a VC for the first time. The solid line represents the annual probability that a prior patent is sold. The dashed line corresponds to the average probability that matching patents are sold. Consistent with Table 16, Panel B, Figure 2 shows that even after controlling for the likelihood that a matching patent is sold, the likelihood that a prior patent is traded increases after companies are financed by a VC.

To test whether this increase in the likelihood that a patent is sold, following the financing event of the issuing company, can explain the effect of VCs on knowledge diffusion, I split my sample of prior patents into two groups, those that are sold and those that are not sold by 2012. I then compare how citations to both of these groups of patents increase after their companies are first financed by a VC, and relative to matching patents. Table 16, Panel C summarizes results, and shows that for both groups of patents, citations increase after the financing event. Thus, although there is an increase in the likelihood that patents are traded after VCs invest in companies, this alone cannot explain the whole effect of VCs on citations to prior patents.

One drawback from the reassignment data is that it is not exhaustive of all the forms in which a company's IP can be traded. Since there is no systematic data on license transactions at the patent level, I cannot be sure whether the effect of VCs on the likelihood that patents are traded, also applies to alternative methods of technology transfer such as the licensing of patents. In addition, I cannot be sure that the effect of VCs on the propensity to license (assuming there is one) cannot entirely explain the effect of VCs on citations to prior patents. My claim is that observable patent trades cannot entirely explain the effect that VCs exert on citations to prior patents.

3.4 Discussion

In this section I examined three mechanisms for knowledge diffusion that can be affected by VCs: company networks, inventor mobility and IP trade. Consistent with the active investor effect of VCs, I find suggestive evidence that the increase in citations is concentrated inside VC portfolios. Additional results suggest that this concentration is associated to inventor turnover across companies financed by the same VC.

Consistent with the certification effect of VCs, I find that the increase in citations is not exclusive to VC portfolios or to the VC industry, cannot be entirely traced back to inventor mobility, and that after the VC investment, companies are more likely to sell their prior patents. As suggestive evidence of this certification effect, I looked for evidence that the popularity of companies increases after they are financed by a VC. To do that, I take the names of companies financed by VCs in 2006 reported in VentureXpert and download from Google Insights weekly hits for these names in Google from 2004 until 2011. I standardize names by stripping them of punctuation, capitalization and common acronyms. Figure 3 compares normalized searches to these names (weekly searches are divided by the maximum number of searches in the entire period and multiplied by 100), and to the word "Gold" for the same period. The figure shows an increase in the number of hits after 2006 for the VC-backed companies and relative to the word Gold. This is consistent with VC financing increasing the exposure of their targets.

3.5 Extension: Effect of VCs on the distribution of citations across technological classes

Having a VC as an investor may also affect the distribution of citations across technological classes. To test this idea, in unreported results I look at the dispersion of citations received by prior patents across technology classes. To that end, I construct the "generality" of citations to prior and matching patents following Hall et al. (2001). 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.

I find that on average the generality of prior patents, relative to matching patents, increases post VC financing. However, the effect recedes once I control for patent heterogeneity. My results suggest that the VC financing event has no significant effect on the dispersion of forward citations to prior patents, relative to matching patents, across technology classes.

4 Conclusion

In this paper I investigate how VCs affect innovation. I focus on the effect of VCs on three mechanisms for knowledge diffusion: company networks, inventor mobility and patent sales. Following the innovation literature, I use data on citations to patents as a measure of knowledge diffusion. My main finding is that after companies are financed by a VC there is a causal increase in the use by third-party inventors of the company's knowledge to create more innovation.

Consistent with VCs facilitating the transfer of knowledge inside their portfolios, I find suggestive evidence that the increase in citations is more pronounced inside VC portfolios. Additional results suggest that this concentration is associated to inventor turnover across companies financed by the same VC. Consistent with VCs certifying the IP of their targets, I find that the increase in citations is not exclusive to VC portfolios or the VC industry, but instead diffuses more generally. Additional results show, that the increase in citations outside VC portfolios cannot be entirely explained by an exodus of inventors after the VC investment or after the VC exit. Finally, I also find that patents are more likely to be sold post-financing.

Estimating the effect of VCs on knowledge diffusion is challenging due to the unobservability of knowledge diffusion, heterogeneity in the types of knowledge, trends in production of knowledge, and non randomness in the choices made by VCs. To overcome these challenges, my broad empirical strategy estimates the increase in the likelihood of a citation to the same patent after a VC invests for the first time in the issuing company, and relative to other patents classified in the same technology-class and filed the same year. My identification strategy controls parametrically for patent heterogeneity and trends in citations at the technology-class and application- year level. To address concerns of non random timing in VC selection, I use variations in the size of public pension funds in the home-state of companies as an exogenous determinant of the timing of VC investments. The main insight behind this instrumental variable strategy is that in states and periods where pension pools are larger, domestic VC firms are more likely to raise capital and invest it locally. I also address concerns of a potential violation of the exclusion restriction, by using as dependent variable relative citations at the state-level which removes any state-level correlation between the instrument and the dependent variable.

The findings of this paper contribute to our understanding of how financial intermediaries affect innovation and has implications for the design of innovation policy. Governments around the globe have increasingly sought to increase innovation by stimulating VCs. However, most of these efforts have failed (Lerner, 2009), and led to questions regarding the efficiency of these policies. My paper finds evidence that VCs have a multiplier effect on innovation that is not only concentrated inside the VC industry and which can have important distributional consequences. My finding that knowledge transfer inside VC portfolios is associated to inventor turnover, suggests that regulation regarding mobility of workers, such as non compete covenants, could negatively affect the multiplier effect of VCs on innovation. This result is consistent with other papers in the literature (Samila and Sorenson, 2010).

However, this paper does not address the general equilibrium effects of VCs on innovation. It could be that by imposing a commercial focus VCs encourage innovation only on certain areas of research with high short term rewards but that can slow innovation in the long run. My findings provide some evidence that this does not seem to be prevalent as the distribution of citations across technology classes does not seem to significantly change after the investment. Yet, a better understanding of this potential effect is relevant in order to analyze welfare consequences of such policies.

References

Aghion, Dewatripont and Stein (2008). Academic Freedom, Private-Sector Focus and the Process of Innovation. The RAND journal of Economics. Vol. 39 (3) 617-635.

Agrawal and Singh (2011). Recruiting for Idea: How Firms Exploit the Prior Inventions of New Hires. Management Science. Vol. 57, No. 1, pp. 129–150

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

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

Angrist and Pischke (2008) Mostly Harmless Econometrics. Princeton University Press.

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

Azoulay, Graff Zivin and 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), pp. 107-155, 2012. Chicago, IL: University of Chicago Press.

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

Bernstein (2012). Does Going Public Affect Innovation? HBS mimeo.

Bernstein, Lerner, Sørensen and Stromberg (2008) Private Equity and Industry Performance. HBS Working Paper

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

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

Dahl and Pederson (2005). Social Networks in the R&D Process: The Case of the Wireless Communication Industry Around Aalborg, Denmark, Journal of Engineering and Technology Management. Vol. 22, No. 1-2, pp. 75-92

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

Florida and Kenney (1988), Venture Capital, High Technology and Regional Development, Regional Studies 22 (1): 33-48.

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

Gourieroux (1984) Pseudo Maximum Likelihood Methods: Application to Poisson Models, INSEE mimeo, Paris, 1981

Hall, Jaffe and Trajtenberg (2001) The NBER Patent Citations Data File: Lessons, Insights and Methodological Tools. NBER working paper 8498

Hall, Jaffe and Trajtenberg (2005) Market value and patent citations. RAND Journal of Economics, Vol 36. No 1.pp16-38

Hall, Mairesse and Turner (2007). Identifying Age, Cohort and Period Effects in Scientific Research Productivity - Discussion and Illustration Using Simulated and Actual Data on French Physicists. Economics of Innovation and New Technology, Taylor and Francis Journals, vol. 16(2), pages 159-177.

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

Hellman and Puri (2000) The Interaction between Product Market and Financing Strategy: The Role of Venture Capital. Review of Financial Studies. Vol. 13, No. 4, pp. 959-984,

Hellman and Puri (2002) Venture capital and the Professionalization of Start-Up Firms: Empirical Evidence, Journal of Finance 57, 169-197.

Hirukawa and Ueda (2011) Venture Capital and Innovation: Which is first? Pacific Economic Review. Vol 16, Issue 4, pp. 421-465.

Hochberg, Ljungqvist and Lu (2007) Whom you Know Matters: Venture Capital Networks and Investment Performance. The Journal of Finance, Vol 62, Issue 1, 251-301.

Hochberg and Rauh (2011) Local Overweighting and Underperformance: Evidence from Limited Partner Private Equity Investments.

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

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

Jaffe, Adam B., (1989). "Characterizing the "technological position" of firms, with application to quantifying technological opportunity and research spillovers," Research in Policy, Elsevier, vol. 18(2), pages 87-89, April.

Jaffe, Henderson and Trajtenberg (1993). Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations. Quarterly Journal of Economics 108, 577â€“598.

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

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

Kogut and Zander (1993). Knowledge of the firm and the evolutionary theory of the multinational corporation. Journal of International Business Studies, 24(4) 625-645.

Kim and Marschke (2005) Labor Mobility of Scientists, Technological Diffusion, and the Firm's Patenting Decision. RAND Journal of Economics. Vol 36. No. 2, pp.298-317

King (1989).Variance Specification in Event Count Models: From Restrictive Assumptions to a Generalized Estimator. American Journal of Political Science, 33, 762-784.

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

Kortum and Lerner (1998) Assessing the Contribution of Venture Capital to Innovation. RAND Journal of Economics. Vol. 31, No.4, 674-692.

Imbens and Angrist (1994) Identification and Estimation of Local Average Treatment Effects. Econometrica, 467-475.

Lai, D'Amour and Fleming (2009) The Careers and Coauthorship Networks of U.S. Patent-Holders, since 1975, http://hdl.handle.net/1902.1/12367 UNF:5:daJuoNgCZlcYY8RqU+/j2Q== Harvard Business School;Harvard Institute for Quantitative Social Science [Distributor] V3 [Version]

Lemley and Sampat (2010). Examiner Experience and Patent Office Outcomes. Working Paper, Columbia University

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

Lerner (2009) *Boulevard of Broken Dreams.* Princeton University Press. Princeton, New Jersey.

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

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

Megginson and Weiss (1991) Venture Capitalist Certification in Initial Public Offerings. The Journal of Finance. Vol. 46, No. 3.

Mollica and Zingales (2007) The Impact of Venture Capital on Innovation and the Creation of New Businesses, mimeo.

Nanda and Rhodes-Kropf (2011). Investment Cycles and Startup Innovation. HBS Working Paper No. 12-032.

Popov and Rosenboom (2009) Does Private Equity Investment Spur Innovation? Evidence from Europe. ECB Working Paper No. 1063.

Roach and Cohen (2010) Patent citations as indicators of knowledge flows from public research. Mimeo

Rogers, E., J. Larsen. 1984. Silicon Valley Fever. Basic Books, New York.

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

Romer (1990). Endogenous Technological Change. The Journal of Political Economy, Vol. 98, No. 5, Part 2: The Problem of Development: A

Conference of the Institute for the Study of Free Enterprise Systems. (Oct., 1990), pp. S71-S102.

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

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

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

Sampat and Ziedonis (2004) Patent Citations and The Economic Value of Patents in Handbook of Quantitative Science and Technology Research, p. 277-297

Saxenian (1990) Regional Networks and the Resurgence of Silicon Valley. California Management Review. Vol. 33 No. 1, pp. 89-103

Saxenian (1994) Regional advantage. Cambridge, MA: Harvard University Press.

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

Song, Almeida and Wu (2003). Learning-by-Hiring: When is Mobility More Likely to Facilitate Interfirm Knowledge Transfer?, Management Science 49(4): 351-365

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

Sorenson and Stuart (2001). Syndication Networks and the Spatial Distribution of Venture Capital Investments . The American Journal of Sociology. Vol. 106, No. 6, pp.1546-1588

Sorenson and Fleming (2004). Science and the Diffusion of Knowledge. Research Policy, Vol. 33, 1615-1634, 2004

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

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

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

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

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

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

Figure 1- Estimated temporal trends in citations to prior patents

The solid lines in the plot correspond to the coefficient estimates of a QMLE specification in which the dependent variable corresponds to annual citations to prior patents, and the explanatory variables are Event Year dummies. I restrict the sample to a [-2,6] 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 financing event (Event Year 0).

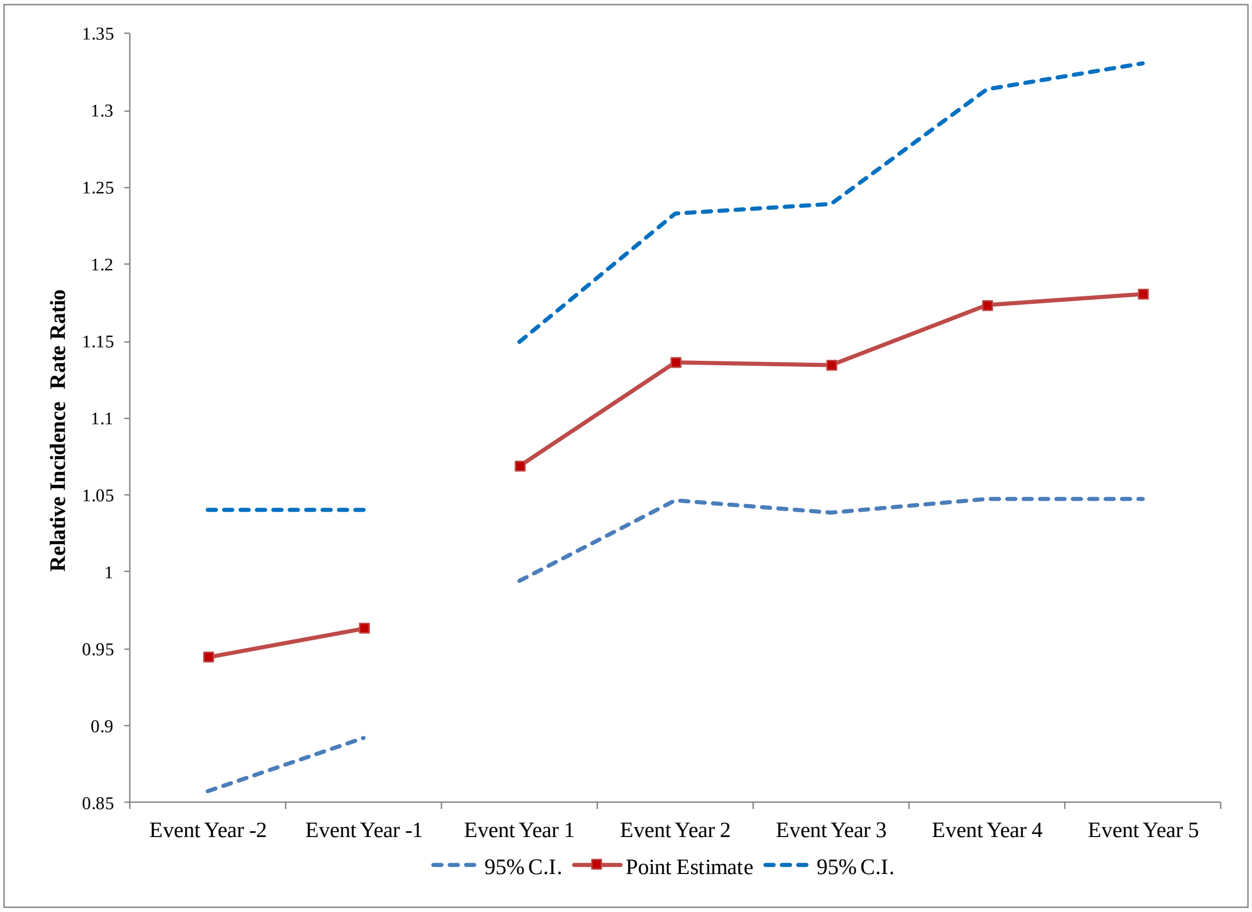

Figure 2- Patent sale likelihood

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

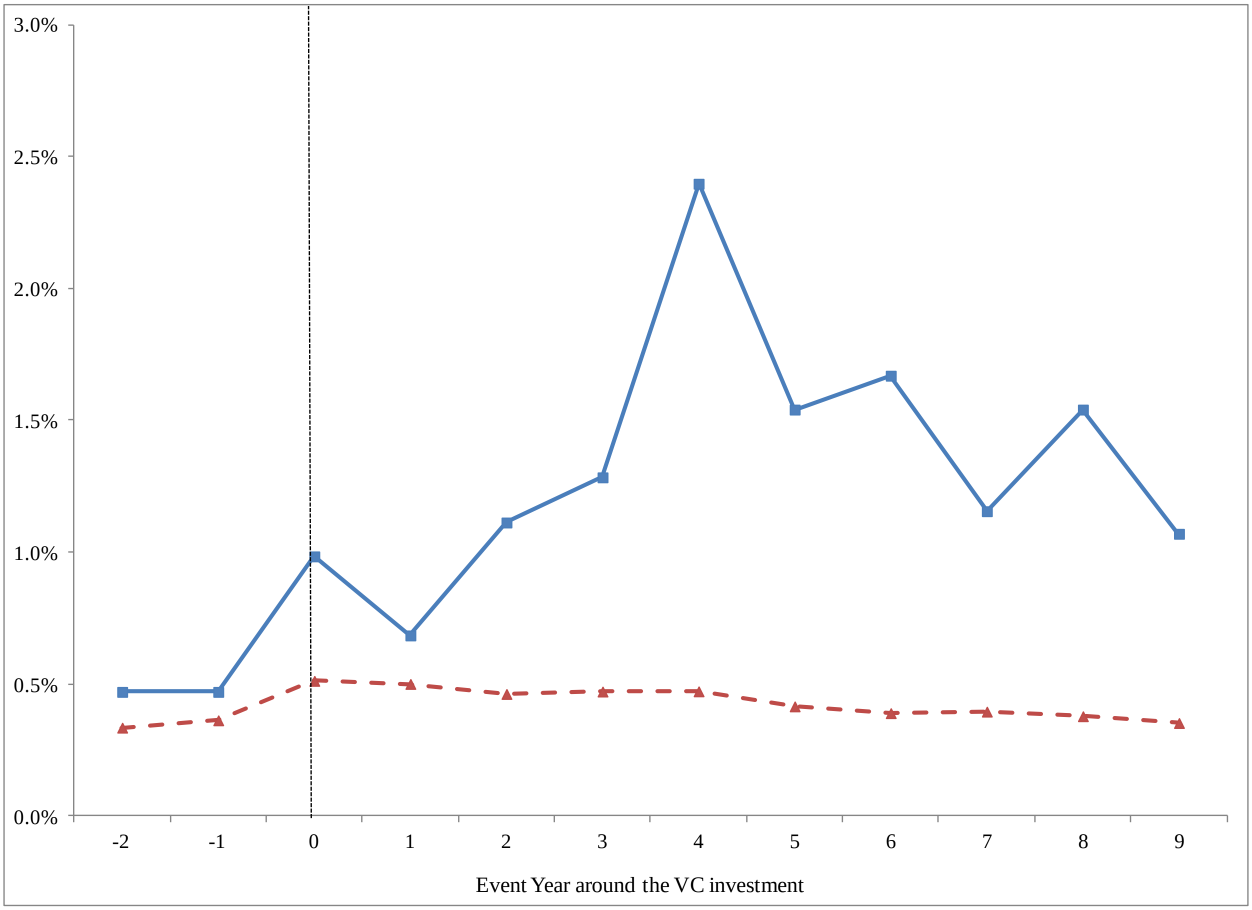

Figure 3- Exposure Effect of VC financing

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.

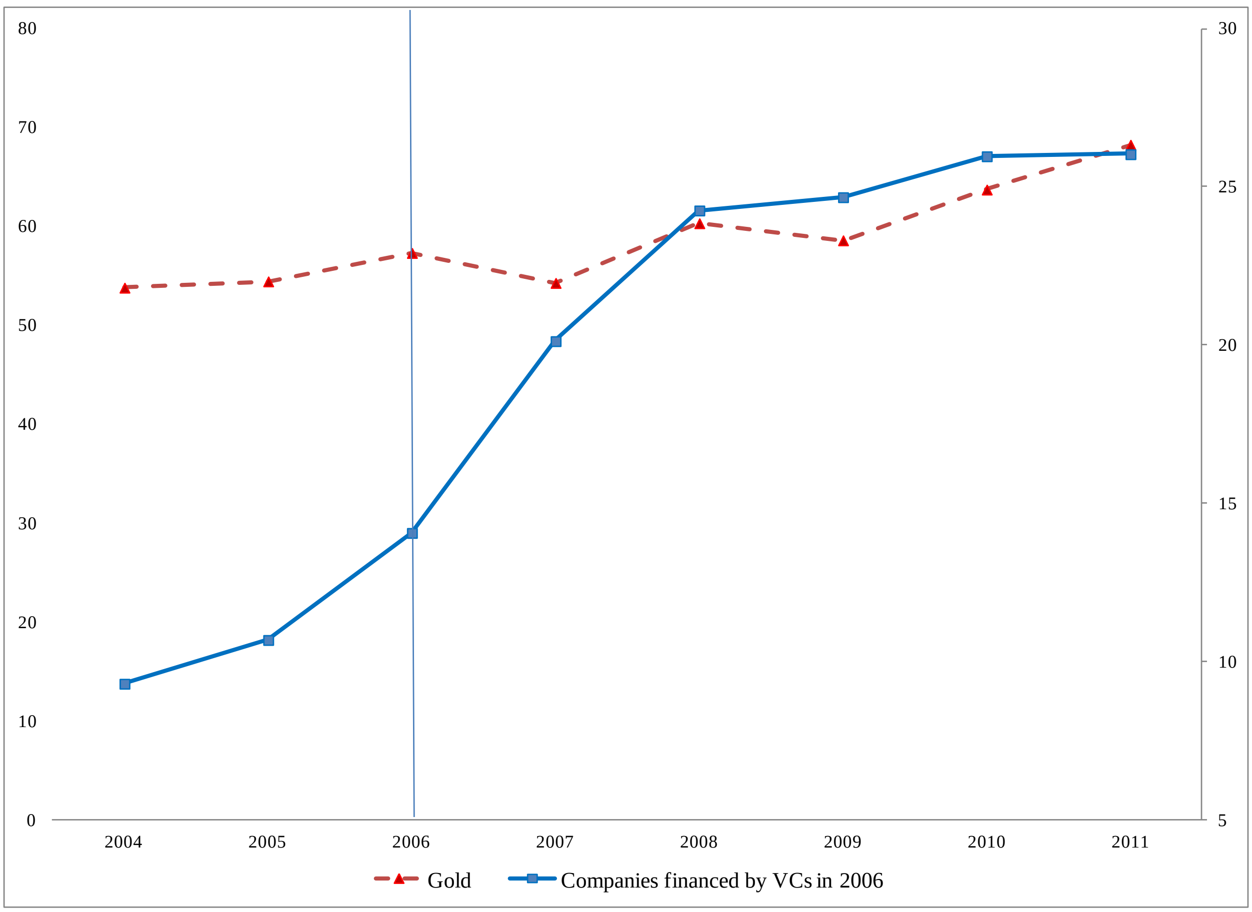

Table 1 - Summary statistics analysis sample

The sample consists of 2,336 patents (prior patents) that were awarded to 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, E and F, 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. The industry classification used in Panels D and E, is based on the VentureXpert files. Panel H excludes all patents in the full matched sample granted after 2005.

Panel A. Application and grant years of prior patents, and transaction years for the VC deals involving the companies with prior patents in my sample

	Prior Patents		VC deals
Year	Applications	Grants	First time financing
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 VC-backed companies with patenting and associated patents: Top States in Analysis Sample

	% of Companies			% of Patents	
State	Analysis Sample	Full Matched Sample	Overall VC Population	Analysis Sample	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 type of investment by VC firms in companies with prior patents

		% of Companies		
	Number of Companies	Analysis Sample	Full matched Sample	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. Industry distribution of VC-backed companies with prior patents

	Number of companies	% of Companies		
		Analysis Sample	Full matched Sample	Overall VC Population
Biotechnology	63	8.4%	9.9%	6.1%
Communications and Media	75	10.0%	11.0%	10.3%
Computer Hardware	51	6.8%	8.9%	6.3%
Computer Software	94	12.5%	16.3%	21.3%
Consumer Related	33	4.4%	2.0%	4.8%
Industrial Energy	97	12.9%	8.0%	5.1%
Internet Specific	37	4.9%	8.5%	20.7%
Medical Health	145	19.3%	16.8%	11.6%
Other Products	30	4.0%	2.6%	6.6%
Semiconductors	127	16.9%	16.0%	7.2%
Total	752			

Panel E. Industry distribution of prior patents

	Number of patents	% of Patents	
		Analysis Sample	Full matched Sample
Biotechnology	199	8.5%	11.6%
Communications and Media	215	9.2%	8.9%
Computer Hardware	104	4.5%	17.2%
Computer Software	180	7.7%	13.6%
Consumer Related	125	5.4%	1.1%
Industrial Energy	377	16.1%	4.4%
Internet Specific	50	2.1%	2.0%
Medical Health	505	21.6%	15.1%
Other Products	101	4.3%	0.8%
Semiconductors	480	20.5%	25.3%
Total	2,336		

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

		% of Companies		
	Number of companies	Analysis Sample	Full matched Sample	Overall VC Population
Acquisition	282	37.5%	34.9%	30.8%
Active	209	27.8%	29.9%	35.8%
Bankruptcy - Chapter 11	4	0.5%	0.5%	0.5%
Bankruptcy - Chapter 7	5	0.7%	0.8%	0.8%
Defunct	140	18.6%	14.4%	19.9%
In Registration	1	0.1%	0.4%	0.2%
LBO	7	0.9%	0.7%	0.8%
Merger	10	1.3%	1.6%	1.6%
Other	2	0.3%	0.4%	0.4%
Pending Acquisition	1	0.1%	0.2%	0.2%
Went Public	91	12.1%	16.3%	9.1%
Total	752			

Panel G. Distribution of patent age at the time of first time VC financing

	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 H. Annual Citations and Generality

	Mean	S. D.	Med.	Min	Max	Obs.
Annual Citations	0.92	2.45	0.00	0.00	60.00	43,519
Annual Generality	0.20	0.27	0.00	0.00	0.80	14,469
Annual Adjusted Generality	0.58	0.40	0.67	0.00	1.00	7,336

Table 2 - Summary statistics restricted sample 1993-2008

Information on public state pension funds is available from 1993 to 2008. The sample of prior patents restricted to this period consists of 1,170 patents (prior patents) that were awarded to 434 VC-backed companies, at least two years before they were first financed by a VC (289 VC firms). 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 by the producer Price Index and expressed in 2008 U.S. millions. For Panels B, C, D, E and F, the percentage of companies used for comparisons consists of 752 companies with prior patents from the full analysis sample.

Panel A. Application and grant years of prior patents, and VC financing years for the issuing companies of prior patents

Year	Applications	Grants	First time financing
1993	95		
1994	99	25	
1995	139	67	13
1996	117	67	18
1997	188	80	25
1998	207	131	44
1999	117	150	49
2000	126	158	81
2001	82	148	53
2002		106	74
2003		96	77
2004		51	
2005		44	
2006		30	
2007		9	
2008		8	
Total	1,170	1,170	434

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

	Pension Funds' Assets		% of Patents		% of Companies	
	Mean	Std. Dev	Restricted Sample	Analysis Sample	Restricted Sample	Analysis Sample
AL	0.16	0.03	0.9%	0.4%	0.5%	0.3%
AZ	0.19	0.05	1.8%	1.8%	1.2%	1.6%
CA	2.59	0.79	41.2%	32.6%	38.2%	35.0%
CO	0.2	0.06	3.3%	3.6%	2.8%	2.4%
CT	0.15	0.03	1.6%	3.3%	2.3%	2.7%
DC	0.03	0.01	0.5%	0.3%	0.2%	0.1%
FL	0.64	0.21	2.4%	2.0%	2.1%	2.0%
GA	0.34	0.1	0.8%	1.3%	1.8%	2.0%
ID	0.04	0.02	1.3%	0.7%	0.5%	0.4%
IL	0.6	0.16	1.9%	2.1%	2.3%	2.5%
IN	0.12	0.04	0.2%	0.2%	0.2%	0.4%
LA	0.17	0.04	0.4%	0.6%	0.7%	0.4%
MA	0.27	0.09	8.2%	10.5%	11.8%	14.1%
MD	0.26	0.06	2.3%	3.4%	1.4%	2.4%
ME	0.05	0.02	0.3%	0.2%	0.2%	0.1%
MI	0.45	0.1	1.1%	1.2%	1.6%	1.6%
MN	0.27	0.06	1.5%	1.3%	2.3%	2.3%
MO	0.25	0.07	0.6%	0.7%	0.7%	0.9%
NC	0.35	0.1	1.5%	0.8%	2.1%	1.5%
NH	0.03	0.01	0.9%	1.2%	1.2%	1.5%
NJ	0.35	0.07	1.9%	2.3%	2.8%	2.7%
NM	0.09	0.03	0.7%	0.3%	0.7%	0.4%
NV	0.09	0.03	0.1%	0.0%	0.2%	0.1%
NY	1.67	0.41	3.9%	3.4%	4.4%	4.0%
OH	0.77	0.15	2.0%	2.2%	1.6%	1.9%
OR	0.21	0.11	0.6%	1.9%	0.7%	0.7%
PA	0.56	0.13	3.0%	5.0%	3.0%	3.5%
TN	0.19	0.05	2.3%	2.0%	0.2%	0.8%
TX	0.86	0.25	8.8%	9.7%	6.5%	5.3%
UT	0.09	0.03	0.3%	0.7%	0.7%	0.8%
VA	0.29	0.08	0.7%	0.9%	1.4%	1.2%
VT	0.02	0	0.2%	0.3%	0.2%	0.3%
WA	0.3	0.08	2.8%	1.9%	3.2%	2.7%
WI	0.42	0.1	0.3%	0.3%	0.5%	0.4%
Total	0.38	0.11				

Panel C. Distribution of type of investment by VC firms in companies with prior patents

	Number of Companies	Percentage of sample	
		Restricted Sample	Analysis Sample
Bridge Loan	13	3.0%	2.8%
Early Stage	183	42.2%	34.2%
Expansion	156	35.9%	39.8%
Later Stage	49	11.3%	12.1%
Seed	33	7.6%	11.2%
Total	434		

Panel D. Industry distribution of VC investments in companies with prior patents

	Number of companies	% of Companies	
		Restricted Sample	Analysis Sample
Biotechnology	53	12.2%	8.4%
Communications and Media	43	9.9%	10.0%
Computer Hardware	24	5.5%	6.8%
Computer Software	63	14.5%	12.5%
Consumer Related	9	2.1%	4.4%
Industrial Energy	30	6.9%	12.9%
Internet Specific	30	6.9%	4.9%
Medical Health	94	21.7%	19.3%
Other Products	12	2.8%	4.0%
Semiconductors	76	17.5%	16.9%
Total	434		

Panel E. Industry distribution of prior patents

	Number of patents	% of Patents	
		Restricted Sample	Analysis Sample
Biotechnology	147	12.6%	8.5%
Communications and Media	115	9.8%	9.2%
Computer Hardware	42	3.6%	4.5%
Computer Software	121	10.3%	7.7%
Consumer Related	51	4.4%	5.4%
Industrial Energy	85	7.3%	16.1%
Internet Specific	38	3.2%	2.1%
Medical Health	270	23.1%	21.6%
Other Products	24	2.1%	4.3%
Semiconductors	277	23.7%	20.5%
Total	1,170		

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

		% of Companies	
	Number of companies	Restricted Sample	Analysis Sample
Acquisition	152	35.0%	37.5%
Active	170	39.2%	27.8%
Bankruptcy - Chapter 11	3	0.7%	0.5%
Bankruptcy - Chapter 7	4	0.9%	0.7%
Defunct	50	11.5%	18.6%
In Registration	1	0.2%	0.1%
LBO	2	0.5%	0.9%
Merger	3	0.7%	1.3%
Other	0	0.0%	0.3%
Pending Acquisition	1	0.2%	0.1%
Went Public	48	11.1%	12.1%
Total	434		

Panel G. Distribution of patent age at the time of first time VC financing

	Number of patents	Percentage of sample
2 Years	329	28.12
3 Years	407	34.79
4 Years	186	15.9
5 Years	106	9.06
Between 6 years and 10 years	142	12.14
Total	1,170	

Panel I. Annual Citations and Generality

	Mean	S. D.	Med.	Min	Max	Obs.
Annual Citations	1.38	3.28	0.00	0.00	60.00	13,965
Annual Generality	0.25	0.29	0.00	0.00	0.88	5,564
Annual Adjusted Generality	0.61	0.38	0.67	0.00	1.00	3,335

Table 3– Incidence Rate Ratios and Relative Incidence Rate Ratios

The table presents the Incidence Rate Ratio (IRR) and Relative Incidence Rate Ratio (RIRR) of the exposure to first time VC financing on citations to patents. Columns (1) and (2) summarize average annual citations to prior patents before and after the VC investment. Columns (3) and (4) summarize average annual citations to matching patents, which match prior patents by technology-class and application-year. In Columns (1)-(4) standard deviations are included in parentheses and the number of observations in squared brackets. In Columns (5) and (7) the p-values of the t-test are reported in parentheses, and in Columns (6) and (8) the p-values are reported in parentheses. Column (5) presents the difference in annual average citations to prior patents, and column (6) presents the IRR for prior patents. The IRR is constructed as the ratio between average annual citations post VC-financing, and average annual citations pre VC-financing. Column (7) presents the "difference in difference", defined as the difference between the change in average annual citations post financing to prior patents, and the change in average annual citations post financing to matching patents. Column (8) presents the RIRR, defined as the ratio between the IRR of prior patents and the IRR of matching patents. *, **, and *** indicate statistical significant at the 10%, 5% and 1% level.

Prior patents		Matching Patents					
Average Annual citations		Average Annual citations		Diff.	IRR	Diff. in Diff.	RIRR
Pre	Post	Pre	Post				
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
0.64	1.04	0.54	0.66	0.40***	1.63***	0.28***	1.33***
(1.69)	(2.69)	(0.61)	(0.83)	(0.00)	(0.00)	(0.00)	(0.00)
[12,767]	[40,096]	[12,767]	[40,096]				

Table 4 –VC Financing and patent citations

The table presents the estimated effect of VC financing on citations to prior patents using Pooled and Fixed Effects Poisson Models. An observation is a patent-year. The dependent variable is annual forward citations. VC_{pt} is an indicator variable that equals 1 after the issuing company of the patent is first financed by a VC. The reported coefficients are incidence rates. Column (1) presents results using a Pooled Poisson model. Column (2) presents results using a Pooled Poisson model and offsetting average annual citations to matching patents in the estimation. Column (3) summarizes coefficient estimates of a patent fixed-effects Poisson model. 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	QMLE
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 average annual citations at the tech-class and app. year level	No	Yes	Yes
Patent FE	No	No	Yes

Table 5 – VC investments in new companies and local and state pension funds' assets

The table reports the relation between number of VC investments in new companies and local and state pension funds' assets. The dependent variable is stated at the beginning of each column. Observations are at the state-year level. Standard errors are clustered at the state level. $Pension\ Size$ corresponds to the value of assets held by local and state pension funds (deflated by PPI and expressed in 2008 US$ millions) lagged by 1 year. In column (2) the reported coefficient is an incidence rate. *, **, and *** indicate statistical significance at the 10%, 5% and 1% level.

	(1)	(2)
Dependent Variable	Value of new Investments	Number of new Investments
Model	OLS	QMLE
$Pension\ Size$	0.020***	1.070**
	(0.005)	(0.036)
Constant	-0.003	
	(0.002)	
Obs.	765	765
Wald		3401.93
F test	10.89	
Time F.E.	Yes	Yes
State F.E.	Yes	Yes

Table 6– GMM-IVs estimation of within-patent relationship between VC financing and patent citations

This table reports the effect of VC financing on citations to prior patents relative to matching patents. An observation is a patent-year. The dependent variable is annual forward citations. VC_{pt} is a dummy variable that equals one after the issuing company of the patent is first financed by a VC and zero otherwise. z_{pt} corresponds to the value of assets held by local and state pension funds, in the home-state of the company, deflated by PPI, expressed in 2008 US$ millions, lagged by 1 year, and demeaned by state and time. For columns (2)-(4) the regression includes the average citation intensity at the technology class and application year level with a coefficient fixed to 1. For columns (2)-(4) the estimated 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 state level. *, **, and *** indicate statistical significance at the 10%, 5% and 1% level.

	(1)	(2)	(3)	(4)
Dependent Variable	VC_{pt}	Y_{pt}^{sa}	Y_{pt}^{sa}	Y_{pt}^{sa}
Regression	First Stage	Within- patent	Reduced Form	IV
Model	OLS	QMLE	QMLE	GMM-IVs
VC_{pt}		1.214***		1.805**
		(0.076)		(0.540)
z_{pt}	0.713***		1.440***	
	(0.037)		(0.052)	
Constant	0.684***			
	(0.002)			
Observations	10,071	10,071	10,071	10,071
Number of cited	1,058	1,058	1,058	1,058
Number of Companies	411	411	411	411
Offset citations baseline at the tech-class and app. year level	No	Yes	Yes	Yes
F test for Weak Instruments	367.73			

Table 7– First robustness check GMM-IVs estimation of within-patent relationship between VC financing and patent citations

This table reports the effect of VC financing on citations to prior patents relative to matching patents issued in the same state. An observation is a patent-year. The dependent variable is annual forward citations. VC_{pt} is a dummy variable that equals one after the issuing company of the patent is first financed by a VC and zero otherwise. z_{pt} corresponds to the value of assets held by local and state pension funds, in the home-state of the company, deflated by PPI, expressed in 2008 US$ millions, lagged by 1 year, and demeaned by state and time. For columns (2)-(4) the regression includes the average citation intensity at the technology-class, application-year, and state level with a coefficient fixed to 1. For columns (2)-(4) the estimated 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 state level. *, **, and *** indicate statistical significance at the 10%, 5% and 1% level.

	(1)	(2)	(3)	(4)
Dependent Variable	VC_{pt}	Y_{pt}^{sa}	Y_{pt}^{sa}	Y_{pt}^{sa}
Regression	First Stage	Within- patent	Reduced Form	IV
Model	OLS	QMLE	QMLE	GMM-IVs
VC_{pt}		1.286***		1.837***
		(0.052)		(0.343)
z_{pt}	0.737***		1.416***	
	(0.041)		(0.042)	
Constant	0.675***			
	(0.003)			
Observations	8,072	8,072	8,072	8,072
Number of cited	951	951	951	951
Number of Companies	388	388	388	388
Offset citations baseline at the tech-class, app. year and state level	Yes	Yes	Yes	Yes
F test for Weak Instruments	326.40			

Table 8– Second robustness check GMM-IVs estimation of within-patent relationship between VC financing and patent citations

This table reports the effect of VC financing on out-state citations to prior patents relative to matching patents issued in the same state. An observation is a patent-year. The dependent variable is out-state citations, which correspond to the number of citations received by the patent at time t from patentees located in a different state. VC_{pt} is a dummy variable that equals one after the issuing company of the patent is first financed by a VC and zero otherwise. z_{pt} corresponds to the value of assets held by local and state pension funds, in the home-state of the company, deflated by PPI, expressed in 2008 US$ millions, lagged by 1 year, and demeaned by state and time. For columns (2)-(4) the regression includes the citations baseline at the technology class, application year and state level with a coefficient fixed to 1 For columns (2)-(4) 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 state level. *, **, and *** indicate statistical significance at the 10%, 5% and 1% level.

	(1)	(2)	(3)	(4)
Dependent Variable	VC_{pt}	Out-state citations	Out-state citations	Out-state citations
Regression	First Stage	Within- patent	Reduced Form	IV
Model	OLS	QMLE	QMLE	GMM-IVs
VC_{pt}		1.293***		1.851**
		(0.050)		(0.537)
z_{pt}	0.733***		1.456***	
	(0.040)		(0.054)	
Constant	0.678***			
	(0.002)			
Observations	7,741	7,741	7,741	7,741
Number of cited	915	915	915	915
Number of Companies	379	379	379	379
Offset citations baseline at the tech-class, app. year and state level	Yes	Yes	Yes	Yes
F test for Weak Instruments	328.30			

Table 9- Originality

This table reports Originality and Relative Originality measures for prior patents that are funded in hot versus cold markets. A patent is said to have been financed in a hot market if the variation in local public pension funds' assets at the time and state level is above the 75th percentile of all years and states in the sample. Analogously, a patent is said to have been financed in a cold market if the variation in local public pension funds' assets at the time and state level is below the bottom 25th percentile of all years and states in the sample. *, **, and *** indicate statistical significance at the 10%, 5% and 1% level.

	Top 75%	Bottom 25%	Difference
Originality	0.57	0.47	0.10***
Originality Adjusted	0.65	0.54	0.11***
Relative Originality	0.15	0.11	0.05*
Relative Originality Adjusted	0.15	0.10	0.05*

Table 10 – VC Portfolio Size and Citations to Prior Patents by Type

This table summarizes the number of members in the VC portfolios for companies in the analysis sample, and citations to prior patents by type of citation. Panel A reports average size of the VC portfolio for companies with prior patents. Panel B, shows the distribution of average size of the VC portfolio for companies with prior patents by year of VC financing. Panel C, reports distribution of citations to prior patents by type of citation and home-state of companies.

Panel A. Average size of VC portfolios for companies with prior patents

	Mean	Std. Dev.	Min	Max	p50
Portfolio Members	16.66	29.89	0	351	6

Panel B. Average size of VC portfolios for companies with prior patents by year of VC investment

	Average	Standard Deviation
1978	2.33	2.08
1979	2.83	2.4
1980	5.5	7.55
1981	5.93	4.85
1982	12.53	13.38
1983	10.93	9.65
1984	9.07	13.22
1985	12.25	13.01
1986	11.18	12.15
1987	6.85	7.91
1988	10.92	13.77
1989	15.47	22.18
1990	15.25	18.3
1991	19.13	29.21
1992	10.38	7.84
1993	23.89	36.71
1994	20.17	22.64
1995	15.42	38.18
1996	14.13	23.59
1997	13.03	22.47
1998	21.54	30.05
1999	28.34	59.84
2000	19.07	28.59
2001	13.16	16.98

2002	18.21	37.28
2003	20.28	29.44
Total	16.66	29.89

Panel C. Average Citations to Prior Patents by type of citation and home-state of company

State	Average Annual Citations to Prior Patents by type		
	Non Portfolio- lined		Portfolio- linked
	Non VC-backed	Regular VC-backed	
AL	1.168	0.042	0.000
AZ	0.600	0.019	0.000
CA	0.957	0.163	0.006
CO	1.165	0.233	0.000
CT	0.452	0.046	0.001
DC	2.275	0.150	0.000
DE	0.524	0.000	0.000
FL	1.313	0.077	0.027
GA	0.770	0.089	0.009
IA	0.250	0.000	0.000
ID	0.327	0.122	0.000
IL	0.941	0.247	0.000
IN	0.241	0.009	0.000
KS	0.040	0.000	0.000
LA	1.324	0.055	0.000
MA	0.696	0.094	0.003
MD	0.666	0.099	0.023
ME	0.391	0.023	0.000
MI	0.515	0.130	0.000
MN	1.084	0.402	0.109
MO	0.548	0.003	0.017
NC	0.306	0.100	0.000
ND	0.534	0.000	0.000
NE	0.336	0.008	0.000
NH	0.856	0.152	0.000
NJ	0.534	0.042	0.000
NM	0.359	0.000	0.000
NV	0.533	0.000	0.000
NY	0.983	0.098	0.001
OH	1.186	0.104	0.011
OR	0.546	0.068	0.024

PA	0.455	0.053	0.004
RI	0.306	0.000	0.000
SC	0.453	0.012	0.000
TN	0.589	0.034	0.000
TX	0.643	0.057	0.000
UT	0.744	0.158	0.000
VA	1.789	0.193	0.000
VT	0.401	0.030	0.000
WA	1.193	0.095	0.000
WI	0.587	0.156	0.000
WY	0.267	0.000	0.000
Total	0.799	0.113	0.006

Table 11– Incidence Rate Ratios and Relative Incidence Rate Ratios inside and outside VC portfolios

The table presents the Incidence Rate Ratio (IRR) and Relative Incidence Rate Ratio (RIRR) of the exposure to first time VC financing on citations to patents. Columns (1) and (2) summarize average annual citations to prior patents before and after the VC investment. Columns (3) and (4) summarize average annual citations to matching patents, which match prior patents by technology-class and application-year. Standard deviations are included in squared-bracket. Column (5) presents the difference in annual average citations to prior patents, and column (6) presents the IRR for prior patents. The IRR is constructed as the ratio between average annual citations post VC-financing, and average annual citations pre VC-financing. Column (7) tests whether the estimated IRR for portfolio- linked citations exceeds the IRR for non portfolio- linked citations. Column (8) presents the "difference in difference", defined as the difference between the change in average annual citations post financing to prior patents, and the change in average annual citations post financing to matching patents. Column (9) presents the RIRR, defined as the ratio between the IRR of prior patents and the IRR of matching patents. Finally, Column (10) tests whether the estimated RIRR of portfolio- linked citations exceeds the RIRR of non portfolio- linked citations. *, **, and *** indicate statistical significant at the 10%, 5% and 1% level.

	Average Annual citations													
	Prior patents				Matching Patents									
	Pre		Post		Pre		Post		Diff	IRR	Diff. IRR	Diff. in Diff.	RIRR	Diff. RIRR
	(1)		(2)		(3)		(4)		(5)	(6)	(7)	(8)	(9)	(10)
Portfolio-linked	0.00	[0.08]	0.01	[0.15]	0.00	[0.01]	0.00	[0.02]	0.006***	4.05***	2.43**	0.004***	2.44***	1.11*
Non Portfolio- linked	0.64	[1.68]	1.03	[2.68]	0.54	[0.61]	0.66	[0.82]	0.39***	1.62***		0.27***	1.33***	
Non VC-backed	0.59	[1.50]	0.89	[2.30]	0.49	[0.82]	0.58	[0.69]	0.30***	1.52***		0.21***	1.29***	
VC-backed	0.05	[0.40]	0.14	[0.82]	0.05	[0.12]	0.08	[0.17]	0.09***	2.84***		0.06***	1.64***	

Table 12 – Distribution of forward citations to prior patents inside and outside VC portfolios

The table reports changes in the distribution of forward citations to prior patents by type of citation, following the VC financing event. The dependent variable corresponds to the number of citations received by prior patents. An observation is at the patent, type of citing patentee, and year level. D_P (D_{NP}) 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 issuing company of the prior patent is first financed by a VC. 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 QMLE model requires variation in the dependent variable for each patent-type of citing assignee group for estimation, which explains the difference in observations across columns (1) and (3), and, (2) and (4). 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 for columns (1)-(4) and at the state level for columns (5)-(6). *, **, and *** indicate statistical significance at the 10%, 5% and 1% levels, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)
Model	Poisson	Poisson	QMLE	QMLE	QMLE Sample GMM-IVs	GMM-IVs
A. Estimated IRRs						
D_{NP}	0.635***	1.176***				
	(0.024)	(0.040)				
D_P	0.002***	0.868				
	(0.001)	(0.497)				
$VC_{pt} * D_{NP}$ (I)	1.620***	1.325***	1.464***	1.186***	1.215***	2.102***
	(0.067)	(0.050)	(0.047)	(0.035)	(0.083)	(0.233)
$VC_{pt} * D_P$ (II)	4.052***	2.437	2.898***	2.785**	2.416**	2.176***
	(1.619)	(1.541)	(0.790)	(1.276)	(1.040)	(0.171)
B. Difference in IRRs						
II-I	5.37	0.93	6.29	3.44	2.09	0.04
	(0.025)	(0.335)	(0.012)	(0.064)	(0.149)	(0.838)
Observations	87,038	45,064	42,191	39,299	9,062	9,062
Number of patents	2,336	2,336	2,183	2,183	1,048	1,048
Number of companies	752	752	726	726	409	409
Baseline by tech-class, app. year and type of citation	No	Yes	No	Yes	Yes	Yes
Patent-type of citation FE	No	No	Yes	Yes	Yes	Yes

Table 13 – Robustness Check: Distribution of forward citations to prior patents by type of citation

The table reports the association between VC financing and patent citations outside and inside the VC industry. The dependent variable corresponds to the number of citations received by prior patents. An observation is at the patent, type of citing patentee, and year level. D_{NVC} is a dummy that equals one if the type of citation is Non VC-backed. D_{VC} is a dummy that equals one if the type of citation is VC-backed. D_P is a dummy that equals one if the type of citation is portfolio- linked. VC_{pt} is a dummy that equals one after the issuing company of the prior patent is first financed by a VC. The Poisson model requires that the citation baseline be different from zero, which explains the difference in observations across columns (1)-(2). The QMLE model requires variation in the dependent variable for each patent-type of citing assignee group for estimation, which explains the difference in observations across columns (1) and (3), and, (2) and (4). 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 for columns (1)-(4), and at the state level for columns (5)-(6). *, **, and *** indicate statistical significance at the 10%, 5% and 1% levels, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)
Model	Poisson	Poisson	QMLE	QMLE	QMLE Sample GMM	GMM-IVs
A. Estimated IRRs						
D_{NVC}	0.585***	1.082**				
	(0.022)	(0.036)				
D_{VC}	0.049***	1.043				
	(0.005)	(0.100)				
D_P	0.002***	0.868				
	(0.001)	(0.497)				
$VC_{pt} * D_{NVC}$ (I)	1.517***	1.238***	1.396***	1.136***	1.221***	1.869***
	(0.061)	(0.046)	(0.043)	(0.034)	(0.067)	(0.281)
$VC_{pt} * D_{VC}$ (II)	2.839***	1.641***	2.221***	1.192**		
	(0.298)	(0.173)	(0.193)	(0.090)		
$VC_{pt} * D_P$ (III)	4.052***	2.437	2.898***	2.785**		
	(1.619)	(1.541)	(0.790)	(1.276)		
B. Difference in IRRs						
II-I	42.43	8.24	32.75	0.39		
	(0.00)	(0.00)	(0.00)	(0.53)		
III-I	6.16	1.15	7.20	3.79		
	(0.01)	(0.28)	(0.01)	(0.05)		
III-II	0.76	0.38	0.87	3.29		

	(0.39)	(0.54)	(0.35)	(0.07)		
Observations	130,557	71,253	54,824	48,958	8,855	8,855
Number of patents	2,336	2,336	2,183	2,183	1,033	1,033
Number of companies	752	752	726	726	406	406
Baseline by tech-class and app. year	No	No	No	No	Yes	Yes
Baseline by tech-class, app. year and type of citation	No	Yes	No	Yes	No	No
Patent-type of citation FE	No	No	Yes	Yes	Yes	Yes

Table 14– Annual Average inventor-linked citations and not inventor-linked citations

The table presents average annual inventor- and not inventor- linked citations to prior and matching patents. Panel A, reports citations from all assignees. Panel B, reports citations inside VC portfolios (portfolio- linked). Panel C reports citations outside VC portfolios (non portfolio – linked).

Panel A. All Citations

	Prior Patents				Matching Patents				IRR	RIRR
	Pre		Post		Pre		Post			
	(1)		(2)		(3)		(4)		(5)	(6)
Not 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

	Prior Patents				Matching Patents					
	Pre		Post		Pre		Post		IRR	RIRR
	(1)		(2)		(3)		(4)		(5)	(6)
Not 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 B. Non Portfolio-linked Citations

	Prior Patents				Matching Patents					
	Pre		Post		Pre		Post		IRR	RIRR
	(1)		(2)		(3)		(4)		(5)	(6)
Not 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 15 – Distribution of Not inventor- linked citations to prior patents inside and outside VC portfolios

The table reports changes in the distribution of forward citations to prior patents by type of citation, following the VC financing event. The dependent variable corresponds to the number of citations received by prior patents. An observation is at the patent, type of citing patentee, and year level. D_P (D_{NP}) is a dummy that equals one if the citation is portfolio-liked (non portfolio- linked). VC_{pt} is a dummy that equals one after the issuing company of the prior patent is first financed by a VC. The Poisson model requires that the citation baseline be different from zero, which explains the difference in observations across columns (1)-(2). The QMLE model requires variation in the dependent variable for each patent-type of citing assignee group for estimation, which explains the difference in observations across columns (1) and (3), and, (2) and (4). 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 for columns (1)-(4) and at the state level for columns (5)-(6). *, **, and *** indicate statistical significance at the 10%, 5% and 1% levels, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)
Model	Poisson	Poisson	QMLE	QMLE	QMLE Sample GMM-IVs	GMM-IVs
A. Estimated IRRs						
D_{NP}	0.635***	1.174***				
	(0.024)	(0.040)				
D_P	0.002***	0.868				
	(0.001)	(0.497)				
$VC_{pt} * D_{NP}$ (I)	1.517***	1.406***	1.387***	1.281***	1.330***	2.228***
	(0.063)	(0.053)	(0.043)	(0.038)	(0.095)	(0.288)
$VC_{pt} * D_P$ (II)	3.869***	2.620	2.789***	2.891**	6.227***	1.129
	(1.557)	(1.667)	(0.756)	(1.394)	(2.420)	(0.171)
B. Difference in IRRs						
II-I	5.51	0.96	6.66	2.83	8.29	7.53
	0.02	0.33	0.01	0.09	0.00	0.01
Observations	87,038	44,991	41,926	39,115	8,987	8,987
Number of patents	2,336	2,336	2,170	2,170	1,038	1,038
Number of companies	752	752	726	726	408	408
Baseline by tech-class, app. year, not-inventor linked, type of citation	No	Yes	No	Yes	Yes	Yes
Patent-type of citation FE	No	No	Yes	Yes	Yes	Yes

Table 16 – Summary Statistics Patent Trade

This table reports summary statistics of patent sales around the financing event of the issuing companies. Panel A presents number of prior patents that were traded before and after a VC first finances the issuing companies. Panel B compares prior patents and matching patents and their respective likelihood of being traded 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 prior patents and their respective matching patents according to whether the patents were traded or not throughout the sample.

Panel A. Number of prior patents traded

	Number	Percentage of Total
Total prior patents sold during the sample	375	16%
Prior patents sold at least once pre VC financing	62	3%
Prior patents sold at least once post VC financing	327	14%

Panel B. Annual Likelihood that a patent is traded in percentages

Prior Patents		Matching Patents		Difference	Diff. in Diff.
Pre	Post	Pre	Post		
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 traded and not traded patents before and after the VC financing event

	Prior patents		Match. Patents		Diff.	IRR	Diff.-Diff.	RIRR
	Pre	Post	Pre	Post				
Traded	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 Traded	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]				
Traded - Not Traded	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

In this section I give a detailed account on how the data set was constructed.

Investments by U.S. Venture Capital firms

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. (include footnote: In robustness checks presented in Section 3 I use the sample of PIPEs. For details on this sample see Appendix 2). Finally, I only include investments made by U.S. VC firms in U.S. companies. After these eliminations, the data contain 116,574 transactions.

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] I restrict my sample to primary assignments of utility patents (99\%) awarded to US companies from 1976 onwards.

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 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 small number of matches between

[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.

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.

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.[3] 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

[3]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.

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 (salf-citations) and citations made by other assignees (no self citations). Panel B, shows this is also true for the generality measure. Finally, 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			