Perspectives on Recombination

Elizabeth G. Pontikes
University of Chicago Booth School of business

April 2014

Draft. Do not cite or quote.

Abstract

Research in economics and sociology over the past century has pointed to recombination as the source for novel social and economic developments. This study suggests that the categorical structure a person uses to understand a domain is fundamental to this concept. This is studied in an investigation of venture capital financing of software organizations. Findings show that venture capitalists are more likely to invest in companies that engage in recombination based on *market categories*, but that traditional measures of recombination based on patent classes do not have predictive value. Results are strongest for private equity venture capitalists and weakest for corporate venture capitalists, suggesting that people who value novelty based on breaking down existing boundaries will favor recombination, while those who prefer progress that reinforces existing categories will avoid it.

Schumpeter's insight that novelty emerges from "the carrying out of new combinations" spawned a large and diverse literature. Scholars point to new combinations of existing elements, or "recombination," as the primary antecedent for novel developments. Recombination is reified as the "ultimate source of novelty" (Schoenmakers and Duysters 2010; Fleming 2001), and "the 'holy grail' of innovation research" (Gruber, Harhoff and Hoisl 2013). Research on recombination is prolific. Studies show that new combinations of existing components lead to radically new inventions, innovation, and knowledge creation (Henderson and Clark 1990; Tushman and Anderson 1986; Hargadon and Sutton 1997), resulting in especially successful outcomes (Fleming 2001; Schoenmakers and Duysters 2010; Nerkar 2003).

The concept of recombination relies on the assumption that some elements are similar and others are different. Recombination results when *different* elements are brought together. This leads to social, economic, and technological breakthroughs. But researchers have not unpacked how a person's *a priori* understanding of similarity and difference affects when recombination occurs. Instead, studies either define recombination in an ad hoc manner for a particular product or technology, or assume that technological distinctions derived from patent classification are natural divisions across which recombination can occur. This has led to a proliferation of definitions of radical as compared to incremental development without convergence on what constitutes a "new combination" across settings.

This study proposes that recombination depends on the categorical system people use to interpret a domain. For example, a venture capitalist uses market categories to understand organizations and products, whereas an inventor uses technological distinctions reflected in patent classification. A new development comprised of two elements may be seen by a venture capitalist as bringing together very different components, but by an inventor as drawing on similar knowledge. The first case is recombination that leads to radical change; the second is an incremental advance. This casts recombination, and rewards associated with it, as dependent on the categorical system used in a domain.

A person's perspective also affects whether recombination is valued. Recombination is important because innovations that combine distant elements fundamentally change current structures and transform the value system. For example, Schumpeter (1934) describes "new combinations" as the source of "spontaneous and discontinuous change … disturbance of equilibrium, which forever alters and displaces the equilibrium state previously existing," (p. 64). New combinations are critical because they lead to a "process of industrial mutation…that incessantly revolutionizes the economic structure *from within*, incessantly destroying the old one, incessantly creating a new one," (1942:83). As Van de Ven (1986) states, "innovation is a new *idea*, which may be a recombination of old ideas, a scheme that challenges the present order… a unique approach which is perceived as new by the individuals involved," (p. 591). The idea that recombination is the source of developments that alter the social and economic structure underlies researchers' prolonged interest in the subject. It also suggests that whether development based on new combinations will be embraced or resisted depends on a person's orientation toward fundamental structural change (Pontikes 2012).

This study explores two ways in which perspective is central to recombination. First, the relevant categorical system a person uses to understand a domain establishes the boundaries over which recombination can occur. Second, people who wish to challenge boundaries will seek out recombination, while those who want to reinforce existing boundaries will avoid it.

Recombination and Boundaries

Scholars have proposed that recombination is the source of novelty in technical and non-technical domains. As Nelson and Winter (2010) state, "the creation of any sort of novelty in art, science, or practical life—consists to a substantial extent of a recombination of conceptual and physical materials that were previously in existence," (p. 130). Studies show that novel products and technologies are based on new combinations of existing components (Hargadon and Sutton 1997; Henderson and Clark 1990). New organizational forms emerge when activists blend elements across social boundaries (Rao 1998; DiMaggio 1991; Clemens 1996; Phillips and Owens 2004; Rindova and Petkova 2007). Many of these studies take an *ex post* view (Eggers and Kaul 2014), investigating a new development that became successful and tracing it back to its innovative roots.

Studies that take an *ex ante* approach typically use patent data. Patent data track a wide array of inventions, ranging from incremental advances to radically new technologies. This provides researchers an opportunity to investigate whether there is a systematic relationship between recombination and the subsequent importance of an invention. Findings underscore the inherent uncertainty associated with recombination. Patents that build on similar elements are more useful on average, but those that combine distant components have variable outcomes. Distant combinations lead to breakthroughs or failure (Fleming 2001; Fleming and Sorenson 2001). This distinction maps to the concept of exploration and exploitation in the organizational learning literature. Exploiting local knowledge is reliable but predictable, while exploring distant knowledge is riskier but has the potential to be revolutionary (March 1991; Nerkar 2003; Kogut and Zander 1992; Katila and Ahuja 2002). Given the importance of recombination, recent research has turned toward investigating individual characteristics, structures, and processes that lead to recombinatory inventions (Gruber, Harhoff and Hoisl 2013; Fleming and Sorenson 2004; Schoenmakers and Duysters 2010; Fleming and Waguespack 2007; Fleming, Mingo and Chen 2007).

Previous studies recognize that recombination is based on perceptions of similarity, but the implications of this are not fully explored. Scholars suggest that people's views of similarity and difference develop through experiences and social construction (Fleming 2001; Gruber, Harhoff and Hoisl 2013). Studies show that patent classification is not a mere reflection of natural technological differences, but is a system of institutionalized categories that influences how impactful patents become (Wezel, Kovacs and Carnabuci 2014). Technological boundaries are understood to change over *time* as technologies evolve: for example, a material based on sand and aluminum would have seemed strange to a scientist in the 1940s but is widely recognized as a semiconductor today (Fleming 2001). But innovation researchers have not considered that category boundaries differ based on a person's perspective, even at the same point in time. This paper incorporates insights from categorization research into the concept of recombination. The category literature suggests that categorical boundaries are defined in relation to an audience, or a set of people who classify objects in a similar way (Hannan, Pólos and Carroll 2007; Hsu and Hannan 2005). Different audiences can have different conceptions of, and reactions to, categorization. For example, investors that are "market makers" prefer ambiguous categories, while consumers that are "market takers" have an aversion to ambiguity (Pontikes 2012).

If two people use two different categorical systems to interpret a domain, then an invention that combines technologies may be distant "recombination" in one person's view but may be a local combination for another. This means that recombination depends on the perspective of the person evaluating the technology. Previous studies focus on the technical perspective, measuring recombination across patent classes and outcomes based on patent citations (Fleming 2001; Nerkar 2003; Fleming and Sorenson 2001; 2004; Wezel, Kovacs and Carnabuci 2014). But recombination based on patent classes will not resonate with a person who does not use patent classification. For a person who understands a domain using market categories, novel developments will be those that combine elements across markets. More generally:

> *Proposition 1*: Novel developments are those that combine elements across *relevant categorical boundaries*; that is, from the categorical system a person uses to understand a domain.

Valuing Recombination

Most studies that investigate *ex ante* effects study how recombination affects citations, which represent the usefulness of a patent. But the broader literature on technological change and industry evolution suggests that recombination has more expansive implications. Novel developments have the potential to radically change an industry and are the source of value creation (Schumpeter 1934; Nelson and Winter 1982; Weitzman 1998). Major technological shifts undermine the position of established competitors while providing opportunities for new upstarts (Tushman and Anderson 1986; Abernathy and Clark 1985). Studies of successful novel developments demonstrate that they are based on recombination of disparate ideas, concepts, and technologies (Henderson and Clark 1990; Hargadon and Sutton 1997). But *ex ante* approaches show that not all recombination is successful. Inventions based on new combinations have variable outcomes, either becoming especially useful or overlooked (Fleming and Sorenson 2001; Fleming 2001). This has been studied for patent citations. But it should also be evident for other measures of novelty in a domain.

In many contexts, venture capitalists are an important constituency who usher in new technologies and products that transform a market. These investors look for the next "new, new thing" (Lewis 1999), investing in organizations with unique products that have the potential to create new markets (MacMillan, Zemann and Subbanarasimha 1987; Hisrich and Jankowicz 1990). As a result, they embrace companies that do not conform to existing categories. Consumers favor organizations that are clearly classified. But venture capitalists are "market makers" and prefer organizations in ambiguous market categories, which provide flexibility to initiate major changes in an industry (Pontikes 2012). The venture capitalist investment strategy carries substantial risk, with most investments resulting in losses and a few generating substantial returns (Sahlman 1990).

The challenge for venture capitalists is to select promising companies from a set of unproven organizations where quality is uncertain (Hall and Lerner 2010; Stuart, Hoang and Hybels 1999). Patents are observable attributes that can signal quality (Hsu and Ziedonis 2013). Qualitative research indicates that both investors and entrepreneurs use patents as an indicator of technical expertise, branding, or differentiation (Lemley 2000; Mann 2005; Graham and Sichelman 2008). Quantitative studies of venture capital backed companies also provide evidence that investors use patents as quality signals. Patents are more important at attracting prominent venture capital investors and higher valuations when there are fewer alternative quality signals (Hsu and Ziedonis 2013). Even in the software industry, where patents are not as easily protected as compared to other technology domains, patents are correlated with financing rounds and total investment (Mann and Sager 2007). The few studies that investigate a set of start-ups that have not received prior financing also suggest that patents attract funding: biotechnology start-us with patents raise more money (Baum and Silverman 2004), and early stage companies with both patents and a prototype are more likely to receive equity financing (Audretsch, Bönte and Mahagaonkar 2012).

Studies of how patents influence venture capital funding analyze the number of patents an organization has. But venture capitalists look for organizations that can transform an industry, and patents are not necessarily an indicator of this type of potential. Patents are granted for a wide range of inventions that range from trivial to groundbreaking. To the extent that patents are used as a signal, venture capitalists will be less influenced by whether an organization has patents, but will value organizations with patents that have the potential to be transformative. The literature on the creation of novelty suggests that it is organizations that engage in recombination that will be most promising. This suggests:

Hypothesis 1a: Venture capitalists are more likely to invest in organizations that develop technologies based on recombination.

According to proposition 1, it is important to identify the relevant categorical system venture capitalists use to understand a domain, in order to identify organizations that engage in recombination. Venture capitalists do not interpret the world according to technical categories reflected in the patent system. Rather, they use market categories to understand a domain. This suggests that venture capitalists will positively value organizations that combine technologies across market category boundaries. In contexts where market categories are tightly coupled with patent classification, recombination across market categories will be equivalent to recombination across technical categories from the venture capitalist perspective. For example, Lerner (1994) finds that biotechnology organizations that patent across technical boundaries receive higher venture capital valuations. Biotechnology organizations directly market their scientific discoveries, so patent classification reflects the relevant category system used by an investor.

In contexts where market categories are not tightly coupled with technical categories, such as the software industry, the underlying categorical system will have important implications for how venture capitalists reward recombination. For example, collaboration software, a product that allows people to work remotely, is based on a number of different patented technologies: mathematical algorithms, graphical interface developments, and microphone technology. From a technical perspective, these are very different technologies, but from a market-based perspective, they are all associated with the same market category, and so do not constitute a distant combination. Conversely, recombination from a technical perspective may be an incremental advance from the perspective of the market. A software organization that draws on different knowledge bases to create a more effective algorithm may develop a revolutionary patent, but not a breakthrough product. Venture capitalists will favor organizations engaged in recombination across market boundaries. When market classification is different from technical classification, this effect will be stronger than recombination across technical boundaries:

> *Hypothesis 1b*: The positive effect of recombination on venture capital funding is more pronounced when measured across market boundaries, as compared to technical boundaries.

A corollary is that venture capitalists will be less likely to fund organizations that create technologies that build on knowledge from within their category boundaries:

> *Hypothesis 2*: Venture capitalists are less likely to invest in organizations that build on the technologies from within their market category.

Previous studies show that recombination leads to novelty in a domain, but not necessarily to higher average performance. In fact, Fleming (2001) finds that inventions based on distant combinations have more variance in citations, but lower citations on average. This reflects the trade-off between exploration and exploitation described in March (1991). Distant combinations are risky: either resulting in breakthroughs or failure. This is why novel developments that become successful are traced to recombination (Henderson and Clark 1990; Hargadon and Sutton 1997; Tushman and Anderson 1986), and why recombination is the foundation of theories of transformative social and economic change (Nelson and Winter 1982; Schumpeter 1934; Van de Ven 1986; Weitzman 1998).

This suggests recombination will appeal to people who value novelty, breaking down categorical barriers, and risking failure for a chance at transformative success. This characterizes the prototypical private equity venture capitalist. People who are interested in more measured, predictable development –

those who wish to reinforce categorical boundaries – will shy away from recombination. Segmenting venture capitalists by type allows for an investigation of these two perspectives. Most venture capital investment is made by private equity firms, who are traditional market-makers, interested in transforming the structure of an industry. But corporations are increasingly creating venture capital arms that take minority investments in entrepreneurial ventures. Corporate venture capital (CVC) firms have many of the same goals as traditional venture capitalists, but are different in key ways. These firms are successful within existing market categories and have an interest in retaining the status quo. Established organizations are unlikely to introduce new developments (Tushman and Anderson 1986; Sørensen and Stuart 2000; Christensen 1997). Compared to private equity firms, CVCs invest in less risky ventures, in larger syndicates, and many do not award performance pay (Dushnitsky and Shapira 2010). Where private equity firms are typical "market makers" who prefer organizations in ambiguous market categories, corporate venture capital act as "market takers" and avoid organizations that are ambiguously classified (Pontikes 2012). Corporate venture capitalists are not interested in the "destructive" part of creative destruction. Rather, they are drawn to organizations that develop new knowledge that reinforces categorical boundaries. This suggests that corporate venture capitalists will have an aversion to recombination:

> *Hypothesis 3a*: Corporate venture capitalists are less likely to invest in organizations that recombine technologies across market categories.
>
> *Hypothesis 3b*: Corporate venture capitalists are more likely to invest in organizations that build on technologies from within their market category.

Empirical Test

These ideas are studied in the software industry between 1990 and 2002. The software industry is a fast-paced, innovative context, and a fair amount of innovation in this industry can be tracked through patents. Studies suggest that patents influence venture capital financing, IPOs, and acquisition (Mann 2005; Mann and Sager 2007; Cockburn and MacGarvie 2011). Venture capital financing has been critical for software organizations. The software industry was the largest or second largest recipient of venture capital financing for each year in the 1990s (Onorato 1997). In 2002 venture capitalists invested $691 million in 156 different software organizations (Mann 2005). Software products are difficult to make sense of, and as a result market categories are important to helping investors and customers make sense of its many organizations and products in this context (Pollock and Williams 2007; Wang 2009; Pontikes 2012). Although products are built on patented technologies, market categories are only loosely coupled with patent classification. This makes software a good context to study recombination from the perspectives of both technical and market classification.

Data and Methods

Hypotheses are tested using a unique data set of patenting behavior and venture capital financing events for all software organizations that issued a press release between 1990 and 2002. These data are well suited to test the hypotheses because they contain small and young organizations that never received venture capital funding. Much previous research on how patents influence venture capital financing use data on organizations that received at least one round of venture capital investment (Hsu and Ziedonis 2013; Mann and Sager 2007). To the extent that patents act as signals, they will be most important in the earliest stages of funding when little is known about the organization (Hsu and Ziedonis 2013). The data used in this study include organizations that do not receive funding in the risk set, and allow for tests of how patents – and recombination – affect the likelihood of receiving initial funding. Press release data provide a historical record of the market categories organizations are in, based on claims by the producers. This captures the market categories used by industry experts at the time.

All press releases in *Businesswire, PR Newswire, and Computerwire* that contained at least three mentions of the word “software” were the initial source of data. A combination of text-matching programs and manual inspection of their output extracted 4,566 software organizations from these documents. These data were matched to IPO data from Thomson Financial and IPO data compiled by Jay

Ritter, to identify private organizations.[1] There are 4,113 organizations that are private at some point during the time period studied. Through searches through press releases, Hoovers, company Web sites (if still alive), and (at last resort) Wikipedia founding dates could be located for 3,316 private organizations. The fact that almost 800 organizations could not be tracked using standard data sets indicates the wide breadth of organizations included in press releases, including those that would otherwise be unknown. Patent portfolios were gathered from the NBER U.S. Patent Citations data file (Hall and Lerner 2010). The NBER data do not contain information on subclass assignments for patents, so these data were also matched to the Patent Network Dataverse for subclass assignments (Li et al. 2014). Venture capital financing for these organizations was gathered from the Venture Economics database from Thomson Financial. Data on market categories were extracted from press releases. Press releases typically contain a one-sentence description where the organization affiliates with a market category. This was used to identify the market categories organizations were in for each year in the analysis (Pontikes 2008). A list of market categories was compiled from articles in leading industry publications *Software Magazine* and *Computerworld*, and by reading the self-descriptions from press releases. The data contain 456 market categories. Gartner, the preeminent industry analyst, issued reports on over half of these market categories between 1995 and 2002. This suggests that market categories from press releases reflect the taxonomy used by industry insiders.[2]

Dependent variables. The main dependent variable is whether the organization receives venture capital financing in the current year. Empirical tests are run on all private organizations. Organizations that IPO exit the data as a censored observation. The complete data contain 4,113 private organizations over 13,555 years, with 1,625 financing events for 826 organizations. Additional analyses investigate whether the organization receives financing from a large venture capital firm, defined as being in the top twenty in

---

[1] http://bear.warrington.ufl.edu/ritter/ipodata.htm

[2] Overlap with markets covered by Gartner is a conservative estimate of whether categories in press releases reflect the common parlance. Gartner reports are only available from 1995 onward, and they cover a wide range, but not all, types of software. Gartner issues reports on high-technology product markets that are important for their clients (including most of the Fortune 500), and so do not include categories innovative organizations try to pioneer.

the amount invested under capital in the current year (652 events for 406 organizations). Analyses are also run on: all private organizations that have previously patented (512 organizations over 1,951 years; 194 events for 120 organizations), all private organizations that are less than 15 years old, or where a founding date could not be located (2,803 organizations over 8,496 years; 1,325 events for 668 organizations), and private organizations less than 15 years old who have patented previously (339 organizations over 1,018 years; 160 events for 98 organizations). Models are also run to estimate first round financing. These are run on all private organizations, that have not previously received venture capital financing, that are less than 10 years old (1,947 organizations over 4,827 years). There are 280 events of first round financing, 145 events of first round financing from private equity firms only, and 48 events of first round financing from a corporate venture capital firm. [3]

Independent variables. Patent data are used to measure recombination across market categories, whether an organization builds on knowledge from within its categories, and recombination across technical categories. All measures are timed by the application year of patents (only patents that were later granted are used).

Market category recombination: Recombination across market categories is measured based on whether an organization's patents combine technologies from categories it is not in. To compute this measure, an n-dimensional "knowledge space" is first constructed, using all patents in the Computers & Communications class (Hall, Jaffe and Trajtenberg 2001). This includes all patents relevant to software, not just those issued to organizations in the press release data. Knowledge space includes a five-year window of all patents issued in the current year and the previous four years, for every year from 1990 – 2002. Patents are located in this space based on citation overlap (Podolny, Stuart and Hannan 1996).

[3] Only funding events specified as round 1 financing are used. Thomson's Venture Economics does not include round 1 funding events for all organizations. Most financing is from private equity firms, and CVC firms typically invest along with other private equity firms. Private equity only indicates that *all firms* investing in the round were PE firms. CVC indicates that there was *at least one* CVC firm investing in the round.

Similarity between two patents is measured by dividing the number of shared citations by citations made by the focal patent: $\alpha_{mn} = \frac{s_{mn}}{s_m}$, where s_{mn} is the number of shared citations between patent *m* and *n*, and s_m is the total number of citations by patent *m*. Second degree similarity is also computed to capture if two patents are similar based on shared citations with a third patent. This is calculated by multiplying the similarity of patent *m* and *n* to a third patent, *k*, $\sigma_{mn} = \max_{\alpha_{mk}>0 \wedge \alpha_{kn}>0} \{\alpha_{mk} \cdot \alpha_{kn}\}$, capturing similarity based on shared citations with a third patent. The patent *k* that yields the maximum similarity between *m* and *n* is used.

An organization's market category recombination is measured based on its bringing together knowledge associated with market categories it is not in. Areas of knowledge space are mapped to market categories based on patenting behavior of organizations in the category (Pontikes and Hannan 2013). It is important to ensure that organizations in multiple categories do not have a larger influence than organizations in fewer categories. Therefore a grade of membership is computed, which assigns "membership" scores in each category between (0,1] based on the frequency with which the organization affiliates with each category. Grade of membership divides (the number of press releases in which organization *i* claims category *A*) by (the number of press releases in which it claims any category) in a given year. Market category recombination is computed using the knowledge space similarity (σ_{mn}) between each patent *m* issued to organization *i*, and each patent *n* that is affiliated with a category the organization is not in (D_i). A patent *n* is affiliated with D_i if it was issued to an organization *j* that is a member of D_i, weighted by *j*'s grade of membership in D_i:

$$prox_{m,D_i} = \sum_{(m \in P_i) \wedge (n \in P_j) \wedge (\mu_j(D_i)>0)} \mu_j(D_i) \times \sigma_{mn} \qquad (1)$$

Here, P_i is the patent portfolio of organization *i*, and P_j are the patent portfolios of all organizations j that are members of category D_i. This is summed over all categories D_i that the organization is not in:

$$prox_{m,D} = \sum_{(\mu_i(D_i)=0)\wedge(m\in P_i)} prox_{m,D_i} \quad (2)$$

Patent level proximities are then averaged over the number of patents issued to the organization. Because the distribution is skewed, the natural log is used:

$$\text{Market category recombination} = ln\left(\frac{\sum_{m\epsilon P_i} prox_{m,D}}{npat_i}\right) \quad (3)$$

In analyses that include organizations that do not patent, those without patents are given a value of zero for recombination.

Patents based on existing categories: To test hypothesis 2 and 3b, patent data are used to measure the extent to which an organization builds on knowledge related to its own categories. This is constructed in the same way as market category recombination (equation 3), but an organization's knowledge space proximity to patents in its own categories, S, is used. In analyses that include organizations that do not patent, this measure is assigned a value of zero.

Technical category recombination: To test hypothesis 1b, recombination across technical categories is computed. Previous studies use a number of different measures of recombination based on patent classes (see Gruber *et al* (2013) for a detailed review). Three different measures are used here:

1. Traditional Herfindahl index: a widely used measure of recombination is derived from a Herfindahl index based on the class assignments of patent *j*'s cited patents. This is the measure of patent originality proposed by Trajtenberg et al (1997):

   $$Herf_p = 1 - \sum_{k=1}^{K}\left(\frac{cite_{p,k}}{J}\right)^2 \quad (4)$$

   Here patent p has J backward citations to k = 1 … K patent classes, and $cite_{p,k}$ is the number of p's citations in class k. Patents are sometimes reclassified. The NBER patent database contains the original class assignments of all patents, and so original class assignments are used. For an organizational level of recombination, the Herfindahl is averaged over the organization's patents, *P,* for a given year.

Recombination (citations' patent class) = $\frac{\sum_{p=1}^{P} Herf_p}{P}$ (5)

2. Technical niche width: A second measure of recombination uses the subclass assignments of the focal patent. This is a Herfindahl that measures whether the patents' subclasses are in the same or different class.

$$niche_p = 1 - \sum_{c=1}^{C} \left(\frac{sub_{p,c}}{S}\right)^2 \quad (6)$$

Here, patent p is assigned to *S* subclasses in *C* classes, where $sub_{p,c}$ is the number of assigned subclasses in class *C*. This is averaged over the organization's patents, *P,* in the given year. Subclass assignments are available from the Patent Network Dataverse for the patent's current subclass assignment only. Again, for an organization level measure, the niche width is averaged over the organization's patents for a given year:

Recombination (technical niche width) = $\frac{\sum_{p=1}^{P} niche_p}{P}$ (7)

3. Patent breadth: The third measure used counts the number of classes to which the patent is assigned. This is similar to the measure is used by Lerner (1994). This is averaged over the organization's patents for a given year.

Controls. A number of controls are included that may influence venture capital financing. These include the number of patents issued to the organization and the number of citations received, in the previous year. Previous studies show that venture capitalists prefer ambiguous categories and so category fuzziness is included. At the category level, this is measured as one minus the average grade of membership of organizations in the category. At the organizational level, category fuzziness is averaged over the categories the organization is in, weighted by grade of membership (Pontikes 2012). The number of other organizations in the same market categories (weighted by grade of membership) is included to control for the popularity and competitiveness of the organization's current market. The number of acquisitions made by the organization, its tenure in the press release data, number of previous rounds of financing, and whether it was ranked in *Software Magazine*'s Software 500 are included to control for organizational size, quality, and other attributes. Supplementary analyses control for organizational age for organizations whose founding dates could be located. Because it is less likely that founding dates were found for unsuccessful organizations that fail early, it is important to include these in the analysis. Therefore a control is included that indicates if the organization's founding date is known. Supplementary analyses also include the total number of citations to the organization's patents in the four years *following* its grant date, as a control for the quality of the organization's patents.[4] All independent and control variables are measured as of the beginning of each time period. Tables 1 and 2 contain descriptive statistics for the variables included in the analysis. Correlations are in the appendix.

--- Insert tables 1 and 2 about here ---

Model and Estimation. The hypotheses are tested by estimating the instantaneous likelihood that the organization receives venture capital financing during the time period Δt in the limit where $t \to 0$. This can be operationalized in terms of two random variables: Y(t), which indicates whether an organization receives venture capital funding at time t, and t_n, the time it receives funding:

$$r(t) = \lim_{\Delta t \to 0} \frac{\Pr(Y(t - t_n + \Delta t) \mid Y(t - t_n) = 0)}{\Delta t} \tag{8}$$

This rate is estimated as a function of the independent and control variables listed above:

$$r(t - t_n) = r_o(t - t_n) \cdot \exp(\beta_{ind} \cdot \mathbf{x}_{ind} + \alpha_{control} \cdot \mathbf{x}_{control}) + \varepsilon \tag{9}$$

The estimation uses piecewise continuous hazard rate models employing the stpiece routine in Stata written by Jesper Sørensen. The piecewise exponential specification allows the base rate of receiving

[4] Forward citations are measured after the patent's grant date, as a measure of the quality of the patent. Patents are attributed to the organization at their application date.

funding to vary in time "pieces" according to the number of years in which the organization has been "waiting" for funding. Pieces are included for less than one year, [1-3) years, [3-5) years, and 5 or more years. In models with repeated events, organizations exit and enter with a new ID. In these models standard errors are clustered by organization to correct for departures from statistical independence. Estimations on subsets of venture capital investment (large VCs, PE only, and CVC) are modeled as competing risks.

Results

The relationship between market category recombination and venture capital financing is evident through a graphical inspection of the data. Figure 1 plots the mean number of venture capital funding events by bands of technical recombination, for patenting organizations. The blue line is the mean value and the grey shaded area is the confidence interval. The graph shows that organizations that recombine technologies across market categories are more likely to receive venture capital financing.

--- Insert figure 1 about here ---

Table 3 provides statistical tests of hypothesis 1a and 2. These are piecewise continuous hazard rate models on an organization's likelihood to receive venture capital financing in a given year. Models 1 – 4 are run on all private organizations. Models 1 and 2 contain controls only. Model 1 includes the number of patents issued to the organization, and model 2 includes the number of citations. The number of citations, frequently used to indicate the importance of a patent, does not reach conventional levels of significance. This is likely because venture capitalists invest in early stage organizations, before enough time has passed for citation counts to be credible signals of quality. Model 1 is used as the base model for hypothesis tests. Model 1 shows that the number of patents has a negative effect on an organization's likelihood to receive financing. This effect does not retain significance in all models, but the coefficient is consistently negative. This indicates that the count of an organization's patents is not itself a draw for venture capital investment.

Model 3 tests hypothesis 1a by including market category recombination. The effect is positive and significant ($p < 0.01$), providing support for the hypothesis. Model 4 tests hypothesis 2 by including the extent to which an organization's patents build on knowledge from market categories it is in. The effect is negative and marginally significant ($p < 0.10$). Model 5 is run on patenting organizations only. The effect of market category recombination remains positive and significant ($p < 0.05$), providing additional support for hypothesis 1a. The effect of patents building on existing categories loses significance. This effect is negative across models but does not reach significance at conventional levels. Therefore hypothesis 2 is not supported, although there is evidence of a negative trend.

--- Insert table 3 about here ---

Table 4 contains tests of hypothesis 1b, which states that the positive effect of recombination on venture capital funding is more pronounced when elements are combined across market category boundaries, as opposed to technical categories. Models include variables measuring recombination based on citations' patent classes, technical niche width, and patent breadth. The citation patent class measure is positive and marginally significant ($p < 0.10$) in model 6, run on all organizations, but loses significance in model 7, run on patenters only. The niche width measure does not reach significance in model 8, run on all organizations (or in models run on patenters only). The patent breadth measure is positive and significant at $p < 0.05$ in model 9, run on all organizations, but also loses significance in models run on patenters only (model 10). Models 11 and 12 include all recombination variables that reached conventional significance levels when included independently, in models run on all organizations and on patenters only. Results show that when the variables are included together, market category recombination continues to have a positive and significant effect, while the other measures lose significance and even change sign. Together, these results provide support for hypothesis 1b. Venture capitalists are more likely to fund organizations that recombine technologies across *market category boundaries*. Recombination across *technical category boundaries* does not reliably attract venture capital financing.

--- Insert table 4 about here ---

Additional Tests

Table 5 includes additional tests of hypothesis 1a. One question may be whether venture capitalists' preference for market category recombination is due to their choosing organizations that create important technologies. To test for this, the *future citations* of an organization's patents, in four years after the patent was granted, are included in the model. Patent citations are the best indicator researchers have identified to measure the impact of a patent (Trajtenberg 1990). Importance can reliably be measured *ex post*, but there are not reliable present-time indicators of the future importance of patents (Fleming 2001). It is unusual to include future events in a statistical model, but in this case the control provides a conservative test of the hypothesis. If causality is reversed and companies that are funded are more likely to have their patents cited, this biases the results in the favor of the control. Results show that forward citations do not predict venture capital funding. Effects of recombination persist. Venture capitalists prefer to invest in organizations that engage in recombination, but there is no evidence that this is a proxy for picking companies that will have highly cited patents. This suggests that either venture capitalists value market category recombination in itself and are not seeking out companies with impactful patents, or that they are not good at predicting which patents will be most important in the future.

Models 15 and 16 are run on organizations that are less than 15 years old or where the founding date is unknown. This excludes from the risk set old private organizations that may not be looking for venture capital financing. Again, results persist. During the time period of the analysis, there were changes in legal restrictions on the patenting of software. A 1995 ruling lifted most of the restrictions. Previous research shows despite legal barriers, software patents were broadly issued throughout the time period of study (Cohen and Lemley 2001; Mann 2005). Still, it is important to confirm whether results hold for the period after restrictions were lifted. Models 17 and 18 are run after 1995. Results continue to provide support for hypothesis 1a.

--- Insert table 5 about here ---

Venture capitalist type and investment stage

Table 6 contains models that test effects based on the venture capitalist type and the stage of investment. Model 19 estimates effects on whether an organization reeives financing from a large venture capital firm, based on the total amount of investment under capital. Influential venture capital firms are able to attract more investment, so this is an indicator of the prominence of the investment. Findings show that large venture capital firms also prefer organizations that engage in market category recombination.

Models 20 – 23 estimate the likelihood that a firm receives first round financing. They are run on the set of firms that have not previously been funded, that are less than 10 years old or where the founding date is unknown. This substantially reduces statistical power. As a result, some insignificant controls are excluded, and time pieces that do not significantly differ are condensed, to reduce statistical noise (including these does not substantially change results). Model 20 estimates the likelihood of first round financing. Results show that market category recombination has a positive effect on first round financing ($p < 0.05$). The size of the effect is larger than in models on all financing rounds ($b = 0.356$ (0.168)). The effects are not statistically different, but the trend supports the idea that patents are most important as signals when there is the most uncertainty surrounding the investment (Hsu and Ziedonis 2013). Model 21 estimates effects on first round investment that is comprised of only private equity venture capitalists. Results again show that market category recombination has a positive and significant effect and the coefficient is even larger ($b = 0.500$ (0.248)). The trend supports the suggestion that private equity venture capitalists will have the strongest preference for recombination.

Models 22 and 23 test hypotheses 3a and 3b. These models estimate effects on first round investment that includes a corporate venture capital firm. Corporate venture capital investment is not frequent and as a result this model does not have much statistical power. Neither market category recombination nor patents building on existing categories reach statistical significance, though the coefficients are in the proposed direction: the coefficient for recombination shows a negative trend, and patents building on existing categories yield a positive trend, the opposite preference as private equity venture capitalists. Model 23 excludes recombination and uses a continuous control for year, instead of year dummies, to reduce statistical noise. This model also shows a positive effect of organizations with patents that build on existing categories at higher confidence, but not at conventional levels of statistical significance ($p = 0.11$). Hypotheses 3a and 3b are not supported. However, results do provide some directional evidence of these trends, which perhaps can be explored in the future using data with more statistical power.

--- Insert table 6 about here ---

Effects of controls

Results show that the fuzziness of an organization's categories has a positive effect on venture capital investment, consistent with Pontikes (2012). This effect is insignificant in models that include patenting organizations only, which suggests that both effects reflect venture capitalists' preference for organizations with the potential for novelty. For organizations that patent, combining technical elements across market categories may be a better signal of this potential than having an ambiguous market identity. The number of organizations in a category has a positive effect on venture capital investment, which may reflect the underlying quality of certain market categories, or may be an indicator of industry fads.

Discussion and Conclusion

The idea that recombination is the source of novelty underlies research of technical, economic, and social change. This paper suggests that the concept of recombination is inherently categorical, and that whether recombination is recognized and rewarded depends on a person's categorical perspective. Findings support this view. Venture capitalists prefer to invest in organizations that engage in recombination, a finding that provides support for the view that recombination underlies novelty. But effects also indicate that recombination is based on the categorical system used to understand a domain. Venture capitalists prefer for recombination across *market categories*.

Previous research on recombination has taken a technical perspective and shows that recombination across patent classes results in especially useful inventions. This study compliments this view, but emphasizes the importance of considering the underlying categorical structure when investigating recombination. For inventors, patent classes are relevant categorical boundaries. But these are by no means natural divisions that determine an objective basis for recombination for all people. In some contexts, market categorization is based on technical distinctions. But in many contexts, technical categories are not the basis of how people understand the market. Findings here reinforce this point. Recombination based on technical categories has only weak effects on venture capital investment when included alone in the estimation. If one did not consider what was the relevant categorical structure from the perspective of a venture capitalist, he might come to the erroneous conclusion that new combinations do not drive venture capital investment. Results show that venture capitalists do prefer to invest in organizations that develop technologies based on recombination, when new combinations are defined with respect to the relevant categorical structure. This supports an audience-based view of categorization, which allows for different segments of people to have different views on how categories are defined (Pontikes 2012; Hannan, Pólos and Carroll 2007).

This paper also suggests that whether people value recombination will depend on their orientation toward developments that promise to fundamentally change the categorical structure in a domain. This was investigated by comparing first-round investment between private equity and corporate venture capitalists. Results show that private equity venture capitalists strongly prefer organizations that engage in market category recombination. But results for corporate venture capital investment were not strong enough to reject the null. There is directional support in line with the hypotheses, providing some indication that corporate venture capitalists may not value recombination as strongly and perhaps have an aversion to it. But future research is needed for more conclusive evidence.

Results also speak to the literature that investigates whether patenting influences venture capital financing. Findings indicate that the number of patents is not as important as the type of patenting an organization engages in. An organization’s number of patents has a negative or null effect on its likelihood to receive venture capital funding in this context. Previous studies have shown that patenting has a positive effect on venture capital funding, mostly in studies of industries like biotechnology and semiconductors, where intellectual property is an important competitive differentiator (Hsu and Ziedonis

2013; Baum and Silverman 2004). Studies of the software industry provide qualitative evidence of the importance of patents (Mann 2005; Graham and Sichelman 2008). A quantitative investigation in this industry of how patents affect repeat financing and total amount invested suggest a positive correlation, but results are not conclusive (Mann and Sager 2007). Findings here show that it is the types of technologies organizations develop, rather than number of patents, that is most important to venture capitalists. Given that venture capitalists gather in depth knowledge on each of their investments, and that patents are issued for a wide range of new developments, it should not come as a surprise that the content of patents is as or more important than patent counts. This is likely to be most pronounced in contexts where intellectual property protection is weak, but we should expect that recombination would be preferred by novelty-seeking venture capitalists even in where it is strong. Future research investigating this in other industries will be informative.

This work has implications for research on categorization. Much of this literature emphasizes how categories reflect institutional boundaries that constrain individual action, leading to a reproduction of the existing structure (Zuckerman 1999; 2000). More recently, scholars have turned their focus to the observation that many actors challenge institutional boundaries, causing categorical structures to change over time (Rao, Monin and Durand 2005; Lounsbury and Rao 2004; Pontikes 2008). This study shows that existing classification has important consequences for the emergence of novelty in a domain and underlies processes that transform the social and economic structure.

References

Abernathy, William J, and Kim K. B. Clark. 1985. "Innovation: Mapping the Winds of Creative Destruction." *Research Policy* 14:3-22.

Audretsch, David B, Werner Bönte, and Prashanth Mahagaonkar. 2012. "Financial Signaling by Innovative Nascent Ventures: The Relevance of Patents and Prototypes." *Research Policy* 41:1407-1421.

Baum, Joel AC, and Brian B. S. Silverman. 2004. "Picking Winners or Building Them? Alliance, Intellectual, and Human Capital As Selection Criteria in Venture Financing and Performance of Biotechnology Startups." *Journal of Business Venturing* 19:411-436.

Christensen, C. 1997. *The Innovator's Dilemma: When New Technologies Cause Great Firms to Fail.* Boston, MA: Harvard University Press.

Clemens, Elisabeth S. 1996. "Organizational Form As Frame: Collective Identity and Political Strategy in the American Labor Movement, 1880-1920." Pp. 204-226 in *Comparative Perspectives on Social Movements*, edited by Doug McAdam, John D McCarthy and Mayar N Zald. Cambridge University Press.

Cockburn, I M, and . J. MacGarvie. 2011. "Entry and Patenting in the Software Industry." *Management Science* 57:915-933.

Cohen, Julie, and Mark Lemley. 2001. "Patent Scope and Innovation in the Software Industry." *California Law Review* 89:1-57.

DiMaggio, Paul J. 1991. "Constructing An Organizational Field As a Professional Project: U.S. Art Museums, 1920-1940." Pp. 267-292 in *The New Institutionalism in Organizational Analysis*, edited by Walter J Powell and Paul J DiMaggio. Chicago: The University of Chicago Press.

Dushnitsky, G, and Z Shapira. 2010. "Entrepreneurial Finance Meets Organizational Reality: Comparing Investment Practices and Performance of Corporate and Independent Venture Capitalists." *Strategic Management Journal* 31:990-1017.

Eggers, J.P., and Aseem Kaul. 2014. "Fools Rush In: A Behavioral Perspective on Incumbent Pursuit of Radical Technologies." *Working Paper, NYU Stern.*

Fleming, L, and O Sorenson. 2004. "Science As a Map in Technological Search." *Strategic Management Journal* 25:909-928.

Fleming, Lee. 2001. "Recombinant Uncertainty in Technological Search." *Management Science* 47:117-132.

Fleming, Lee, and David D. M. Waguespack. 2007. "Brokerage, Boundary Spanning, and Leadership in Open Innovation Communities." *Organization Science* 18:165-180.

Fleming, Lee, and Olav Sorenson. 2001. "Technology As a Complex Adaptive System: Evidence From Patent Data." *Research Policy* 30:1019-1039.

Fleming, Lee, Santiago Mingo, and David Chen. 2007. "Collaborative Brokerage, Generative Creativity, and Creative Success." *Administrative Science Quarterly* 52:443-475.

Graham, S, and T Sichelman. 2008. "Why Do Start-ups Patent." *Berkeley Technology Law Journal* 23:1063-97.

Gruber, M, D Harhoff, and K Hoisl. 2013. "Knowledge Recombination Across Technological Boundaries: Scientists Vs. Engineers." *Management Science* 59:837-851.

Hall, Bronwyn, Adam Jaffe, and Manuel Trajtenberg. 2001. "The NBER Patent Citations Data File: Lessons, Insights, and Methodological Tools." in *NBER Working Paper Series.* Cambridge, MA: National Bureau of Economic Research.

Hall, Bronwyn H, and Josh Lerner. 2010. "The Financing of R&D and Innovation." Pp. 609-639 in *Handbook of the Economics of Innovation*, vol. 1, edited by Bronwyn Hall and Nathan Rosenberg.

Hannan, Michael T, László Pólos, and Glenn Carroll. 2007. *Logics of Organization Theory: Audiences, Codes and Ecologies.* Princeton, NJ: Princeton University Press.

Hargadon, Andrew, and Robert Sutton. 1997. "Technology Brokering and Innovation in a Product Development Firm." *Administrative Science Quarterly* 42:716-749.

Henderson, Rebecca, and Kim Clark. 1990. "Architectural Innovation: The Reconfiguration of Existing Product Technologies and the Failure of Established Forms." *Administrative Science Quarterly* 35:9-30.

Hisrich, R D, and A. D. Jankowicz. 1990. "Intuition in Venture Capital Decisions: An Exploratory Study Using a New Technique." *Journal of Business Venturing* 5:49-62.

Hsu, David H, and Rosemarie R. H. Ziedonis. 2013. "Resources As Dual Sources of Advantage: Implications for Valuing Entrepreneurial-firm Patents." *Strategic Management Journal* 34:761-781.

Hsu, Greta, and Michael Hannan. 2005. "Identities, Genres, and Organizational Forms." *Organization Science* 16:474-490.

Katila, R, and G Ahuja. 2002. "Something Old, Something New: A Longitudinal Study of Search Behavior and New Product Introduction." *Academy of Management Journal* 45:1183-1194.

Kogut, Bruce, and Udo Zander. 1992. "Knowledge of the Firm, Combinative Capabilities, and the Replication of Technology." *Organization Science* 3:383-397.

Lemley, M A. 2000. "Reconceiving Patents in the Age of Venture Capital." *J. Small & Emerging Bus. L* 4:137.

Lerner, Joshua. 1994. "The Importance of Patent Scope: An Empirical Analysis." *The RAND Journal of Economics* 25:319.

Lewis, Michael. 1999. *The New New Thing: A Silicon Valley Story.* WW Norton & Company.

Li, Guan-Cheng et al. 2014. "Disambiguation and Co-authorship Networks of the U.S. Patent Inventor Database (1975–2010)." *Research Policy.*

Lounsbury, Michael, and Hayagreeva Rao. 2004. "Sources of Durability and Change in Market Classifications: A Study of the Reconstitution of Product Categories in the American Mutual Fund Industry, 1944-1985." *Social Forces* 82:969-999.

MacMillan, Ian C, lauriann Zemann, and P. N. Subbanarasimha. 1987. "Criteria Distinguishing Successful From Unsuccessful Ventures in the Venture Screening Process." *Journal of Business Venturing* 2:123-137.

Mann, R J, and T. W. Sager. 2007. "Patents, Venture Capital, and Software Start-ups." *Research Policy* 36:193-208.

Mann, Ronald J. 2005. "Do Patents Facilitate Financing in the Software Industry?" *Texas Law Review* 83:1-71.

March, James G. 1991. "Exploration and Exploitation in Organizational Learning." *Organization Science* 2:71-87.

Nelson, R, and S Winter. 1982. *An Evolutionary Theory of Economic Change.* Cambridge: Harvard University Press.

Nerkar, Atul. 2003. "Old Is Gold? The Value of Temporal Exploration in the Creation of New Knowledge." *Management Science* 49:211-229.

Onorato, Nicole. 1997. *Trends in Venture Capital Funding in the 1990s.* Washington, DC: U.S. Small Business Administration Office of Advocacy.

Phillips, Damon J, and David Owens. 2004. "Incumbents, Innovation, and Competence: The Emergence of Recorded Jazz, 1920 to 1929." *Poetics* 32:281-295.

Podolny, Joel, Toby Stuart, and Michael M. T. Hannan. 1996. "Networks, Knowledge, and Niches: Competition in the Worldwide Semiconductor Industry, 1984-1991." *American Journal of Sociology* 102:659-689.

Pollock, Neil, and Robin Williams. 2007. "Technology Choice and Its Performance: Towards a Sociology of Software Package Procurement." *Information and Organization* 17:131-161.

Pontikes, Elizabeth G. 2008. "Fitting in or Starting New? An Analysis of Invention, Constraint, Contrast, and the Emergence of New Categories in the Software Industry." PhD Dissertation, Stanford University, Stanford, CA: Stanford University.

------. 2012. "Two Sides of the Same Coin: How Ambiguous Classification Affects Multiple Audience Evaluations." *Administrative Science Quarterly* 57:81-118.

Pontikes, Elizabeth G, and Michael M. T. Hannan. 2013. "The Geometry of Social Classification." *Working Paper.*

Rao, Hayagreeva. 1998. *"Caveat Emptor:* The Construction of Nonprofit Consumer Watchdog Organizations." *American Journal of Sociology* 103:912-961.

Rao, Hayagreeva, Phillipe Monin, and Rodolphe Durand. 2005. "Border Crossing: Bricolage and the Erosion of Categorical Boundaries in French Gastronomy." *American Sociological Review* 70:968-991.

Rindova, V P, and . P. Petkova. 2007. "When Is a New Thing a Good Thing? Technological Change, Product Form Design, and Perceptions of Value for Product Innovations." *Organization Science* 18:217-232.

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

Schoenmakers, Wilfred, and Geert Duysters. 2010. "The Technological Origins of Radical Inventions." *Research Policy* 39:1051-1059.

Schumpeter, Joseph. 1942. *Capitalism, Socialism, and Democracy.* New York: Harper & Brothers Publishers.

------. 1934. *The Theory of Economic Development.* Cambridge, MA: Harvard University Press.

Stuart, Toby E, Ha Hoang, and Ralph R. C. Hybels. 1999. "Interorganizational Endorsements and the Performance of Entrepreneurial Ventures." *Administrative Science Quarterly* 44:315.

Sørensen, J B, and T. E. Stuart. 2000. "Aging, Obsolescence, and Organizational Innovation." *Administrative Science Quarterly* 45:81-112.

Trajtenberg, Manuel. 1990. "A Penny for Your Quotes: Patent Citations and the Value of Innovations." *The RAND Journal of Economics* 21:172.

Trajtenberg, Manuel, Rebecca Henderson, and Adam Jaffe. 1997. "University Versus Corporate Patents: A Window on the Basicness of Invention." *Economics of Innovation and New Technology* 5:19-50.

Tushman, Michael, and Philip Anderson. 1986. "Technological Discontinuities and Organizational Environments." *Administrative Science Quarterly* 31:439-465.

Van de Ven, Andrew H. 1986. "Central Problems in the Management of Innovation." *Management Science* 32:590-607.

Wang, Ping. 2009. "Popular Concepts Beyond Organizations: Exploring New Dimensions of Information Technology Innovations." *Journal of the Association for Information Systems* 10:1-30.

Weitzman, Martin L. 1998. "Recombinant Growth." *The Quarterly Journal of Economics* 113:331-360.

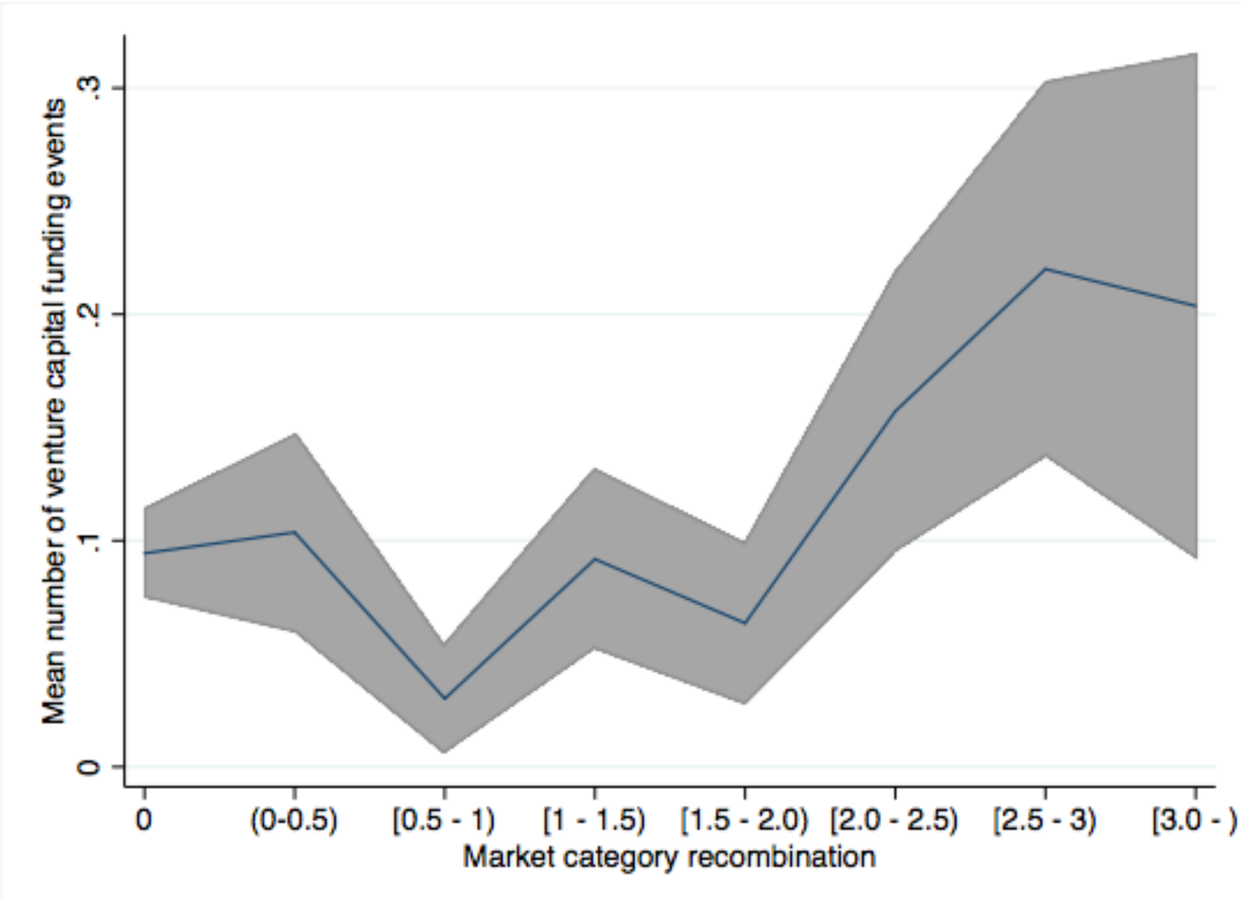
Wezel, Filippo, Balazs Kovacs, and Gianluca Carnabuci. 2014. "Being There: Patent Class Contrast and the Impact of Technological Innovations." *Working Paper.*

Zuckerman, E. 2000. "Focusing the Corporate Product: Securities Analysts and De-Diversification." *Administrative Science Quarterly* 45:591-619.

Zuckerman, Ezra W. 1999. "The Categorical Imperative: Securities Analysts and the Illegitimacy Discount." *American Journal of Sociology* 104:1398-1438.

Figures

Figure 1. Venture capital funding events by market category recombination.



Tables

Table 1. Descriptive statistics for all organizations.

	Mean	Standard Deviation	Minimum	Maximum
Received venture capital financing	0.1199	0.3248	0	1
Recombination (market category)	0.1142	0.4641	0	4.153
Patents build on existing categories	0.0139	0.0977	0	2.462
Recombination (citations' patent class)	0.0402	0.1461	0	0.9017
Recombination (niche width)	0.0231	0.0995	0	0.8639
Recombination (patent breadth)	0.0845	0.2853	0	2
Number of citations to received in future 4 years (logged)	0.1733	0.7595	0	8.405
Number of patents	1.335	20.74	0	821
Number of citations	9.307	180.05	0	9519
Category fuzziness	0.3806	0.2696	0	0.8332
Number of organizations in categories (weighted, logged)	2.407	1.8923	0	6.618
Number of acquisitions	0.0091	0.1083	0	4
Tenure in data	2.178	2.442	0	13
Age (since founding)	11.15	12.06	0	165
Number of previous rounds of VC funding	0.5465	1.561	0	20
Ranked in Software 500	0.0808	0.2725	0	1
Year	1998.10	2.968	1990	2002

N= 13,555 (11,316 for age)

Table 2. Descriptive statistics for patenting organizations.

	Mean	Standard Deviation	Minimum	Maximum
Received venture capital financing	0.0994	0.2993	0	1
Recombination (market category)	0.7935	0.9787	0	4.153
Patents build on existing categories	0.0966	0.2417	0	2.462
Recombination (citations' patent class)	0.2790	0.2859	0	0.9017
Recombination (niche width)	0.1605	0.2160	0	0.8639
Recombination (patent breadth)	0.5870	0.5201	0	2
Number of citations received in future 4 years (logged)	1.204	1.664	0	8.405
Number of patents	9.276	54.00	0	821
Number of citations	64.66	470.90	0	9519
Category fuzziness	0.5217	0.1363	0	0.7974
Number of organizations in categories (weighted, logged)	3.311	1.278	0	6.605
Number of acquisitions	0.0251	0.1997	0	4
Tenure in data	4.237	2.771	0	13
Age (since founding)	16.94	17.86	1	154
Number of previous rounds of VC funding	0.8273	1.973256	0	15
Ranked in Software 500	0.1199	0.3250	0	1
Year	1998.11	2.809	1990	2002

N= 1,951 (1,699 for age)

Table 3. Piecewise continuous hazard rate models on organization's likelihood to receive VC funding in the given year. All organizations and patenting organizations only.

	Model 1	**Model 2**	**Model 3**	**Model 4**	**Model 5**
	All orgs	All orgs	All orgs	All orgs	Patenters only
Recombination (market category)			0.165**	0.229***	0.233*
			(0.0574)	(0.0673)	(0.0938)
Patents build on existing categories				-0.540+	-0.518
				(0.317)	(0.342)
Number of patents	-0.0289*		-0.0729*	-0.0721*	-0.102**
	(0.0146)		(0.0319)	(0.0305)	(0.0384)
Number of citations		-0.00956			
		(0.00598)			
Category fuzziness	1.514***	1.506***	1.489***	1.441***	0.129
	(0.214)	(0.214)	(0.214)	(0.217)	(0.700)
Number of organizations in categories (weighted, logged)	0.0570*	0.0576*	0.0555*	0.0629*	0.115+
	(0.0277)	(0.0277)	(0.0278)	(0.0282)	(0.0590)
Number of acquisitions	-0.0629	-0.0659	-0.0650	-0.0509	0.509+
	(0.277)	(0.277)	(0.272)	(0.284)	(0.260)
Tenure in data	-0.0736**	-0.0724**	-0.0789**	-0.0777**	-0.0723
	(0.0263)	(0.0263)	(0.0264)	(0.0265)	(0.0440)
Number of previous rounds	0.162***	0.161***	0.163***	0.163***	0.116**
	(0.0268)	(0.0268)	(0.0264)	(0.0265)	(0.0407)
Ranked in Software 500	-0.173	-0.170	-0.180	-0.177	-0.584*
	(0.129)	(0.129)	(0.128)	(0.128)	(0.246)
Time piece: 0-1 year	-1.848***	-1.861***	-1.840***	-1.841***	-1.495
	(0.227)	(0.227)	(0.226)	(0.226)	(1.030)
Time piece: 1-3 years	-3.668***	-3.679***	-3.648***	-3.649***	-2.921**
	(0.236)	(0.236)	(0.236)	(0.236)	(1.007)
Time piece: 3-5 years	-4.491***	-4.499***	-4.456***	-4.461***	-3.953***
	(0.275)	(0.275)	(0.275)	(0.275)	(1.057)
Time piece: 5+ years	-5.338***	-5.336***	-5.277***	-5.285***	-4.989***
	(0.375)	(0.374)	(0.373)	(0.373)	(1.147)
Year Dummies	Yes	Yes	Yes	Yes	Yes
Log pseudo likelihood	-4390.5	-4389.5	-4386.8	-4385.5	-433.9
Degrees of freedom	23	23	24	25	25

+ p<0.10; * p<0.05; ** p<0.01, *** p<0.001

Table 4. Piecewise continuous hazard rate models on organization's likelihood to receive VC funding in the given year. All organizations and patenting organizations only.

	Model 6 All orgs	**Model 7** Patenters	**Model 8** All orgs	**Model 9** All orgs	**Model 10** Patenters	**Model 11** All orgs	**Model 12** Patenters
Recombination (market category)						0.176+	0.232*
						(0.0946)	(0.117)
Patents build on existing categories						-0.548+	-0.494
						(0.311)	(0.334)
Recombination (citations' patent class)	0.391+	0.104				-0.164	-0.252
	(0.221)	(0.247)				(0.398)	(0.395)
Recombination (technical niche width)			0.389				
			(0.307)				
Recombination (patent breadth)				0.312*	0.216	0.219	0.103
				(0.123)	(0.155)	(0.251)	(0.295)
Number of patents	-0.0599*	-0.0829*	-0.0459*	-0.0874*	-0.102*	-0.0856*	-0.1000*
	(0.0276)	(0.0327)	(0.0211)	(0.0389)	(0.0400)	(0.0394)	(0.0408)
Category fuzziness	1.512***	0.278	1.499***	1.498***	0.218	1.439***	0.0965
	(0.214)	(0.695)	(0.215)	(0.215)	(0.697)	(0.217)	(0.699)
Number of organizations in categories	0.0541+	0.0766	0.0565*	0.0545*	0.0748	0.0629*	0.116*
(weighted, logged)	(0.0276)	(0.0549)	(0.0278)	(0.0278)	(0.0547)	(0.0280)	(0.0591)
Number of acquisitions	-0.0596	0.456+	-0.0663	-0.0619	0.452+	-0.0513	0.501+
	(0.276)	(0.244)	(0.275)	(0.276)	(0.249)	(0.285)	(0.260)
Tenure in data	-0.0744**	-0.0818+	-0.0730**	-0.0770**	-0.0730+	-0.0782**	-0.0779+
	(0.0262)	(0.0450)	(0.0263)	(0.0262)	(0.0434)	(0.0266)	(0.0444)
Number of previous rounds	0.161***	0.109**	0.161***	0.162***	0.109**	0.163***	0.119**
	(0.0265)	(0.0414)	(0.0266)	(0.0263)	(0.0414)	(0.0265)	(0.0408)
Ranked in Software 500	-0.170	-0.552*	-0.173	-0.172	-0.563*	-0.176	-0.587*
	(0.129)	(0.247)	(0.129)	(0.128)	(0.244)	(0.128)	(0.246)
Time piece: 0-1 year	-1.845***	-1.413	-1.848***	-1.847***	-1.503	-1.844***	-1.506
	(0.226)	(1.024)	(0.226)	(0.226)	(1.010)	(0.226)	(1.030)
Time piece: 1-3 years	-3.664***	-2.890**	-3.666***	-3.662***	-2.981**	-3.654***	-2.924**
	(0.236)	(1.003)	(0.236)	(0.236)	(0.989)	(0.236)	(1.010)
Time piece: 3-5 years	-4.486***	-3.958***	-4.489***	-4.477***	-4.036***	-4.464***	-3.947***
	(0.275)	(1.054)	(0.276)	(0.275)	(1.039)	(0.275)	(1.054)
Time piece: 5+ years	-5.327***	-4.992***	-5.335***	-5.309***	-5.086***	-5.288***	-4.967***
	(0.374)	(1.148)	(0.374)	(0.373)	(1.136)	(0.373)	(1.149)
Year Dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Log pseudo likelihood	-4388.9	-436.8	-4389.6	-4387.5	-436.1	-4385.2	-433.7
Degrees of freedom	24	24	24	24	24	27	27

+ p<0.10; * p<0.05; ** p<0.01, *** p<0.001

Table 5. Piecewise continuous hazard rate models on organization's likelihood to receive VC funding in the given year. All organizations, young organizations, and patenting organizations.

	Model 13	**Model 14**	**Model 15**	**Model 16**	**Model 17**	**Model 18**
	All orgs	Patenters	All; age <15	Patenters; age< 15	All year >1995	Patenters, year >1995
Recombination (market category)	0.184*	0.186+	0.189**	0.187*	0.223**	0.229*
	(0.0731)	(0.0967)	(0.0656)	(0.0944)	(0.0688)	(0.0953)
Patents build on existing categories	-0.558+	-0.518	-0.428	-0.424	-0.489	-0.533
	(0.313)	(0.341)	(0.298)	(0.325)	(0.326)	(0.356)
Number of patents	-0.0971*	-0.129*	-0.0433	-0.0816*	-0.0509+	-0.0800*
	(0.0457)	(0.0551)	(0.0264)	(0.0405)	(0.0261)	(0.0355)
Patents' future citations	0.0939	0.113				
	(0.0874)	(0.103)				
Category fuzziness	1.439***	0.125	1.194***	-0.221	1.363***	0.402
	(0.217)	(0.698)	(0.220)	(0.671)	(0.237)	(0.849)
Number of organizations in categories (weighted, logged)	0.0637*	0.116+	0.0297	0.101	0.0563+	0.108+
	(0.0282)	(0.0593)	(0.0297)	(0.0618)	(0.0295)	(0.0597)
Number of acquisitions	-0.0420	0.549*	0.0757	0.754*	-0.0687	0.514*
	(0.288)	(0.268)	(0.308)	(0.306)	(0.298)	(0.259)
Organization age			-0.118***	-0.0899**		
			(0.0121)	(0.0310)		
Founding date is known			1.049***	0.740**		
			(0.100)	(0.267)		
Tenure in data	-0.0780**	-0.0721			-0.0677**	-0.0572
	(0.0266)	(0.0440)			(0.0261)	(0.0427)
Number of previous rounds	0.163***	0.116**	0.190***	0.141**	0.154***	0.0979*
	(0.0264)	(0.0407)	(0.0236)	(0.0439)	(0.0274)	(0.0402)
Ranked in Software 500	-0.179	-0.593*	-0.0552	-0.559+	-0.166	-0.629*
	(0.128)	(0.249)	(0.165)	(0.293)	(0.128)	(0.253)
Time piece: 0-1 year	-1.849***	-1.678	-2.065***	-1.070	-2.038***	-1.304**
	(0.227)	(1.057)	(0.256)	(0.900)	(0.104)	(0.488)
Time piece: 1-3 years	-3.662***	-3.111**	-3.622***	-2.358**	-3.866***	-2.815***
	(0.237)	(1.030)	(0.262)	(0.861)	(0.151)	(0.468)
Time piece: 3-5 years	-4.473***	-4.139***	-4.301***	-3.171***	-4.746***	-3.794***
	(0.276)	(1.077)	(0.304)	(0.927)	(0.218)	(0.556)
Time piece: 5+ years	-5.293***	-5.171***	-5.370***	-4.633***	-5.495***	-4.902***
	(0.374)	(1.173)	(0.452)	(1.098)	(0.340)	(0.742)
Year Dummies	Yes	Yes	Yes	Yes	Yes	Yes
Log pseudo likelihood	-4384.9	-433.3	-3913.3	-380.9	-3803.5	-392.6
Degrees of freedom	26	26	26	26	19	19

+ p<0.10; * p<0.05; ** p<0.01, *** p<0.001

Table 6. Piecewise continuous hazard rate models on organization's likelihood to receive VC financing from large VCs, round 1 financing, and round 1 financing from only private equity firms and CVCs.

	Model 19	**Model 20**	**Model 21**	**Model 22**	**Model 23**
	Large VC	Round 1; Age <10	Round 1; PE only; Age < 10	Round 1; CVC included; Age < 10	Round 1; CVC included; Age < 10
Recombination (market category)	0.191*	0.356*	0.500*	-0.117	
	(0.0922)	(0.168)	(0.248)	(0.645)	
Patents build on existing categories	-0.599		-0.837	1.793	1.557
	(0.442)		(1.423)	(1.300)	(0.953)
Number of patents	-0.0638	-0.123	-0.156	-0.00703	-0.00488
	(0.0472)	(0.125)	(0.180)	(0.0632)	(0.0614)
Category fuzziness	1.780***	1.697+	2.502+	-2.236*	-2.513*
	(0.346)	(0.986)	(1.315)	(1.033)	(0.994)
Number of organizations in categories (weighted, logged)	0.0535				
	(0.0436)				
Number of acquisitions	0.302	-11.86	-12.82	-14.21	-10.72
	(0.316)	(365.1)	(844.3)	(3496.6)	(603.0)
Tenure in data	-0.0878*	-0.241*	-0.185	-0.0703	-0.0765
	(0.0386)	(0.108)	(0.139)	(0.329)	(0.323)
Number of previous rounds of financing	0.182***				
	(0.0298)				
Ranked in Software 500	-0.225	-0.108	-0.889	0.271	0.195
	(0.223)	(0.394)	(0.722)	(1.054)	(1.046)
Time piece 1	-2.471***	-2.715***	-2.766***	-20.31	-614.1***
	(0.299)	(0.450)	(0.453)	(2944.2)	(142.0)
Time piece 2	-4.397***	-4.084***	-4.377***	-21.21	-615.0***
	(0.325)	(0.714)	(0.867)	(2944.2)	(142.0)
Time piece 3	-5.786***	-4.548***	-5.563***		
	(0.440)	(1.085)	(1.526)		
Time piece 4	-5.888***				
	(0.549)				
Year (continuous measure)					0.305***
					(0.0710)
Year Dummies	Yes	Yes	Yes	Yes	No
Log pseudo likelihood	-2300.9	-1128.7	-674.4	-262.4	-273.5
Degrees of freedom	25	21	22	21	9

+ p<0.10; * p<0.05; ** p<0.01, *** p<0.001

Appendix

Table A1. Correlations for all organizations

		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
Received venture capital financing	(1)																
Recombination (market category)	(2)	0.01															
Patents build on existing categories	(3)	0.00	0.60														
Recombination (citations' patent class)	(4)	-0.01	0.73	0.40													
Recombination (niche width)	(5)	-0.01	0.62	0.35	0.76												
Recombination (patent breadth)	(6)	-0.02	0.81	0.48	0.90	0.77											
Number of citations received in future 4 years (logged)	(7)	-0.03	0.67	0.39	0.65	0.60	0.76										
Number of patents	(8)	-0.02	0.14	0.15	0.18	0.15	0.21	0.48									
Number of citations	(9)	-0.02	0.12	0.13	0.15	0.12	0.17	0.38	0.90								
Category fuzziness	(10)	-0.03	0.12	0.09	0.13	0.11	0.14	0.09	0.03	0.03							
Number of organizations in categories (weighted, logged)	(11)	-0.02	0.11	0.13	0.12	0.10	0.13	0.08	0.04	0.04	0.86						
Number of acquisitions	(12)	-0.01	0.05	0.09	0.05	0.04	0.06	0.02	0.02	0.02	0.06	0.08					
Tenure in data	(13)	-0.11	0.11	0.10	0.16	0.12	0.17	0.14	0.11	0.11	0.55	0.52	0.06				
Age (since founding)	(14)	-0.18	0.08	0.07	0.17	0.13	0.17	0.23	0.22	0.18	0.13	0.12	0.03	0.35			
Number of previous rounds of VC funding	(15)	0.40	0.04	0.03	0.04	0.04	0.04	0.00	-0.02	-0.02	0.13	0.12	0.01	0.07	-0.10		
Ranked in Software 500	(16)	-0.04	0.04	0.05	0.02	0.02	0.03	0.02	0.03	0.03	0.21	0.21	0.11	0.21	0.11	0.00	
Year	(17)	0.08	-0.06	-0.03	-0.04	-0.06	-0.07	-0.15	-0.03	0.00	0.24	0.28	0.03	0.19	-0.03	0.06	0.04

N= 13,555 (11,316 for age)

Table A2. Correlations for patenting organizations

		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
Received venture capital financing	(1)																
Recombination (market category)	(2)	0.08															
Patents build on existing categories	(3)	0.03	0.52														
Recombination (citations' patent class)	(4)	0.01	0.55	0.24													
Recombination (niche width)	(5)	0.02	0.43	0.20	0.62												
Recombination (patent breadth)	(6)	0.01	0.68	0.35	0.81	0.63											
Number of citations received in future 4 years (logged)	(7)	-0.07	0.51	0.25	0.45	0.41	0.61										
Number of patents	(8)	-0.05	0.06	0.10	0.10	0.08	0.14	0.48									
Number of citations	(9)	-0.04	0.06	0.09	0.09	0.06	0.11	0.37	0.90								
Category fuzziness	(10)	0.06	-0.06	0.06	-0.12	-0.06	-0.12	-0.18	-0.03	0.01							
Number of organizations in categories (weighted, logged)	(11)	0.10	-0.03	0.27	-0.06	-0.07	-0.08	-0.14	0.05	0.07	0.46						
Number of acquisitions	(12)	0.00	0.02	0.10	0.01	0.01	0.03	-0.02	0.01	0.01	0.05	0.09					
Tenure in data	(13)	-0.16	-0.27	-0.05	-0.23	-0.22	-0.27	-0.16	0.12	0.16	0.18	0.17	0.05				
Age (since founding)	(14)	-0.19	-0.08	0.00	0.07	0.04	0.08	0.25	0.33	0.28	0.00	0.01	0.01	0.33			
Number of previous rounds of VC funding	(15)	0.37	0.00	0.02	-0.04	-0.01	-0.05	-0.09	-0.06	-0.05	0.04	0.10	0.02	0.03	-0.15		
Ranked in Software 500	(16)	-0.05	0.02	0.08	-0.06	-0.04	-0.05	-0.05	0.04	0.05	0.12	0.17	0.20	0.11	0.03	-0.03	
Year	(17)	0.09	-0.21	-0.08	-0.16	-0.20	-0.30	-0.49	-0.08	-0.01	0.41	0.39	0.07	0.35	-0.02	0.09	0.07

N= 1,951 (1,699 for age)