Venture Capital and the Diffusion of Knowledge

Juanita González-Uribe*

Columbia GSB

Preliminary and incomplete

This version: May 29, 2012

Abstract

I consider how Venture Capitalists (hereafter VCs) affect innovation by facilitating the diffusion of knowledge. I show that after a company is first financed by a VC, its ideas are more likely to be used by other agents in the economy, as measured by the number of citations received by its patents. To control for patent heterogeneity, my broad empirical strategy compares the likelihood of a citation to the same patent before and after the issuing company is first financed by a VC, relative to other patents in the same technology- class and applied for the same year. To address concerns of non-random timing of VC financing I implement an instrumental variables approach that uses variations in the size of local and state pension funds' assets as an exogenous determinant of VC selection. My findings suggest that knowledge spillovers from VC-backed companies constitute an important part of the effect of VCs on innovation.

*Columbia Graduate School of Business, e-mail: jgonzalez14@gsb.columbia.edu. I am indebted to my advisor Morten Sorensen for his invaluable guidance, encouragement and numerous discussions. I am deeply grateful to Bruce Kogut, William Greene, Daniel Paravisini and Daniel Wolfenzon for their suggestions. Finally, I am also grateful for funding from the Ewing Marion Kauffman Foundation.

In this paper, I ask whether Venture Capitalists (hereafter VCs) affect innovation by facilitating the diffusion of knowledge. I show that after a company is first financed by a VC, its ideas are more likely to be used by other agents in the economy, as measured by the number of citations received by its patents.

Existing work provides evidence of a causal effect of VCs on patent production at the industry-level (Kortum and Lerner, 2000; Mollica and Zingales, 2007; Popov and Rosenboom, 2009; Hirukawa and Ueda, 2011). Specifically, Kortum and Lerner (2000) estimate that VCs accounted for 14% of U.S. patent production from 1983 to 1998. However, only 2% of US patentees are VC-backed companies, and less than 6% of US patent production can be attributed to them.[1][2] The question of how such a relatively few number of VC investments can have this impact on innovation, remains. In addition, worldwide public efforts to increase innovation by stimulating VC activity have mostly failed.[3] A better understanding of how VCs affect innovation can be important to design more effective innovation policy.

In this paper I argue that one mechanism through which VCs affect innovation is by facilitating the diffusion of knowledge across agents. The existence of knowledge flows induced by VC activity, can help explain the large impact of VCs on innovation, in spite of the relatively few patents that can be attributed to VC-backed companies. To measure knowledge flows, I follow the innovation literature and use data on forward citations to patents (Jaffe, 1986; Jaffe et al., 1993; Jaffe and Trajtenberg, 2002; Hall et al, 2001).[4]

[1] VCs are the dominant form of equity financing in the U.S. for privately held high-technology businesses. However, 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 crowd funding in privately-held companies.

[2] The contribution falls to 3% once I exclude well-known outliers such as Intel, Cisco, Microsoft and Apple. For details on the data sources and construction of sample see Section 2.

[3] See Lerner (2009) for a discussion at length on this topic.

[4] Inventors have the legal duty to disclose relevant prior art when filing their patent applications. Thus, citations are informative of knowledge links between patented innovations. Because the interest is in the transfer of knowledge across parties, I drop all self-citations: citations made by the same patentee to its own

Prior studies have pointed out that spillovers from private equity -backed companies are important, but have been unable to quantify them due to the lack of comprehensive data (Bernstein et al., 2011). In this paper, I overcome that limitation by constructing a dataset at the patent-level of all patents granted in the U.S. to VC-backed companies, together with the annual citations received by these patents. The novelty of my approach is to link VCs to specific innovations, and measure the level, and distribution across agents, of the knowledge spillovers generated by these innovations over time.

There are three empirical challenges in identifying the effect of VCs on knowledge diffusion using data on patent citations. The first challenge is unobserved heterogeneity. Patents issued by companies that eventually receive VC financing may be of high quality and receive more citations than the average patent. This could produce a correlation between VC financing and patent citations in the cross section, without the former really causing the latter. To overcome this concern, my methodology relies on prior patents: patents that were applied for at least two years before the issuing company is first financed by a VC. I estimate the effect of VCs on knowledge diffusion using within-patent changes in forward citations, following the VC financing event of the issuing companies.

The second challenge is the aggregate increase in patent citations over time. As the pace of patenting accelerates worldwide, the frequency of patent citations has increased (Lerner, 2011). This positive trend in citations could lead to an upward bias in the correlation between within-patent changes in citations and VC financing. To address this concern, I define a set of matching patents as follows. For each prior patent in my sample, I determine all U.S. patents that were filed the same year, assigned to the same technology class and were not assigned to a VC-backed company. Using data on forward citations to matching patents, I calculate the average citation intensity of the matching patents for every year after the application date. I then use this average intensity in my within-patent

prior work.

estimation to control for technology- and year- specific variations in citation patterns. Specifically, my broad empirical strategy estimates the increase in the likelihood of a forward citation to the same prior patent, relative to other patents in the same technology-class and application year, after the issuing company is first financed by a VC. To the extent that this within-patent estimator fully absorbs the time-invariant quality of patents and controls for technology-class wide changes in citation patterns, the estimated increase in the likelihood of a citation can be associated to the financing event of the issuing company, and not to heterogeneity in the quality of patents, or to changes in aggregate citation behavior. Using a conditional 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 the issuing company is first financed by a VC.

The third estimation challenge 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 (even within very narrow technological classes) will be more likely to be cited in the future, and invest based on that prediction. This would generate a positive correlation between VC financing and within-patent increases in relative citations, without the former really causing the latter.

To address this third identification challenge, I implement an instrumental variables methodology. To identify exogenous variation in the timing in which companies are selected by a VC, and similar in spirit to other papers in the VC literature (Mollica and Zingales, 2007; Bernstein et al. ,2011), I use variations in the size of public pension funds' assets in the home-state of companies. Public pension funds are among the largest sponsors of the VC industry and they are home biased in their private equity 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 idea behind this instrumental variables approach, is that in states and periods where pension pools are larger, domestic VC firms are more likely to raise capital and invest it locally.

A valid instrument must satisfy two requirements.[5] First, it must be correlated with the timing in which companies are first financed by a VC. Similar to Mollica and Zingales (2007), I show that variations in the size of state public pension funds are positively correlated with the value of VC investments in local new companies. Second, the instrument must satisfy the exclusion restriction, i.e., the instrument must be related to changes in the likelihood of forward citations to prior patents only through the VC financing event of the issuing companies. This condition is likely to be satisfied because pension policy is driven by broader socioeconomic considerations, rather than the investment opportunities in high-technology businesses inside the VC industry.

Nevertheless, one concern is that variations in pension funds' assets may be indicative of innovation opportunities within states. If this is the case, then variations in the size of state pension funds, and changes in relative citations to prior patents may be correlated via a state-effect, i.e., the exclusion restriction will be violated. I address this concern in two ways. First, I define relative citations at the state-level. Under this approach, my estimation compares the change in the likelihood of a forward citation to a prior patent post-financing, to the change in the likelihood of a forward citation to matching patents that have been issued in the same state. This robustness check is useful because if in fact variations in pension fund size reflect a state-effect, this state-effect shouldn't affect differentially citations to prior patents, and citations to local matching patents.

The second way in which I address the potential violation to the exclusion restriction is by using as dependent variable out-state citations*:* citations to patents from assignees that are located in a different state. This second robustness check compares the change in the

[5]In the case of heterogeneous treatment effects, monotonicity is also required to estimate a local average treatment effect (Imbens and Angrist, 2004). The monotonicity condition states that there should be no patent whose issuing company is financed by a VC when the availability of capital for VCs is low, but it is not chosen when there is excess capital, which seems reasonable.

likelihood of a forward out-state citation to a prior patent post -financing, to the change in the likelihood of a forward out-state citation to matching local patents. Similar to the first robustness check, the exclusion restriction is unlikely to be violated in this setting because if the size of local and state pension funds is correlated to changes in innovation within a state, which in turn affects out-state citations, such a change should affect equally all patents issued in the same state, and is therefore unlikely to affect relative citations measures at the state level.

Using a Generalized Method of Moments (GMM-IVs) approach to estimate a conditional fixed-effects Poisson model with endogenous regressors, I find that results continue to hold. Compared to the QMLE estimator, the GMM-IVs estimator is larger, although the difference in the two coefficients is not statistically significant across all different specifications. Since my GMM-IVs approach estimates the Local Average Treatment Effect (LATE), the GMM-IVs estimates are based on those patents whose companies are selected by a VC only because there was an excess of funding capital for VC investors as measured by local and state pension funds' assets. The interpretation of the positive difference between the GMM-IVs approach and the QMLE estimator, is that an abundance of capital allows investors to experiment and invest in patents (compliers) that are less well known, and which can presumably be more sensitive to VC financing, than VCs' usual investments (always-takers). This interpretation is consistent with recent findings by Nanda and Rhodes-Kropf (2011) regarding changes in the investment behavior of VCs during hot markets.

Having shown suggestive evidence that forward citations to prior patents causally increase after companies are first financed by a VC, I turn to analyzing the distribution of forward citations by type of citing patentee. For each company that issued a prior patent, I define portfolio-connections as the set of VC-backed companies that are financed by the same VC (s) as the target. I then classify forward citations to prior patents into three groups: non VC-backed, if the citing patentee has not been backed by a VC at the time of the citation, portfolio- connected, if the citing patentee is a portfolio connection of the target, and non portfolio-connected, otherwise. Using matching patents I define average citation intensities by type of citing patentee, and estimate a model that compares the changes in the likelihood of forward citations to prior patents after the VC financing event, by type of citing patentee. There are two findings. First, the increase in citations to prior patents following the VC financing event is not concentrated within the VC industry, but diffuses more generally to other patentees outside the VC industry. Second, citations from portfolio-connections see the largest increase post-financing.

This paper contributes to the literature that considers the role of finance in the innovation process (Kortum and Lerner, 2000; Mollica and Zingales, 2007; Popov and Rosenboom, 2009; Hirukawa and Ueda, 2011; Lerner et al. 2010, Bernstein, 2011, Nanda and Rhodes-Kropf, 2011; Bernstein, 2012, Seru, 2008). Although the innovation literature has highlighted the impediments of knowledge flow across institutions (Zander, 1991; Almeida and Kogut, 1993; Gomes-Caceres et al., 2006 and Azoulay et al., 2001), it has mostly focused on the geographical concentration of these flows (Jaffe et al., 1998, Jaffe and Trajtenberg, 1996, Jaffe and Trajtenberg, 1999, Jaffe et al., 1993, 2002). The question of whether financial intermediaries can facilitate knowledge diffusion, has not been systematically addressed before. One exception is Lindsey (2008), who examined the role of VCs in facilitating formal knowledge flows across companies, as measured by research alliances, inside the VC industry. I focus on informal knowledge flows, as measured by patent citations, and examine how VC financing affects the diffusion of a company's knowledge, both inside and outside the VC industry.

This paper finds suggestive evidence that one mechanism through which VCs affect innovation is by facilitating the diffusion of knowledge created by their targets. My findings contribute to explaining how VCs can have a large impact on innovation in spite of the relatively few companies that are VC-backed, and the patents that can be attributed to them. A conservative back of the envelope calculation of my findings, estimates that for every patent issued by a VC-backed company, an additional patent in the economy exploits the same knowledge to generate a new product. After we include this effect, the share of patents that can be attributed to VCs increases from 6% to 12%, a similar number to Kortum and Lerner (2000)'s estimate (14%). The main implication of my paper is that to quantify the effect that VCs have on innovation one must account for knowledge spillovers of VC activity.

My analysis is not without limitations. One unaddressed topic is whether the increase in forward citations to patents is beneficial for target companies. Although there is evidence that citations to patents are associated to value, there is a clear trade-off between exposure and imitation. On the one hand, more citations may indicate that the patent increases in value and that the company may be able to extract those rents via licenses, royalties and sales of patents. On the other hand, more citations can also indicate more imitation and increased competition, which can negatively affect the company. My paper only shows that the likelihood of a citation increases, but I make no claim regarding the implications for the value of the target company.

Another question that remains unaddressed is what are the mechanisms through which VCs generate these knowledge spillovers. There are many ways in which VCs can affect knowledge diffusion, for instance, by increasing the exposure of their targets. The presence of VCs as investors in a company may communicate unobserved qualities about the company to the market (Megginson and Weiss, 1991), including the quality of its IP. Financial resources provided by VCs can help their companies exploit the option value embedded in disembodied patents (Bloom and Van Reenen, 2002). Finally, VCs can help their targets recruit top executives (Hellman and Puri, 2002,). Disentangling among these channels is out of the scope of this paper, and represents an 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 main empirical strategy and present base results. In Section 3, I address concerns of non-random timing in the selection of companies by VCs. I examine the distribution of citations to prior patents in Section 4. Finally, Section 5 concludes.

1 Data Description and Summary Statistics

1.1 Investments by U.S. Venture Capital firms

My starting point is the universe of investments registered in SDC's VentureXpert that closed between January 1976 and December 2008. I eliminate three types of investments. First, I eliminate investments by institutions 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). Second, I drop data on funds by venture capital firms that are not focused on venture capital, such as buyout funds and funds of funds. Finally, I only include investments made by U.S. VC firms in U.S. companies. After these eliminations, the data contain 124,466 investments in 21,887 U.S. based companies by 1,967 VC firms in the U.S., from January 1976 through December 2008.

1.2 Capturing patent data

I restrict my sample to companies with at least one successful patent application. I match companies involved in VC investments to their patenting records based on their 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. I restrict my sample to primary assignments of utility patents (99%) awarded to US companies. The data consist of 1,589,174 patents assigned to 128,346 U.S. companies through December 2008.

In order to search the HBS database for each of the VC-backed companies, I strip company names of punctuation, capitalization and common acronyms. I then match the samples on the normalized company 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)[6]. 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.

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

1.3 Matched Sample

The final matched sample consists of 168,460 patents awarded from January 1976 through December 2008 to 5,346 companies that were financed by at least one of 1,424 VC firms between January 1978 and December 2009 (43,837 investments). The small number of matches between the two data sets likely reflects two facts. First, SDC includes data on all companies that received VC financing, including those that were not ultimately successful. Second, many inventions are not patented. The propensity to patent depends on the industry and type of invention, and an important fraction of sectors where VCs invest, do not tend to use IP protection in the form of patents (for example, internet and media companies).

Table 1 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.[7] 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, investments, VC-backed companies and VC firms, by state. As it is common in the VC literature, California, Texas and

[7] 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.

Massachusetts are overrepresented in the sample.

Panel C shows the distribution of type of investments by VC firms on companies that patent. The types of investments include traditional VC investments such as: Bridge Loans, Early Stage, Expansion, Later Stage and Seed. However, they also include other non traditional VC investments such as: Acquisitions, Leveraged Buyouts (LBOs), Management Buyouts (MBOs), Open Market Purchases (OMPs), Private Investments in Public Equity (PIPEs), Recapitalizations or Turnarounds, Secondary Buyouts and Secondary Purchases. Because the interest is in the effect that VCs have on the diffusion of knowledge of private companies, I exclude from the main analysis sample information on non-traditional investments by VCs.

1.4 Analysis sample

I restrict my matched sample to patents that were applied for at least two years before the issuing company was first financed by a VC, and to companies for which the first investment by a VC was in one of the following stages: Seed, Early, Expansion, Late or Bridge Loan. I refer to these patents as prior patents throughout this paper.[8]

The analysis sample consists of 2,336 prior patents that were applied for by 752 VC-backed companies at least two years before they are first financed by a VC. Table 2, Panel A, describes the distribution of the VC deals (first time VC financing of companies with prior patents) and patent filings (both, application and grant years) over time. Panel B summarizes the number of companies, patents, and patent-year observations, by state.

Panels C and D, summarize characteristics of the issuing companies of prior patents. Panel C, illustrates the stage of the company at the time of the first VC investment. Panel

[8]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, although the sample size diminishes. I report results using the application date due to considerations of statistical power.

D, shows the distribution of VC-backed companies with prior patents according to industry classification. The sample is concentrated in Computer Software, Medical Health and Semiconductors. Panel E, describes the status of the investment by 2010. Finally, Panel F, shows the distribution of prior patents by age.

1.5 Citations to prior patents

Table 2, Panel B presents average annual citations to prior patents by the issuing company's home-state, classified by type of citing patentee. I define three types of citing patentees: non VC-backed, portfolio-connections and non portfolio-connections. I classify a citing patentee as non VC-backed if at the time of the citation it has never been backed by a VC, and as a VC-backed patentee otherwise. Non VC-backed patentees include companies outside the VC industry, as well as patentees who are not companies, such as, government institutions, universities and individual inventors. Because there are many more non VC-backed patentees than VC-backed patentees, annual average citations from non VC-backed patentees are always higher than annual average citations from VC-backed companies.

VC-backed citing patentees are further classified into two sub-groups: portfolio -connected, if they share a common VC investor with the issuing company at the time of the first financing event, and as non portfolio-connected, otherwise. Table 3 shows the average number of portfolio connections for VC-backed companies with prior patents, and the distribution by year in which the companies are first financed by a VC. On average, VC-backed companies with prior patents have 16.6 portfolio -connections. Column (4) shows the average number of citations received by prior patents from portfolio -connections, by year in which companies are first financed by a VC.

1.6 Local and State Pension Funds' Data

Data on the size of local and state pension funds' assets comes from the Census Bureau State and Local Government Public - Employee Retirement Systems Survey. The data is available from 1993 to 2008. Table 4 shows summary statistics for the sample of prior patents restricted to the 1993-2008 period. Panel A, shows the distribution of application and grant years of prior patents, as well as the distribution of VC financing events of the issuing companies. The restricted sample consists of 1,170 prior patents awarded to 434 companies. The first two columns of Panel B show the distribution of local and state public pension funds' assets by state. The last four columns of Panel B also includes the distribution of companies, patents, patent-year observations, and average annual citations to prior patents by state. Similar to the full sample, California is overrepresented in the restricted sample.

2 Empirical Strategy

In this section I discuss the use of patent citations to measure knowledge flows. I then describe the empirical strategy I use to estimate the effect of VC financing on knowledge diffusion.

2.1 Measuring 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.[9] Although citations are not a perfect measure of knowledge flows, for example,

[9] Each patent record includes citations of prior inventions on which the current patent builds. By law, inventors are obliged to include all relevant prior art in their application, including patents, academic papers,

many are added by patent examiners rather than by the inventors themselves, prior research finds they correlate well with actual knowledge flow (Jaffe et al., 2002; Duguet and MacGarvie, 2005; Roach and Cohen, 2010).[10] Thus, the consensus in the literature is that citations are informative (although noisily so) of links between patented innovations, and can be interpreted as "paper trail" evidence of spillovers.[11] Because the interest is in the diffusion of knowledge across patentees, I exclude from my data within-patentee citations.

I start the analysis by performing a simple non parametric test of the effect of VCs on knowledge diffusion. I estimate the change in the likelihood of a citation to a prior patent after the issuing company is first financed by a VC. In Table 5, Column (5), I report Incidence Rate Ratios (IRRs), which corresponds to the ratio between the mean number of citations to prior patents post-financing, to the mean number of citations pre-financing. The IRR of 1.63 in the first row, implies that the likelihood of a citation to a prior patent increases 63% after the issuing company is first financed by a VC.

---

and technical disclosures. Citations serve a useful legal purpose; they help determine the scope of the property rights awarded to the patent. That is, if patent B cites patent A, it implies that patent A represents a piece of previously existing knowledge upon which patent B builds, and over which B cannot have a claim.

[10] There are four parties involved in a patent application: the inventor, the assignee, the lawyer and the patent officer. Inventors usually hire lawyers to process the application and later assign the intellectual property right to an assignee who in most cases is her/his employer. When the application is submitted to the USPTO an official reviewer is assigned to the case. The review process often lasts a number of years, and during which there are negotiations between the examiner and the lawyers (or directly the inventor) over the claims included in the application. Claims are afforded depending on existing prior art, and for this reason, both the reviewer and the lawyer, are also obliged to include in the application all relevant citations to prior art that were not included by the inventor. After the negotiation the patent officer decides on which claims should be granted to the inventor.

[11] A nuanced view is that citations to prior patents increase after the issuing company is financed by a VC without having any effect on knowledge diffusion. For instance, potential targets may strategically use patent citations to attract the attention of potential investors. There is no evidence of this in the literature. In unreported results, I use investments by VCs in public companies as an informal test (See Table 1). I don't observe an increase in citations after public companies are financed by VCs. This is reassuring because public companies have no need for further diffusion, however, if indeed citations only increase strategically one should still observe an increase post financing for the prior patents of these companies. Another nuanced view, is that agents increase citations to the IP of a company, after the company is first financed by a VC, out of litigation concerns and not because of knowledge diffusion. This concern is 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. This means that citations cannot be used as protection against future law suits for infringement, 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 I4I Limited Partnership, 2010).

Citation rates and patent filing propensity vary across time and across technology classes.[12] These variations may stem from changes in the rates of diffusion of technologies over time or from changes in the patent system. Therefore, information on forward citations to patents is meaningful only when used comparatively. Following Hall et al. (2001), for every prior patent, I define matching patents as follows. I determine all U.S. patents filed the same year and assigned to the same USPTO technology class that are not financed by VCs.[13] I estimate the average number of forward citations to matching patents every year since application date. In Table 5, Column (6), I report Relative Incidence Rates (RIRRs) defined as the ratio between the IRR of prior patents and the IRR of matching patents. The RIRR of 1.33 in the first row, implies that the likelihood of a citation to a prior patent increases by 33% after the issuing company is first financed by a VC, relative to the increase in the likelihood of a citation to a matching patent.

To control for potential overdispersion and clustering of standard errors, I turn to a regression analysis. As it is typical in this literature, I find substantial overdispersion in the citation counts. Overdispersed count data can be appropriately handled with a Negative Binomial model. In addition, since all patents issued by the same company are subject to company specific shocks, I construct robust standard errors by allowing correlation of the error term at the issuing company level. Finally, note that all reported coefficients are incidence rates. An incidence rate greater than one corresponds to a positive effect of the characteristic on the intensity. 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 are different from one.

[12] When patents are granted by the USPTO they are assigned to a technology- class. There are approximately 800 technological classes in my dataset. The classification is important because it helps patent examiners review existing art. Examples of the technological classes are: Dynamic Optical Information Storage or Retrieval (technology-class: 720) and Plastic Article of Earthenware Shaping or Treating Apparatus (technology-class: 425).

[13] 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 due to the review process in the Patent Office.

Table 6, Column (1) contains results from a pooled regression using the Negative Binomial model of annual citation counts to prior patents on VC_{pt}, a dummy that equals one after the issuing company of the patent is first financed by a VC. The coefficient for VC_{pt} in the first column implies that the likelihood of a citation to a prior patent significantly increases by 62.7% after the VC financing event of the issuing company.

To control for changes in citations' behavior over time, I control for the baseline citation intensities, using the matching patents described above. This is implemented as follows. For every prior patent and every time period t, I calculate the average citation intensity of the matching patents as

$$\delta_t^s = \ln\left(\frac{Total\ \mathrm{Ci}tations}{Number\ of\ Matching\ Patents}\right) \tag{1}$$

where $Total\ \mathrm{Ci}tations$ is the the number of citations received by all matching patents at time t. By including this average intensity 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. It is similar to including time-fixed effects (although it is more stringent because the time fixed effects are at the technology-class and application-year level). This technique has two additional benefits. First it facilitates the interpretation of the estimated coefficient. The resulting estimates reflect the relative citation intensities to patents in my sample compared to patents in the same technology-class and applied for the same year. Second, 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., 2010).

Table 6, Column (2) presents the estimates of a pooled regression using the Negative Binomial Model of annual relative citation counts to prior patents on VC_{pt}. To illustrate, the coefficient for VC_{pt} in the second column implies that the likelihood of a citation to a prior patent is estimated to increase by 34.6% after the issuing company is first financed by a VC, relative to patents in the same technology-class and application year. Results from the first two columns in Table 6, imply that the non parametric results from Table 5 are robust to overdispersion and clustering.

2.2 Main Empirical Specification

To control for patent heterogeneity, I turn to a fixed-effects regression. My estimating equation relates citations to prior patent p to the VC financing event of the issuing company:

$$E\left[Y_{pt}^{s}|\delta_{t}^{s},\alpha_{p}\right]=\exp\left(\delta_{t}^{s}+\alpha_{p}+\beta VC_{pt}\right) \tag{2}$$

where Y_{pt}^{s} represents the number of citations to patent p, classified in technology-class and application year s, at time t. δ_{t}^{s} corresponds to the average citation intensity for patents similar to patent p at time t, and as above, VC_{pt} is an indicator variable that equals 1 after the issuing company of patent p is financed by a VC for the first time. I include a full set of patent fixed-effects, α_{p}, that account for cross sectional unobserved heterogeneity, and that is consistent with my approach in analyzing changes in p's citations following the financing event of the issuing company. Note that specification (2) does not allow me to include individual year controls, since including the citations baseline at the technology-class and application-year level in the specification, also removes any aggregate year variation.

2.2.1 Econometric Considerations

I estimate equation (2) by conditional quasi-maximum likelihood (QMLE) based on the fixed-effects Poisson model developed by Hausman et al. (1984). Because the Poisson model is in the linear exponential family, the coefficient estimates remain consistent as long as the mean of the dependent variable is correctly specified (Gourieroux et. al, 1984). In fact, the Hausman et al. estimator can be used for any nonnegative dependent variables, whether integer or continuous (Santos Silva and Tenreyro, 2006), as long as the variance/covariance matrix is computed using the outer product of the gradient vector (and therefore does not rely on the Poisson variance assumption). Further, QMLE standard errors are robust to arbitrary patterns of serial correlation (Wooldridge, 1997). The practical implication is that during the QMLE estimation I do not use the (Poisson) assumption that the mean and the variance are equal, or the independence assumption on the observations. I need only assume equation (2), which is not too restrictive given the non negative and discrete nature of the citation data, and address overdispersion concerns by making inference using robust standard errors by clustering at the company level.

2.2.2 Results

Table 6, Column (3), presents the main QMLE results, and shows that the positive association of VC financing on citations to prior patents documented in columns (1) and (2) is robust to including patent-fixed effects. The coefficient of 1.189 for VC_{pt} implies that citations to the same prior patent, relative to patents in the same technology-class and applied for the same year, are estimated to increase by 18.9% after the issuing company is first financed by a VC. The identification strategy relies upon the assumption that the relative changes in citation rates over time for patents within the same technology-class and applied for the same year, are comparable.

2.2.3 Extensions

In unreported results, I repeat the base QMLE specifications reported in Table 6 with the sample restricted to states different from California, Massachusetts and Texas. The effect remains unchanged. Moreover, I find that the effect is not statistically different. I also restrict the sample to the pre - and post - dot com periods. Results hold for both sub-samples. Finally, 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 of the effect using the QMLE approach. I implement this idea by restricting the observations to a [-2,5] window around the financing event, and including time-dummies in the QMLE specification. Figure 1 plots the estimated coefficients and confidence intervals from this specification. The figure shows that the likelihood of a citation to a prior patent increases only after the issuing company receives VC financing, relative to other patents in the same technology-class and applied for the same year. In other words, before the financing event the likelihood of a citation to a prior patents is not statistically different from the likelihood of a citation to other patents in the same technology class and application-year. However, after the financing event, the likelihood of a citation to a prior patent significantly increases. In addition, the figure shows that the increase in the likelihood of a citation does not seem to be driven by a pre-existing trend in the evolution of relative citations to prior patents.

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

The remaining identification concern of non random timing in VC selection can be modeled as a multiplicative unobserved effect on citation counts. Specifically, equation (2) can be modified to include the endogeneity, by including unobserved time-varying heterogeneity at the patent level as follows:

$$E\left[Y_{pt}^{sa}|\delta_t^s, \alpha_p, \varepsilon_{pt}^s\right] = \exp\left(\delta_t^s + \alpha_p + \beta VC_{pt}\right)\varepsilon_{pt}^{sa} \tag{3}$$

where ε_{pt}^{sa} captures time-varying shocks to forward citations to prior patents that are not captured at the technology-class and application year level.

Consistent estimation of this model is complicated by the patent fixed-effects. Following Wooldridge (1997), I use a quasi-differencing transformation to remove the patent fixed effects and construct orthogonality conditions that allow consistent estimation of β using the Generalized Method of Moments (GMM-IVs).

Specifically, given an exogenous determinant of VC selection, I_{pt}^{sa}, the following moment conditions can be used for consistent estimation of the model:

$$E\left[\frac{Y_{pt}^{sa}}{\exp\left[x_{pt}^{sa} - \mu_x\right]B} - \frac{Y_{pt+1}^{sa}}{\exp\left[x_{pt+1}^{sa} - \mu_x\right]B} \mid I_{pt}^{sa}\right] = 0 \tag{4}$$

where $x_{pt}^{sa} = VC_{pt} + \delta_t^s$, $\mu_x = (NT)^{-1}\sum\sum x_{pt}^s$, and $B = [\beta_2 \ 1]$ (the coefficient of the baseline is again offset to 1).

3.1 Local and state pension funds' assets

Similar in spirit to other papers in the literature, I use variations in the size of local public pension funds' assets in a company's home-state as an exogenous determinant for VC selection. The basic idea is that in states and periods were 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 my empirical strategy I use the size of local and state pension funds lagged by 1 year to instrument for the timing in which companies are financed by VCs.[14]

Table 7 illustrates the first condition for a valid instrument, a correlation between the size of local and state pension funds and the timing in which local companies are selected by VCs. I run an Ordinary Least Squares regression on the total value of investments in new companies by VCs in a state, and the size of local and state pension funds in that state. I include time-fixed effects and state-fixed effects. Specifically, I run the following specification,

$$N_{at} = \gamma_a + \phi F_{at-1} + \zeta_t + \xi_a + \epsilon_{pt} \tag{5}$$

where N_{at} is the value of investments in new companies made by VC firms in state a at time t. F_{at-1} is the size of local pension funds' assets deflated by the PPI in state a. I compute robust standard errors by clustering at the state level. Table 7 shows a positive and significant relation between the value of VC investments in new companies in a state, 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.

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

The second condition for a valid instrument, the exclusion restriction, requires that changes in pension funds are independent from the unobserved time varying-heterogeneity, ε_{pt}^{sa}, in specification (3). While this is difficult to establish empirically, pension funds primarily change as a result of pension reforms. 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. However, it is possible that variations in pension funds' assets may be indicative of innovation opportunities within a state. If this is the case, then the pension fund size instrument and changes in relative citations to prior patents may be correlated via a state effect, i.e., the exclusion restriction will be violated. I address this concern in two ways. First, I define relative citations at the state-level. Second, I also restrict the dependent variable to out-state citations. The exclusion restriction is unlikely to be violated in either of 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 citations measures at the state level.

3.2 GMM-IVs main results

Table 8 presents the GMM-IVs results based on the moment conditions (4), using the size of local and state pension funds, demeaned by state and time, and lagged by one year as an instrument, $I_{pt}^{sa} \equiv F_{at-1}$.[15] Column (3) presents the reduced form results, obtained by substituting the endogenous VC_{pt} variable with F_{at-1}. There is a positive correlation between variations in local and state pension funds' assets and relative citations to prior patents.

Column (4) presents the GMM-IVs estimate for β of 1.805. After a company is first

[15] In unreported results I use the size of local and state pension funds', as well as deviations from time and state, separately, as alternative instruments. Results are quantitatively similar.

financed by a VC, the likelihood of a citation to the same prior patent is estimated to increase by 80.5%, relative to patents issued in the same technology-class and applied for the same year. To facilitate the comparison between the QMLE and GMM-IV estimates, in Column (2), I provide the QMLE estimates using the restricted sample of the GMM-IVs approach.[16] 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. As a robustness check, in unreported results I exclude California from the GMM-IVs specification. Results continue to hold and are quantitatively similar.

My GMM-IVs approach estimates the effect of VC financing on forward citations on the prior patents that respond to the instrument, and as such, the GMM-IVs estimates are only representative for those patents whose companies ended up being selected by a VC because there was a higher availability of capital. In other words, my identification strategy estimates the Local Treatment Effect (LATE) of VC financing on *compliers (*patents that were selected by a VC only because the assets of local pension funds were large).[17] The higher GMM-IVs estimate suggests that the marginal patent from the group of *compliers* is more sensitive to the VC effect than the patent in the *always-takers* group (patents whose treatment status doesn't change because of the instrument but were going to be treated any way). This result is reasonable as one would expect that when there is more availability of capital for VCs, the range of investments from which VCs choose their targets widens to include patents with lower diffusion prospects, and that may constitute a riskier investment. The intuition is that patents that are *always takers* are presumably already well known, and that is why they end up selected by VCs, whereas patents that are

[16]The sample of the IVs approach is different from the original analysis sample on two accounts. First, data on the size of the assets of local and state pension funds is only available for the 1993-2008 period. See Table 2. Second, the moment conditions used in the GMM approach differentiate out the fixed effects and therefore drop observations for the last period.

[17]In fact, any instrumental variables estimator uses only the information of the groups of agents that respond to the instrument or *compliers* (Imbens and Angrist, 1994).

compliers have a lower expected diffusion rate, and are more sensitive to the effect of VC selection.

3.3 Robustness Checks

Tables 9 and 10 summarize results from the robustness checks used to address concerns of potential violations of the exclusion restriction. Table 9 replicates the GMM-IVs approach using relative citations at the state-level. Table 10 replicates the GMM-IVs approach using relative citations at the state level and using as dependent variable out-state citations. The effect of VC financing on relative citations is significant for all specifications.[18]

The GMM-IVs results provide suggestive evidence that after companies are first financed by a VC, the likelihood of a citation to their prior patents causally increases. The main implication of this result is that the effect of VCs on innovation is not limited to their targets, but includes externalities on the innovation behavior of other agents.

4 Distribution of forward citations to prior patents

Having shown suggestive evidence that citations to prior patents causally increase after the issuing companies are first financed by a VC, I turn to analyzing the distribution of forward citations by type of citing patentee. As explained in Section 2, I define three types of citing patentees: non VC-backed, portfolio-connections and non portfolio-connections

Table 5, shows annual average citations to prior patents, before and after the first VC

[18] Note that the difference in observations from Tables 8, 9 and 10, 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.

financing event of the issuing company, by type of citing patentee. The likelihood of a citation for all types of citing patentees increases post-financing. For non VC-backed patentees, the IRR is 1.571, this means that after the financing event, the likelihood of a citation from a non VC-backed patentee increases by 57.1%. In contrast, the IRRs for citations from non portfolio-connections and portfolio-connections are respectively, 2.84 and 4.05. This implies that post-financing, the likelihood of a citation to a prior patent from a non portfolio-connection increases 1.87 times the increase in the likelihood of a citation from a non VC-backed patentee. For portfolio-connections, the likelihood of a citation increases 2.67 times the increase in the likelihood of a citation from a non VC-backed patentee.

To control for potential overdispersion and clustering of observations at the patent level, I run a Negative Binomial model with observations at the patent, type of citing patentee and year level. Specifically, I estimate the following specification,

$$E\left[Y_{pCt}\right] = \exp\left(\beta_0 D_{NVC} + \beta_1 D_{NP} + \beta_2 D_P + \sum_C \beta_C D_C * VC_{pt} + \varepsilon_{pCt}\right) \tag{6}$$

where Y_{pCt} are forward citations at time t, to patent p, from patents issued by type of patentee C, where $C \in \{NVC,\, NP,\, P\}$. NVC, NP and P stand for non VC-backed patentees, non portfolio-connections and portfolio-connections, respectively. D_{NVC} is a dummy that equals one when $C = NVC$, D_{NP} is a dummy that equals one when $C = NP$, and D_P is a dummy that equals one when $C = P$. Finally, ε_{pCt} is an $i.i.d$ random variable with mean zero that captures unobserved factors (overdispersion), and VC_{pt} is a dummy that equals one after the issuing company is first financed by a VC. Standard errors are clustered at the patent level.

Column (1) of Table 11, shows the estimated coefficients using specification (6). The coefficients on the type of patentee dummies can be interpreted as average annual citation rates pre-financing by type of citing patentee. For instance, the coefficient for D_{NVC} indicates that the average annual citations from non VC-backed patentees pre financing is 0.585.

The coefficient on the interactions between the type of patentee dummies and VC_{pt}, summarize the estimated IRR by type of citing patentee. For instance, the coefficient for $D_{NVC} * VC_{pt}$ implies that the likelihood of a citation to prior patents from a non VC-backed patentee increases by 51.7% after the financing event of the issuing companies. Similarly, the coefficient on $D_{NP} * VC_{pt}$ implies that post-financing, the likelihood of a citation from a non portfolio-connection increases by 183.9%. Finally, the coefficient on $D_P * VC_{pt}$ implies that post-financing, the likelihood of a citation from a portfolio-connection increases by 305.2%, or 2.671 times the increase in the likelihood of a citation from a non VC-backed patentee.

I am interested in testing whether after companies are first financed by a VC, the increase in the likelihood of a citation from a portfolio-connection and from a non portfolio-connection, is higher than the increase in the likelihood of a citation from a non VC-backed patentee. The first row of Panel B in Table 11 shows the estimated ratio between the IRR of citations from non portfolio-connections to the IRR of citations from non VC-backed patentees. The estimated ratio is 1.87 and is statistically different from 1[19]. The ratio between the IRR of citations from portfolio-connections and the IRR of citations from non VC-backed patentees is 2.67, and is also statistically different from 1.

To control for changes in the citation behavior, and in the technological composition of VC-backed companies over time, I control for the baseline citation intensities by type of citing patentee, using the matching patents described in Section 3. This is implemented as

[19] The standard error for the ratio of the IRR of citations from non portfolio-connections ($\exp(\beta_{NP})$) to citations from non VC-backed patentees ($\exp(\beta_{NVC})$) is calculated as: $\frac{\exp(\beta_{NP})}{\exp(\beta_{NVC})} * \text{sqrt}\left(\left(\frac{SE(\exp(\beta_{NP})}{\exp(\beta_{NP})}\right)^2 - \left(\frac{SE(\exp(\beta_{NVC})}{\exp(\beta_{NVC})}\right)^2\right)$.

follows. For every prior patent and every time period t, I calculate the average citation intensities of the matching patents by type of citing patentee as:

$$\delta^{s}_{C\ t} = \ln\left(\frac{Total\ \text{Ci}tations\ from\ type\ C}{Number\ of\ Matching\ Patents}\right) \tag{7}$$

where the numerator is the total number of citations from type of patentee C received by all matching patents at time t, where $C \in \{NVC,\ NP,\ P\}$. For example, $\delta^{s}_{P\ t}$ corresponds to the logarithm of the average number of citations to matching patents at t from the portfolio connections of prior patents. By including the different types of average intensities 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 citing patentee. This technique is similar to (although more stringent than) including type of citing patentee cross time fixed effects.

Column (2) in Table 11 shows the estimated IRRs using a Negative Binomial Model and including the average citation intensities in the estimation. The coefficient on $D_{NVC} * VC_{pt}$ implies that post-financing, the likelihood of a citation from a VC-backed patentee increases by 30.6% after the financing event, relative to matching patents. The coefficient on $D_{NP} * VC_{pt}$ implies that post-financing, the likelihood of a citation from a non portfolio-connection increases 85.9% post financing, relative to matching patents. Finally, the coefficient on $D_P * VC_{pt}$ implies that the likelihood of a citation from a portfolio-connection is estimated to increase 173.9%, relative to matching patents. Note that the effect for portfolio-connections, although economically significant, is no longer statistically significant. This lack of significance could be associated to the decrease in power due to the many observations with zero average citation intensities, particularly for portfolio-connections. To see this, note the decrease in observations from Column (1) to Column (2). [20]

[20] Note that since the estimation includes the logarithm of average citation intensities with a coefficient

Panel B tests whether the estimated increase in the likelihood of a citation from a portfolio- and from a non portfolio-connection is greater than for a non VC-backed patentee. Consistent with Panel A, I find that both types of VC-backed patentees increase their citations more than non VC-backed patentees. However, the effect is not statistically significant for portfolio-connections.

Column (3) estimates specification (6) including cross patent and type of citing patentee fixed-effects, which control for the time invariant compatibility of each patent to each type of citing patentee. The model with cross patent and type of citing patentee fixed-effects is estimated using the QMLE approach explained in Section 3. Column (3) shows that after controlling for the time -invariant compatibility of each patent to each type of citing patentee, the likelihood of a citation from a non VC-backed patentee is estimated to increase by 39.6% after the financing event. Similarly, for non portfolio connections and portfolio connections, the estimated increases are 122.1% and 189.8%, respectively. The interpretation is that after the VC financing event, the likelihood of a citation from a non portfolio-connection increases 1.591 times the increase in the likelihood of a citation from a non VC-backed patentee, after controlling for the time-invariant compatibility of the technology embedded in the prior patent to the technologies of each type of patentee. The IRR for portfolio-connections is interpreted analogously. Note the decrease in observations from Column (1) to Column (3), due to the implementation of the QMLE approach, which requires variation in the dependent variable within each patent-citing patentee group for estimation.

Finally, Column (4) shows the estimates from the QMLE estimation also including the average citation intensities. Column (4), Panel B, shows that the likelihood of a citation from a portfolio-connection is estimated to increase 2.352 times the increase in the likelihood of a citation from a non VC-backed patentee. This effect is economically and

fixed to one, observations with zero average citation intensity are necessarily dropped from the estimation.

statistically significant. Note that the ratio between the IRR of citations from non portfolio-connections, and the IRR of citations from non VC-backed patentees, shown in the first row of Panel B is very close to 1, and no longer statistically significant. This suggests that after controlling for the time-invariant compatibility of the patent to the technologies of each type of citing patentee, and for aggregate changes in citations, the increase in the likelihood of a citation from a non portfolio-connection is not statistically different from the increase in the likelihood of a citation from a non VC-backed patentee.

4.1 Extensions

The significantly higher increase in citations from portfolio connections estimated using the QMLE approach is consistent with prior research that has shown that companies that share a common VC- investor are more likely to share formal knowledge flows, as measured by research alliances (Lindsey, 2008). However, note that my results are consistent with two non-mutually exclusive explanations. First, VCs may select targets to exploit technological complementarities within their portfolios, which can result in the observed higher knowledge flows among companies backed-by the same VC. Second, VCs may also encourage interaction inside their portfolios post selection, which can also explain the effect. Disentangling the selection and treatment effects is out of the scope of this paper.[21]

In unreported results, I tested whether the higher increase in the likelihood of a citation from a portfolio-connection relative to a non VC-backed patentee post-financing, could be attributed to movement of inventors inside VC portfolios. I didn't find evidence that the higher increase in citations from portfolio-connections is exclusively due to inventor flows inside VC portfolios.

[21] In order to disentangle selection from treatment, I tried running the GMM-IVs approach in this setting. Unfortunately, the algorithm didn't converge using the one year lag in variations of local and state pension funds.

The QMLE approach in Column (4) does not include patent-time fixed effects that can control for patent-specific changes in citations, and allow identification to come from only the differential increase in the likelihood of a citation from each type of VC-backed companies, relative to non VC-backed patentees. Because of the many fixed effects that need to be estimated in this alternative specification, I tried estimating a version of this more saturated model by collapsing time into two periods and introducing patent-time dummies (Note that although the resulting dependent variable is no longer discrete, the QMLE can still be used, see Wooldridge (2007)). However, although it would be interesting to report these results, the QMLE didn't converge for this alternative version.[22]

5 Conclusion

In this paper I hypothesize that the presence of VCs as investors in a company can facilitate the diffusion of the company's ideas in the economy. To test this hypothesis I use data on patent citations as a proxy for the transfer of knowledge among inventors. There are two main findings. First, after a company is first financed by a VC, its ideas are more likely to be used by other agents in the economy. Second, these knowledge gains are not concentrated within the VC industry, but diffuse more generally.

Existing work provides increasing evidence of a VC -innovation channel: a causal effect of VCs on patent production at the industry-level. Investigating the mechanisms behind the ability of VCs to affect innovation has proven difficult given the lack of data that links VCs to innovations over time. In this paper I address this shortcoming by constructing a

[22] An alternative approach is to run a linear model. I don't think this approach is appropriate as the data generating process of citations data is clearly not linear. However, I still tried a linear approach, by estimating relative citation increases using a logarithmic transformation of the data. To avoid dropping observation with zeros (which are most of the observations) I added x to each observation, where x ranges between 0.1 and 1. I found that after the VC financing event, relative citations from VC-backed companies increase more than relative citations from Non VC-backed patentees. However, the significance of the result is not robust to different values of x.

sample of patents issued by VC-backed companies in the US, and measuring the level and distribution of knowledge flows generated by these innovations. My broad empirical strategy estimates the increase in the likelihood of a citation to the same patent after the issuing company is first financed by a VC, relative to the increase in the likelihood of a citation to patents in the same technology-class and applied for the same year. To address concerns of non random timing in VC selection, I use variations in the size of local public pension funds' assets in the home-state of a company as an exogenous determinant of VC selection.

My findings provide a unique glimpse into the distribution of innovation changes in the economy following VC activity. To do that I classify citations received by patents into groups, according to whether the citing patentee is VC-backed or not, and whether the citing patentee has been financed by the same VC as the issuing company. I show that the post-financing increase in citations is not concentrated within the VC industry, but diffuses more generally. From a distribution of welfare perspective, this result is important for policy, as it implies that the VC -innovation channel can have positive distributional consequences.

This paper provides suggestive evidence of spillover effects from VC-backed companies to other companies. The existence of knowledge spillovers induced by VC activity, can help explain why few VC investments can have such a dramatic impact on patent production, as estimated in prior papers.

6 References

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

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.

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

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

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 (1992) 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, pg. 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.

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, 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 (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.

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.

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]

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

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

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)*.

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

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

Wooldridge (1998) Distribution-Free Estimation of Some Nonlinear Panel Data Models. Journal of Econometrics 90 (1999) 77-97

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

Table 1 - Summary statistics of matched sample

The matched sample consists of 168,460 patents awarded between 1976 and December 2008 to 5,346 companies that were financed by at least one of 1,424 VC firms, between January 1978 and December 2009. The total number of investments from VC firms to companies that patented during 1976-2008 was 43,837. VC-investments are divided into two broad classes, traditional and non-traditional investments. Traditional investments include: Bridge loans, early stage, expansion, later stage and seed. Non-traditional investments include: Acquisition for Expansion, Acquisition, LBO, MBO, Open Market Purchase, Other, PIPE, Recap or Turnaround, Secondary Buyout and Secondary Purchase. Patents issued by Texas Instruments and ATT are excluded from the sample as they represented more than 5% of the total sample of patents issued to VC-backed companies. (Both Texas Instruments and ATT had Open Market Purchases from VCs in my sample).

Panel A. Application and grant years of patents issued by VC-backed companies and VC investments in companies that patent

	Patents		VC investments	
	Applications	Grants	Traditional	Non-traditional
1976	1,022	16	26	8
1977	941	487	50	10
1978	1,022	867	78	16
1979	1,079	690	119	11
1980	1,182	912	218	29
1981	1,320	982	486	25
1982	1,364	1,003	679	48
1983	1,296	1,106	910	42
1984	1,497	1,322	1,057	14
1985	1,513	1,428	1,009	21
1986	1,704	1,413	1,160	45
1987	1,927	1,807	964	53
1988	2,207	1,701	928	41
1989	2,475	2,363	1,273	39
1990	2,754	2,192	1,036	21
1991	2,952	2,401	840	13
1992	3,605	2,655	961	10
1993	4,395	2,946	797	9
1994	5,375	3,545	869	18
1995	7,532	4,017	907	21
1996	7,731	4,725	1,151	22
1997	10,037	4,885	1,424	21
1998	10,104	7,567	1,527	34

1999	11,472	8,197	2,079	15
2000	13,681	9,043	2,789	15
2001	14,994	9,220	2,331	14
2002	15,599	10,111	2,120	19
2003	13,015	11,615	2,315	31
2004	11,164	11,837	2,612	15
2005	7,814	11,523	2,510	22
2006	3,505	15,281	2,329	25
2007	929	14,638	2,305	38
2008	35	14,747	1,917	48
2009			1,233	15
Total	167,242	167,242	43,009	828

Panel B. Distribution of Patents and VC investments by state

	Patents	VC Investments	
		Traditional	Non-traditional
AK	5		3
AL	311	26	2
AR	1	1	
AZ	1,031	363	10
CA	84,057	21,705	225
CO	2,258	1,112	39
CT	1,316	591	11
DC	295	71	7
DE	36	7	
FL	868	631	20
GA	712	580	10
HI	6	9	
IA	268	26	3
ID	69	28	6
IL	916	551	32
IN	372	84	9
KS	57	17	5
KY	14	22	1
LA	29	21	
MA	11,351	5,705	69
MD	1,186	720	34
ME	473	10	3
MI	397	314	7

MN	6,274	655	58
MO	190	143	5
MS	36	36	4
MT	30	3	1
NC	1,292	837	24
ND	18	1	3
NE	55	4	
NH	596	405	6
NJ	1,716	1,069	25
NM	56	74	1
NV	67	41	
NY	2,156	663	45
OH	805	376	22
OK	64	37	1
OR	546	592	3
PA	2,644	991	28
RI	69	87	2
SC	19	35	
SD	4	1	
TN	171	109	4
TX	32,429	2,065	63
UT	329	165	6
VA	684	463	11
VT	54	9	2
WA	11,729	1,342	10
WI	388	161	7
WV	5	40	1
WY	6	11	
Total	167,242	43,009	828

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

Type of Investment	Number of deals	Percentage of sample
Acquisition for Expansion	235	0.54
Acquisition	51	0.12
Bridge Loan	2,678	6.11
Early Stage	7,637	17.42
Expansion	15,389	35.11
LBO	288	0.66

Later Stage	11,546	26.34
MBO	4	0.01
Open Market Purchase	694	1.58
Other	152	0.35
Other Acquisition	4	0.01
PIPE	526	1.2
Pending Acquisition	1	0
Recap or Turnaround	274	0.63
Secondary Buyout	5	0.01
Secondary Purchase	200	0.46
Seed	4,153	9.47
Total	43,837	

Table 2 - 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). Panel A describes the distribution of VC deals and patent filings over time. Panel B details the distribution of prior patents and their issuing companies across U.S. states. For panel B, I use the state of the company as reported in the SDC database, as the home state of the patent. Panel C shows the distribution of investments by VC-firms in companies with prior patents. I exclude non-traditional VC investments in companies, such as PIPEs (Private Investments in Public Equity), acquisitions and open market purchases. Panel D details the distribution of VC-backed companies that patent across industries. The industry classification is based on the SDC files. Panel E describes the financial situation of the VC-backed companies with prior patents by 2010. Panel F shows the distribution of prior patents by age.

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

	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. Number of observations and citations to prior patents by issuing company's home-state

				Average annual citations per patent		
State	Companies	Patents	Patent-year observations	All patentees	Non VC-backed patentees	VC-backed patentees
AL	2	10	119	1.210	1.168	0.042
AZ	12	42	800	0.619	0.600	0.019
CA	263	761	12,484	1.126	0.957	0.169
CO	18	84	1,474	1.398	1.165	0.233
CT	20	77	1,802	0.498	0.452	0.047
DC	1	6	80	2.425	2.275	0.150
DE	1	1	21	0.524	0.524	0.000
FL	15	46	767	1.417	1.313	0.104
GA	15	30	652	0.868	0.770	0.098
IA	1	1	20	0.250	0.250	0.000
ID	3	16	196	0.449	0.327	0.122
IL	19	48	879	1.188	0.941	0.247
IN	3	5	112	0.250	0.241	0.009
KS	1	1	25	0.040	0.040	0.000
LA	3	13	219	1.379	1.324	0.055
MA	106	245	5,181	0.793	0.696	0.097
MD	18	79	1,570	0.789	0.666	0.122
ME	1	5	87	0.414	0.391	0.023
MI	12	27	583	0.645	0.515	0.130
MN	17	31	569	1.596	1.084	0.511
MO	7	17	343	0.569	0.548	0.020
NC	11	19	219	0.406	0.306	0.100
ND	1	3	88	0.534	0.534	0.000
NE	1	4	128	0.344	0.336	0.008

NH	11	29	666	1.008	0.856	0.152
NJ	20	54	1,035	0.576	0.534	0.042
NM	3	8	103	0.359	0.359	0.000
NV	1	1	15	0.533	0.533	0.000
NY	30	79	1,364	1.082	0.983	0.099
OH	14	52	1,012	1.300	1.186	0.115
OR	5	44	834	0.638	0.546	0.092
PA	26	116	2,782	0.512	0.455	0.057
RI	2	2	49	0.306	0.306	0.000
SC	2	4	86	0.465	0.453	0.012
TN	6	46	788	0.623	0.589	0.034
TX	40	227	4,349	0.700	0.643	0.057
UT	6	17	450	0.902	0.744	0.158
VA	9	21	398	1.982	1.789	0.193
VT	2	8	167	0.431	0.401	0.030
WA	20	44	675	1.287	1.193	0.095
WI	3	8	167	0.743	0.587	0.156
WY	1	5	161	0.267	0.267	0.000
Total	752	2,336	43,519	0.920	0.799	0.120

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

	Number of Companies	Percentage of sample
Bridge Loan	21	2.79
Early Stage	257	34
Expansion	299	18
Later Stage	91	39.76
Seed	84	12.10
		11.17
Total	752	

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

	Number of companies	Percentage of sample
Biotechnology	63	8.38
Communications and Media	75	9.97
Computer Hardware	51	6.78
Computer Software	94	12.50

Consumer Related	33	4.39
Industrial Energy	97	12.90
Internet Specific	37	4.92
Medical Health	145	19.28
Other Products	30	3.99
Semiconductors	127	16.89
Total	752	

Panel E. Status of VC investments by 2010

	Number of companies	Percentage of sample
Acquisition	282	37.50
Active	209	27.79
Bankruptcy - Chapter 11	4	0.53
Bankruptcy - Chapter 7	5	0.66
Defunct	140	18.62
In Registration	1	0.13
LBO	7	0.93
Merger	10	1.33
Other	2	0.27
Pending Acquisition	1	0.13
Went Public	91	12.10
Total	752	

Panel F. 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	

Table 3 – Portfolio Connections for companies in the analysis sample

This table summarizes the number of portfolio connections for companies in the analysis sample. For every company in the analysis sample, the set of portfolio connections corresponds to all companies in the portfolio of the VC investor when the company is first financed by a VC.

Panel A. Portfolio-connections of VC-backed companies in the analysis sample

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

Panel B. Portfolio-connections of VC-backed companies in the analysis sample and average annual citations from portfolio-connections, by year of first time VC financing

	Portfolio Connections		Patents	
	(1)	(2)	(3)	(4)
	Average	Std. Dev.	Number of patents	Average annual citations from portfolio-connections
1978	2.33	2.08	3	0.000
1979	2.83	2.4	6	0.000
1980	5.5	7.55	10	0.000
1981	5.93	4.85	28	0.000
1982	12.53	13.38	15	0.000
1983	10.93	9.65	14	0.002
1984	9.07	13.22	15	0.032
1985	12.25	13.01	8	0.000
1986	11.18	12.15	17	0.000
1987	6.85	7.91	20	0.003
1988	10.92	13.77	12	0.002
1989	15.47	22.18	15	0.000
1990	15.25	18.3	16	0.017
1991	19.13	29.21	16	0.006
1992	10.38	7.84	13	0.025
1993	23.89	36.71	9	0.018
1994	20.17	22.64	18	0.055
1995	15.42	38.18	24	0.003
1996	14.13	23.59	38	0.003
1997	13.03	22.47	36	0.000

1998	21.54	30.05	67	0.001
1999	28.34	59.84	56	0.000
2000	19.07	28.59	86	0.014
2001	13.16	16.98	55	0.009
2002	18.21	37.28	77	0.008
2003	20.28	29.44	78	0.000
Total	16.66	29.89	752	0.008

Table 4 - Summary statistics restricted sample 1993-2008

Information on public state and local 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). Panel A describes the distribution of VC deals and patent filings over time. Panel B details the distribution of local and state public pension funds assets by state, as well as the distribution of prior patents and their issuing companies across U.S. states. I use the state of the company as reported in the SDC database, as the home state of the patent. 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 US millions.

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. Number of observations and citations to prior patents by issuing companies' home-state, and distribution of local and state pension funds' assets by state

	State Pension Funds' Assets		Prior Patents			
	Mean	Std. Dev	Companies	Patents	Patent-year observations	Annual average citations per patent
AL	0.159	0.027	2	10	119	1.210

AZ	0.185	0.045	5	21	246	0.793
CA	2.593	0.790	166	482	5542	1.600
CO	0.204	0.057	12	38	484	1.269
CT	0.148	0.033	10	19	246	0.679
DC	0.031	0.009	1	6	80	2.425
FL	0.635	0.211	9	28	335	1.182
GA	0.335	0.101	8	9	107	1.748
ID	0.044	0.015	2	15	175	0.486
IL	0.604	0.161	10	22	268	1.481
IN	0.119	0.035	1	2	19	0.211
LA	0.173	0.043	3	5	67	1.164
MA	0.273	0.089	51	96	1168	1.177
MD	0.259	0.058	6	27	310	0.971
ME	0.046	0.019	1	3	37	0.568
MI	0.454	0.104	7	13	165	1.055
MN	0.274	0.064	10	17	236	1.140
MO	0.254	0.072	3	7	85	1.035
NC	0.348	0.100	9	17	182	0.462
NH	0.026	0.011	5	10	131	1.618
NJ	0.352	0.070	12	22	251	0.992
NM	0.094	0.031	3	8	103	0.359
NV	0.091	0.034	1	1	15	0.533
NY	1.673	0.414	19	46	577	1.700
OH	0.774	0.146	7	23	288	2.031
OR	0.205	0.110	3	7	70	0.957
PA	0.561	0.128	13	35	452	0.929
TN	0.191	0.049	1	27	346	0.647
TX	0.86	0.253	28	103	1224	1.257
UT	0.087	0.029	3	4	51	3.314
VA	0.29	0.080	6	8	95	4.011
VT	0.015	0.004	1	2	31	0.323
WA	0.298	0.076	14	33	407	1.550
WI	0.418	0.100	2	4	53	1.868
Total	0.3845	0.105	434	1,170	13965	1.378

Table 5– Incidence Rate Ratios and Relative Incidence Rate Ratios by type of citing patentee

The table presents Incidence Rate Ratios (IRRs) and Relative Incidence Rate Ratios (RIRRs) of the exposure to first time VC financing on citations to patents. IRRs are defined as the ratio between the average number of citations received by a patent after the issuing company is first financed by a VC (post-financing), and the average number of citations received by a patent before the issuing company is first financed by a VC (pre-financing). IRRS are interpreted as the ratio of the probability that a patent receives a citation post-financing, to the probability that a patent receives a citation pre-financing. For example, the IRR (column 5) of 1.63 in the first row, implies that the likelihood of a citation to a prior patent increases by 63% after the issuing company is first financed by a VC. RIRRs adjust IRRs for changes in the likelihood of citations to patents at the technology-class and application-year level. RIIRs are defined as the ratio between the IRR of prior patents and the IRR of matching patents, where patents are matched by technology-class and application year. The RIRR (column 6) of 1.33 in the first row, implies that the likelihood of a citation to a prior patent increases by 33% after the issuing company is first financed by a VC, relative to the increase in the likelihood of a citation to a matching patent. The second and third rows summarize the IRRs and RIRRS by type of citing patentee. A citing patentee is classified as Non-VC backed if at the time of the application of the citing patent the patentee has not received VC financing, and, VC-backed otherwise. Portfolio connections are defined in Table 3. Note that average annual citations from portfolio connections to matching patents are defined as the average number of citations that matching patents receive from *the portfolio connections of the prior patent*.*, **, and *** indicate that the IRR and RIRRs are statistically significant at the 10%, 5% and 1% level.

	Average Annual citations				Incidence Rate Ratio	Relative Incidence Rate Ratio
	Prior patents		Matching patents			
	Pre-financing	Post-financing	Pre-financing	Post-financing		
	(1)	(2)	(3)	(4)	(5)	(6)
All	0.64	1.04	0.54	0.66	1.63***	1.33***
NVC	0.59	0.89	0.49	0.58	1.52***	1.29***
VC	0.05	0.15	0.05	0.08	2.89***	1.64***
Portfolio-connections	0.002	0.008	0.0009	0.003	4.05***	1.32***
Non Portfolio-connections	0.05	0.14	0.05	0.08	2.84***	1.63***

Table 6 –VC Financing and patent citations

The table presents within-patent changes in innovation diffusion around the VC financing event. An observation is a patent-year. 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. 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	Negative Binomial	Negative Binomial	QMLE
VC_{pt}	1.627***	1.346***	1.189***
	(0.106)	(0.065)	(0.045)
Constant	0.636***	1.191***	
	(0.038)	(0.047)	
Observations	43,519	41,172	38,981
Number of patents	2,336	2,336	2,183
Number of companies	752	752	723
Offset citations baseline at the tech-class and app. year level	No	Yes	Yes
Patent FE	No	No	Yes

Table 7 – 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. F_{t-1}^{a} 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
I_{t-1}^{a}	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 8– GMM-IVs estimation of within-patent relationship between VC financing and patent citations

This table reports the effect of VC financing on relative citations. An observation is a patent-year. Y_{pt}^{sa} is the number of citations received by the patent at time t. 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. F_{t-1}^{a} 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 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 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)
F_{t-1}^{a}	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 9– Robustness Checks GMM-IVs estimation of within-patent relationship between VC financing and patent citations

This table reports the effect of VC financing on relative citations. An observation is a patent-year. Y_{pt}^{sa} is the number of citations received by the patent at time *t*. For columns (2)-(4the regression includes the citations baseline at the technology class, application year and state level with a coefficient fixed to 1. 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 F_{t-1}^{a} 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 and demeaned by state and time. 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)
F_{t-1}^{a}	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 10– Robustness Checks GMM-IVs estimation of within-patent relationship between VC financing and patent citations

This table reports the effect of VC financing on relative citations. An observation is a patent-year. Out-state citations correspond to the number of citations received by the patent at time t from patentees located in a different state. 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. 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. F_{t-1}^{a} 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 and demeaned by state and time. 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)
F_{t-1}^{a}	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 11 – Distribution of forward citations to prior patents by type of citing patentee

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. Non VC-backed is a dummy that equals one if the type of citing patentee is Non VC-backed. Non portfolio-connected is a dummy that equals one if the type of citing patentee is a non portfolio connection. Portfolio-connected is a dummy that equals one if the type of citing patentee is a portfolio connection. VC_{pt} is a dummy that equals one after the issuing company of the prior patent is first financed by a VC. The Negative Binomial 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. *, **, and *** indicate statistical significance at the 10%, 5% and 1% levels, respectively.

	(1)	(2)	(3)	(4)
Model	Negative Binomial	Negative Binomial	QMLE	QMLE
A. Estimated IRRs				
Non VC-backed	0.585***	1.198***		
	(0.022)	(0.039)		
Non portfolio-connected	0.049***	1.069		
	(0.005)	(0.097)		
Portfolio-connected	0.002***	0.936		
	(0.001)	(0.575)		
$VC_{pt} *$ Non VC-backed (I)	1.517***	1.306***	1.396***	1.184***
	(0.061)	(0.049)	(0.043)	(0.036)
$VC_{pt} *$ Non portfolio-connected (II)	2.839***	1.859***	2.221***	1.192**
	(0.298)	(0.207)	(0.193)	(0.090)
$VC_{pt} *$ Portfolio-connected (III)	4.052***	2.739	2.898***	2.785**
	(1.619)	(1.866)	(0.790)	(1.276)
B. Ratio of IRRs post-financing				
II/I	1.871***	1.423***	1.591***	1.007
	(0.180)	(0.152)	(0.129)	(0.078)
III/I	2.671**	2.097	2.076***	2.352*
	(1.057)	(1.428)	(0.565)	(1.083)
Observations	130,557	71,191	54,824	48,901
Number of patents	2,336	2,336	2,183	2,183
Number of companies	752	752	726	726
Offset citations baseline at the tech-class and app. year level	No	Yes	No	Yes
Patent-type of citation FE	No	No	Yes	Yes

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 is regressed onto the citation intensity and 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).