Buyouts and Start-up Innovation Incentives

Connie Lee
CERNA, Mines Paristech, PSL Research University
connie.lee@mines-paristech.fr

See Latest Version Here

October 2020

Abstract

This paper investigates how start-up innovation choices are affected by incumbent firm interactions. In particular, incumbent firms have an impact on start-up exit strategies as they can affect their expectations of getting acquired, of succeeding, or of going bankrupt. Using exogenous variation in macroeconomic and financing conditions, I infer a likelihood of getting bought out for entrants. I then estimate how an increase in the expectation of getting acquired affects the new entrant's innovation choices with respect to existing firms. I construct a novel measure of innovation proximity and show that new firms innovate "closer" to their potential acquirers.

JEL classification: O31, L26, G01
Keywords: innovation, entrepreneurship, financial crisis

1 Motivation

There is a large and growing literature on the factors that drive firm innovation. However the *type* of innovations being made by new entrant firms has been lacking in the dialogue. Whether a new firm enters with a minimally differentiated product or a radically innovative product can lead to very different trajectories for the firm and for the industry it is in. Here I will study how firm interactions affect the innovation incentives of the new entrant. Namely, I hypothesize that incumbent firms have an influence on start-up exit options, and start-ups in turn make choices to optimize their exit outcomes.

It is recognized that an exit strategy of getting acquired is increasingly being adopted by start-ups in the US.[1] In this paper I will apply the fact that acquisitions are affected by macroeconomic and financing conditions to estimate an expectation of acquisition for start-ups. I then test whether these expectations affect the start-up's choice in innovation. Using patent data, I build a measure of innovation originality as well as a measure of proximity and complementarity between firms. With data on mergers and acquisitions I identify the firms that are bought out as well as their acquirers. As such, I will provide evidence that start-ups choose to position themselves closer to their potential acquirers when they have a higher expectation of getting bought out.

Why do we care about the type of innovation that firms are doing? Figure 1 shows the average patenting originality of new firms in the US over time. We see patenting originality steadily increasing until 2008 and then clearly falling after. This drop off coincides with a fall in productivity in the wider economy.[2] The peak is a little bit after 2008 however R&D takes time to develop and patenting takes time to be filed so it is expected to have a lag. The literature might explain this fall in originality in different ways, for instance Bloom et al. (2017) suggests that ideas are simply getting harder to find and Akcigit et al. (2013) suggests that this is due to a fall in public funding for basic research. However, the decline in originality coincides with the fall in productivity and the 2008 financial crisis therefore I believe the crisis should have had a role as well.

[1] See Lemley and McCreary (2021)

[2] The declining business dynamism literature dives deeper into the question of how firm entry has been affecting economic growth. See Decker et al. (2014), Decker et al. (2016) and Akcigit and Ates (2019) for an overview of the main concepts in this literature. The innovation provided by new firms is an important aspect in their valuation of entry however the type and originality of innovation is another dimension that has not yet been addressed.

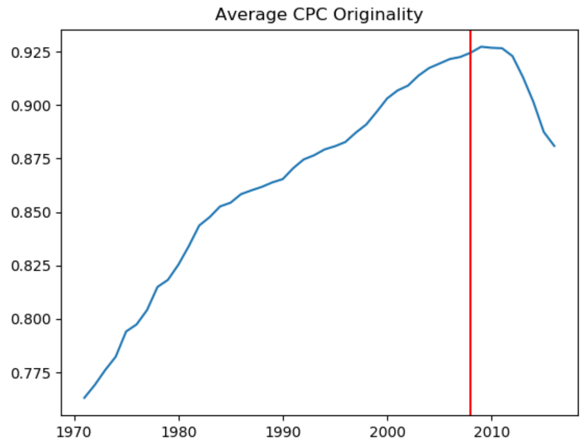


Figure 1: The average firm originality over time in the US
The average originality of new firms in the US. The originality measure is built following Trajtenberg et al. (1997) with full CPC technology codes. The red line is at the year 2008.

In the setting of the 2008 financial crisis and ensuing recession, there was an asymmetric effect on firm access to financing. Large incumbent firms benefited from credit easing policies while small and young firms suffered from changes to financial regulation that imposed more stringent thresholds making it harder to access financing.[3] This asymmetry created more opportunities for capital flows from large firms to young firms and therefore increased buyout expectations for start-ups. I posit that this change in buyout likelihood consequently led start-ups to react strategically in their innovation decisions to further increase their likelihood of getting bought out.

When a start-up enters, it has some degree of choice in what product it will develop, who the founders and employees are, where its funding will come from, where it will locate, where it will sell, what kind of legal status it will take on, etc.[4] The strategic component of firm entry I focus on is the product and corresponding technological content it enters with. What characterizes this product innovation and how does it position with respect to existing products? A new firm

[3]See Davis and Haltiwanger (2019), Greenstone et al. (2020), Bacchetta et al. (2019), and Ayyagari et al. (2018) for some evidence of this asymmetric financial effect.

[4]There is some literature on the initial choices of founding teams when the firm enters. Ouimet and Zarutskie (2014) look at founder characteristics such as age and Choi et al. (2019) look at the composition of founding teams.

that enters with a highly differentiated, original, product will face less competition. However it could also take more effort and experimentation to successfully develop. The risks of failure are higher to invent an original product than one that is largely similar to existing products. Furthermore if the product is radically different, the consumer demand for the product may also be uncertain. On the other hand, if a new firm enters with a product that is complementary and more similar to existing products, it could have an easier, less risky R&D process due to knowledge spillovers. It may be able to benefit from the economies of scale of complementary products and it may increase its probability of getting bought out by potential acquirers with complementary product lines.

The choice of innovation and potential exit options are also important considerations for the start-up when they look for initial sources of capital. When an equity investor, like a venture capitalist or an angel investor invests in a start-up it wants to maximize its return on investment and getting acquired is often the preferred way to achieve this. Of course the start-ups' founders also want to maximize their payoffs with some going so far as to start a company for the sole purpose of selling it quickly - leading to the emergence of serial entrepreneurs. As large existing firms are a critical set of potential acquirers, any factors that influence them are carefully monitored. This is exemplified in a TechCrunch article in response to Elizabeth Warren's announcement of her policy on Big Tech.[5] This article argues that breaking up Big Tech companies will actually have a negative effect on start-ups because it eliminates a major exit option for their investors who will therefore be less willing to invest.

In contrast, when a bank finances a start-up with credit, it primarily cares about getting the interest and principal repaid with minimal risk. A bank does not overly consider the start-ups' exit options and does not take up a seat on the board where it can influence decision making. The financial regulatory changes in response to the crisis, however, added more controls on lending causing access to bank credit for small and new firms to become more difficult. Furthermore, house valuations fell dramatically in the crisis and Davis and Haltiwanger (2019) argue that houses are an important source of collateral for loans to entrepreneurs. While there was a decrease in access to bank financing it was offset to some extent by the flow of funds into venture capital as investors looked for alternative sources of return. As such, there was a decrease in the

[5] https://techcrunch.com/2019/03/08/venture-investors-and-startup-execs-say-they-dont-need-elizabeth-warren-to-defend-them-from-big-tech/

level of traditional bank financing yet an increase in the share of financing from equity investors like angel investors and venture capitalists for new firms.

The innovation literature distinguishes between push and pull effects on innovation. The push effect can come from knowledge spillovers or increased access to financing while the pull effect acts through the demand for innovation. There is an expansive literature on the push drivers such as knowledge spillovers explored in the networks literature, increased resources such as financing, etc. however to the best of my knowledge, the demand pull channel is less explored. It has been discussed in the trade literature as a change in demand comes from the opening up of an export market.[6] It also appears in the environmental economics literature as a regulation change affects the markets for certain products.[7]

While financing has traditionally been considered to have a push effect on innovation by enabling access to more resources, here I suggest it can also exert a pull - the demand for innovative technologies from potential acquirers can affect the direction of innovation firms choose.[8] This is a financial incentive directly implicating the kind of innovation a potential seller-firm is doing. Particularly in the case of equity investors or firm acquirers, there is some pressure to align firm decisions with investors' or acquirers' interests. Tian and Wang (2014) have empirically studied the impact of venture capital tolerance for failure on innovation and Nanda and Rhodes-Kropf (2013) construct a theoretical model of shareholder's failure tolerance and manager's innovation choices to align risk preferences. These are studies that align risk preferences, however there is also a case to be made for aligning technological content. For instance, Atanasova and Chemla (2020) find a familiarity bias in investment decisions made by firm defined benefit pension plans. Investors and acquiring firms exert a demand on their potential target firms' innovation positioning.

The M&A and innovation literature has mainly focused on the ex-post effect of a merger or acquisition on innovation. [9] Chemmanur and Tian (2018) look at the effect of Anti-Takeover

[6]See Aghion et al. (2019)

[7]See Horbach et al. (2012), Nemet (2009), Jones (2011) and Negro and Schorfheide (2004) among others.

[8]See Hall and Lerner (2010) and Kerr and Nanda (2015) for some surveys on the finance and innovation nexus. And along a similar topic, Kerr and Nanda (2009) review the literature on financing constraints and general entrepreneurship.

[9]See Seru (2014), Sevilir and Tian (2015), Ornaghi (2009), Haucap et al. (2019), **?** among others.

Provisions and find a positive effect on amount of innovation that is particularly pronounced in competitive markets and for firms with more information asymmetry. However, the innovation measures is often a count of patents or a citation weighted count of patents and the technological position and type of innovation is overlooked. Arora et al. (2018) develop a model to explain acquisition timing and the role of investment in absorptive capacity. Bena and Li (2014) and Hussinger (2010) are the closest in content to this study. They find that technological overlap between firm pairs increases the likelihood of an M&A deal. I supplement their contribution with a tailored measure of innovation proximity that captures technological complementarity and I further filter on deal pairs that involve a potential acquirer who is a large incumbent firm and a potential seller who is a young and small firm. This provides a better analysis of the motif that start-ups are bought out for innovation acquisition purposes.

Treating the technology in patents as the main dimension of interest, I will assume that a more original patent corresponds to a more differentiated product. Using patent data from Patstat, I measure patent originality and firm differentiation in terms of technological content. I will present results from some existing patent measures and explain their different interpretations then I will introduce some new changes to the measures. Firm differentiation (intuitively the opposite of firm "proximity") is defined based on firm patent portfolios and firm originality is the patent originality averaged to the firm level.[10]

I also use Patstat to identify new entrants with the assumption that firms that develop a new product will apply for patent protection before entering the market. Therefore my entry year is the first year of patenting. If the firm were to start selling before filing the patent, it could then be subject to reverse engineering and imitation. I assume that the set of firm's that enter the market before patenting is small. To identify firms that have been bought out, I use data from Thomson SDC Platinum. I then link the patenting behavior of the target and acquirer to patent applicants in Patstat using a customized fuzzy string matching algorithm based on firm names.

The empirical analysis focuses on two main variables, likelihood of buyout and innovation proximity. The analysis needs to be done with caution because I am positing that the innovation distance affects the likelihood of buyout but also that the likelihood of buyout affects the choice

[10]These measures are explained in more detail in the data section.

of innovation distance. However, this is in fact not an issue as I focus on new entrants. Before they start a research project, they do not have any apriori innovation measures. They do however have information on buyout trends, market sentiment, etc. as well as their financing options. Thus before a start-up comes into existence, its founders have beliefs on their likelihood of buyout. The hypotheses is that when the likelihood of buyout is low, new firms may believe their best option is to work on more original innovations and grow organically to eventually compete, while when the likelihood of buyout is high, new firms may be more incentivized to further increase their chances of buyout by innovating strategically closer to their potential acquirer.

As such, I ask two specific questions:

1. Can the proximity of a firm to another firm affect its likelihood of buyout?

2. Do the expectations of being bought out affect new entrants' innovation originality?

In order to first confirm that firms have a reason to believe their innovation positioning choices can affect their buyout likelihood, I build a firm pair dataset with a proximity-complementarity measure for each pair. I regress this complementary proximity measure on an indicator variable indicating whether the firm pair have had a buyout deal.

To address whether new entrants have indeed been changing their innovation behavior in response to their buyout expectations, I build another cross sectional dataset of firms in their first year of patenting. I then use a two step estimation model where I construct a measure of buyout expectations in the first step which I then use in the main regression on entrants' innovation choices. Using financing and macroeconomic variables to capture the conditions of the crisis and sector level concentration measures as controls, I extract a predicted number of buyouts by sector-year. I assume that this is a strong indicator for expectations of buyout and I use it as a proxy in the second step. With this proxy, I find that indeed a higher expectation of buyout decreases innovation originality in new entrants.

In the following section I will describe the setting of the financial crisis. Then in Section 3 I detail the datasets that I use and how the innovation measures were constructed. Section 4 then presents the empirical strategy, the main results and some robustness checks. Section 5 concludes.

2 The Setting

The Great Recession is characterized by the rupture of the subprime lending market, the use of unconventional policies and a prolonged period of low growth. I will investigate how this setting affected expectations on firm exit options. In particular, I suggest that the recessionary environment increased the chances of firm failure. However, conditional on survival, the likelihood of getting bought out increased. Buyouts involve a large sum of funds and are therefore sensitive to financing conditions. The crisis of 2008 was a shock on financial markets that spilled over to the entire economy. Normally in this situation, the Federal Reserve (Fed) would undertake expansionary monetary policy and lower the federal funds rate. However in the early 2000s, the Fed had already began decreasing the fed funds rate and there was not much room for manipulation by the time the crisis hit. As such, the Fed had to employ unconventional policies such as Quantitative Easing (QE) and forward guidance to boost the economy.

Monetary policy has traditionally had the effect of boosting household consumption by decreasing the interest rate to lower returns on savings and lower the cost of short term borrowing. QE, however, consists of large scale purchases of asset backed securities, collateralized debt obligations and other securitized instruments that put downward pressure on long term interest rates to further credit expansion. However long term debt is used for different purchases than short term debt. For households, long term debt is more likely to be used for automobile or house purchases (and student loans for students) - in general, large purchases. Yet the financial crisis was caused by easy credit for house purchases, therefore this effect was much more restrained. Although automobile loans and student loans did increase, this has arguably had a limited effect on the rest of the economy.

Instead I am putting forward that the principal effect of QE was through firms. Firms are entities that often have to make large purchases and investments that may be debt financed.[11] They have many reasons to take out long term debt such as for equipment purchases, R&D investments or simply because they have the means and the rate is low. In fact, the crisis saw a number of firms

[11] See Eaton et al. (2016)

take out debt to finance dividends or stock buybacks as well as firms that took advantage of the low rates to refinance their debt.

I further suggest that the effect of the crisis on firms was asymmetric. The severity of the crisis saw a high degree of economic uncertainty and risk aversion. It also raised awareness of issues in the financial system leading to financial regulatory reforms, such as Dodd Frank and Basel III, that included stricter rules on lending and the creation of a new macroprudential regulatory agency. This made it much more difficult for potential new firms to access financing. Small and young firms without collateral and established income streams found it particularly hard to access bank financing.[12] Furthermore since small business founders often use their house as collateral to access financing and housing prices fell drastically at the start of the crisis, new firms also experienced more limited access to financing through this channel as well.[13]

Since the crisis and following years was a time of high uncertainty, firms were less likely to invest in long term risky R&D projects. It was simply easier for start-ups to work on incremental innovation if they believed they were more likely to get acquired. In addition, in a recessionary setting, it is likely that firm survival was more difficult. New firms that choose to enter are likely to act strategically so as to decrease their likelihood of failure.[14] It was also easier for large existing firms to work on incremental products however they have the added option of using that money to acquire innovations instead. Since a new R&D project requires a large upfront fixed cost with the risk of being unsuccessful, a large firm might decide to take the less risky option and diversify its investments in multiple smaller R&D firms (for instance, through corporate venture capital) or to buyout new firms after they have successfully developed an innovation.

On a whole, the shock of the crisis and following policies clearly made firms reevaluate their decision making process and how they allocate investments. I will investigate whether the trends in buyouts changed and how that in turn affected the innovation choices of new entrants.

[12] See Ayyagari et al. (2018)

[13] See Davis and Haltiwanger (2019)

[14] See Cahn et al. (2019) for an evaluation of the effects of firm failure on the founders' future options.

3 The Data

My primary sources of data are Patstat for innovation measures and Thomson SDC Platinum for data on mergers and acquisitions. Below I discuss the data sources, the cleaning involved and the construction of the final datasets.

Patstat is a comprehensive database maintained by the European Patent Office (EPO) on patent applications and publications. It covers all major patent offices however I will be focusing on patents filed by companies who used an address in the United States. The database includes information on the applicants, inventors, application authority, filing dates, technology codes, whether it was granted, citations of other patents and of the non-patent literature, etc.. It also provides some constructed information such as industry codes and patent family identifiers as well as a preliminary applicant and inventor name cleaning because the information on applicants is subject to typos. Patstat also includes an educated guess on the type of applicant (ex. individual, company, university, etc.) which is what I use to primarily identify a firm (see Appendix A for more details).

A limitation of using patent data for my firm innovation measures is that I miss any firm innovation that has not been patented. The set of firms that patent is much smaller than the set of firms that do not patent. However this does not affect our results if there has not been a change in startup decisions to patent.

With a firm identified as a disambiguated company applicant, I construct its innovation measures.[15] I build a firm-level originality measure as well as a firm-pair-level proximity measure. The originality measure, already seen in Figure 1, was proposed by Trajtenberg et al. (1997) and is like a Herfindahl index:

$$Orig_p = 1 - \sum_{k \in \mathbb{K}} \left(\frac{Ncites_{p,k}}{Ncites_p} \right)^2$$

where $Ncites_{p,k}$ is the number of citations in technology class k from patent p and $Ncites_p$ is the number of patent p citations. This is simply a measure of concentration of the cited patents'

[15]See Appendix A for details on the applicant name cleaning and disambiguation.

technology codes with the implication being that a patent with more concentrated cited technology codes is less original. Originality is a patent level measure which I then aggregate to the firm-year level by taking the average.

Firm proximity is measured with respect to a firm pair following Jaffe (1986).

$$Prox_{i,j} = \frac{\mathbf{F}_i \cdot \mathbf{F}_j'}{\sqrt{\mathbf{F}_i \cdot \mathbf{F}_i'}\sqrt{\mathbf{F}_j \cdot \mathbf{F}_j'}}$$

where $\{i, j\}$ is a firm pair and $\mathbf{F}_i = (F_{i,1}, F_{i,2}, ...F_{i,K})$ is a vector of $F_{i,k}$, defined as the percent of firm i's patents that are in technology code k.

This proximity measure is essentially an uncentered correlation measure between two firms' patent shares in the different 4-digit IPC technology classes. The Jaffe measure however calculates the proximity only when two firms' technology codes overlap. In reality, certain technologies are more connected. Bloom et al. (2013) measure this connection through technology spillovers. They build a weighting matrix, Ω, from the covariance of the firm patent shares in each technology class.

$$Prox_{i,j} = \frac{\mathbf{F}_i \cdot \boldsymbol{\Omega} \cdot \mathbf{F}_j'}{\sqrt{\mathbf{F}_i \cdot \boldsymbol{\Omega} \cdot \mathbf{F}_i'}\sqrt{\mathbf{F}_j \cdot \boldsymbol{\Omega} \cdot \mathbf{F}_j'}} \tag{1}$$

I build the Bloom et al. (2013) measure however I also develop a different weighting matrix from *patent* level technology codes. By building Ω from the patent level, I capture the frequency that technology code pairs appear together in a patent. This more granular distinction better captures the technology codes that are complementary to each other since all the technology codes in a given patent are necessary for the invention in that patent. Building the weighting matrix at the firm level, also captures this effect however the measure is confounded if a firm has numerous product lines that are unrelated. I therefore suggest that building Ω from the patent level better captures complementarity between technology codes.

In practice, I build these measures from the 4-digit IPC codes. Since this aggregates multiple full IPC codes, a 4-digit IPC code can, and in fact does, appear multiple times in one patent. I keep all the repeated codes in the initial calculation to preserve the weights of each code. However this gives me a resulting matrix with a very heavy diagonal. Since the values along the diagonal will get confounded with the substitution effect I remove them and normalize the matrix. My final Ω weighting matrix is a measure of complementarity between technology codes which I then use in the construction of proximity to build a measure of complementarity between firms.

I also build a patent level measure of proximity where i is a patent and j is the set of patents cited by patent i. Since the citations of a patent consist of, in theory, the existing technologies at the frontier of the field of this patent, a proximity measure between the patent and its citations is similar to a measure of originality except it accounts for overlap between the patent's technology codes and its cited patents' technology codes. Intuitively, proximity should have an inverse relationship with originality. Indeed the trends are inversed for patent level proximity and originality as seen in Figure 2 below.

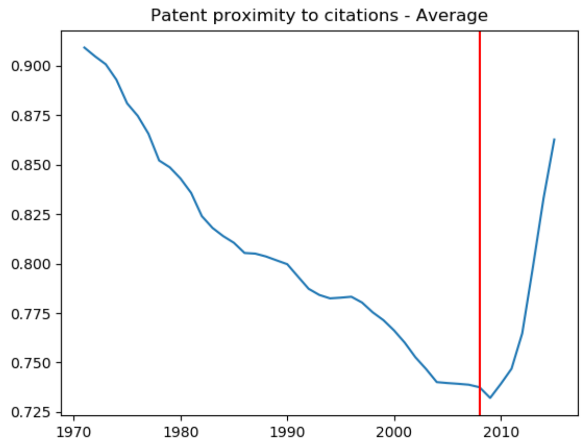


Figure 2: Average patent proximity to its citations in the US over time
The average patent proximity (from IPC codes) in the US. The red line is at the year 2008.

To summarize the time series trends in originality and to comment on the debate over whether entrants or incumbents are the most innovative, we can look at Figure 3 which plots the average originality for new firms and for twenty year old firms.[16] This shows that traditionally, new firms were more original than older firms although both have seen an upward trend. However in more recent years, new firm average originality has fallen below that of older firms and both groups

[16]The age of the firm is inferred from the the year the firm first begins patenting. See Chapter ?? for more discussion on this assumption.

are now showing a fall from their highs in the mid 2000s.

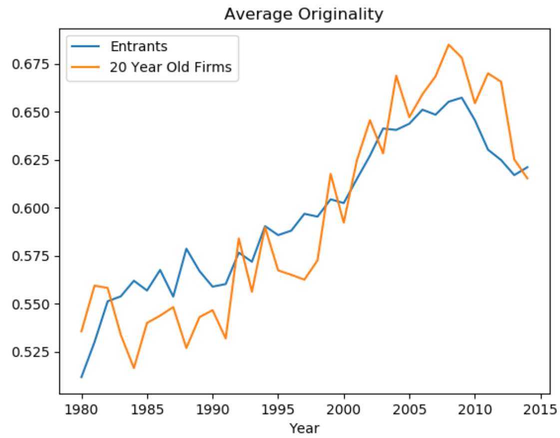


Figure 3: Average patent originality for incumbents and start-ups
Average 4-digit IPC code originality over time for start-ups and incumbent firms, where incumbent firms are defined as 20-year old firms.

If we breakdown the originality by quantiles in the information and communications technology sector (see Figure **??**), we see that the drop off occurs in all quantiles but is much steeper for the lower 25% quantile. The trend is in fact true for all firms however the reasons are different for new, small firms and large firms and I posit that their interaction is part of the reason that originality decreases in both groups.[17]

The acquisitions data comes from Thomson SDC Platinum (henceforth SDC) which offers detailed deal information such as target and acquiror names, address information, immediate and ultimate parents, industry codes, deal announcement date, effective deal date, whether the firm is a financial firm, the deal value, the percent of shares acquired, the source(s) of funding, etc. I extracted the deals involving only US targets as I am primarily interested in the innovation incentives of target US firms.

To connect target firms with their innovation behavior, I merge this with Patstat. Patstat, how-

[17]For instance, Zhao (2009) find that less innovative firms are more engaged in acquisitions.

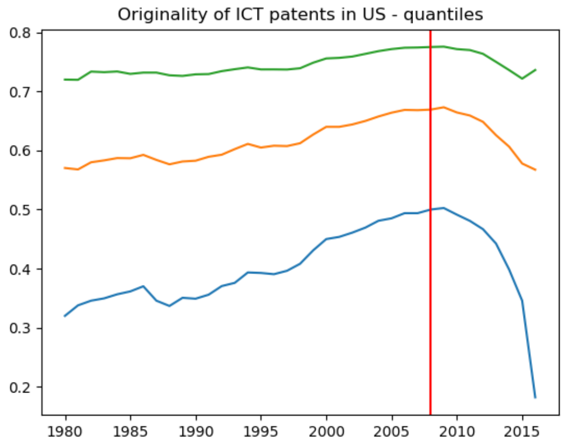


Figure 4: Average patent originality in the ICT sector
The originality quantiles (25th, 50th and 75th) for Information and Communications Technology firms in the US.

ever, does not use any official nor external firm identifier when noting applicant information. The applicant name is therefore prone to misspellings and errors even after the name cleaning done by the EPO. Without a concordance of applicants with an official data source, it is tricky to merge Patstat with any other datasets. The best we can do is to match firm names. This has been done with some other datasets such with French firms (see Aghion et al. (2019)) among others. However, to the best of our knowledge, this is the first time this has been done with Patstat and SDC. SDC uses firm identifiers to define a firm, so their firm names are also subject to some degree of misspellings and inconsistencies. There are different issues of matching firm names and these are discussed in detail in Appendix A. Due to the typos in firm names, I develop a fuzzy string matching algorithm to account for this. A fuzzy string matching algorithm however will inherently introduce errors into the dataset. It is a tradeoff between number of missed matches versus number of wrong matches. I do various checks to remove wrong matches such as checking for common words and matching addresses, however some error will always remain.

Another important issue with name matching is that firm names can change over time. And there is no standard on what name an entity within a firm group would use. Since we use firm names as our firm identifier, we cannot follow firms with name changes over time. This source of measurement error is not an issue in the main specification as I look only at firms in the first year they patent.[18]

Starting from around 147000 merger and acquisition deals in the US between 1990 and 2016, I remove deals where the acquirer was a financial company or an employee stock buyback etc. I also require that the deal resulted in a controlling majority share and count deals that were split into block share acquisitions as one. From Patstat I have about 151000 companies who filed a patent with an address in the US. After the merge process, I end up with 24347 deals with an acquirer who has patented, 28154 deals that involved a patenting target firm and 11393 deals where both the acquirer and target firms have patented and where the acquirer firm is large and the target firm is small.[19] I filter on large and small firms to capture the motive of buying technology and innovation as opposed to other reasons such as market share. The proportion matched seems small at first but this is roughly consistent with the proportion of firms that patent globally.

To address my two questions asked earlier, I will build two datasets:

- a firm pair level dataset with innovation proximity measures between the two firms
- a firm level dataset of new entrants with innovation measures at entry

The firm pair dataset will focus on the software industry to keep the analysis tractable.[20] I construct all possible firm pairs for firms in the software sector. As I am only interested in potential firm pairs that would have one firm acquired for its innovation (as opposed to a merger of equals or acquisitions for market share reasons), I keep only the firm pairs that involve one small firm, and where the firm size ratio is under 50%. I also remove firm pairs where both firms are in the top 1% of patenters as this might not be entirely captured by the firm size ratio restriction due to the skewed distribution of the firm size distribution.

From this smaller set of firm pairs, I build their innovation proximity measures and other innovation controls such as their originality, whether they have collaborated together before, direct

[18] The industry specific focus at the firm pair level however could be subject to this issue. This issue will also give me more entrant firms than in reality and it might give an upwards bias to my estimates later because I expect firms to have some path dependency in their R&D behavior.

[19] A large firm is defined as a firm in the top 10 percent of the firm size distribution where size is proxied by number of patents. Likewise, a small firm is a firm in the bottom 90% of the firm size distribution.

[20] I identify software firms by first identifying patents that are considered software patents following Bessen and Hunt (2007). Then I consider a firm a software firm if over 50% of its patents are software patents. Similarly, for ICT, I first identify patents that are ICT patents based on the OECD concordance with IPC codes then I consider a firm an ICT firm if over 50% of its patents are ICT patents

spillovers between the two firms and a commonality measure $Share\_common$ that Ornaghi (2009) suggests captures complementarity. However the $Share\_common$ measure does not take into account technology codes, it is simply a share of common cited patents over all cited patents. Therefore I prefer the proximity measure described in equation 1 to measure complementarity and I keep this $Share\_common$ measure as a control. In fact since this measure simply measures the share of cited patents the two firms have in common, this measure may capture substitutability more than complementarity. It is difficult to distinguish between the two and therefore I will present results with and without this measure. To measure the other spillovers, let us define P_i and P_j as the patents owned by firms i and j and B_i and B_j as the patents cited by firms i and j. The spillover controls are measured as:

$$Spill_{i,j} = \frac{||B_j \cap P_i||}{||B_j||} \tag{2}$$

$$Spill_{j,i} = \frac{||B_i \cap P_j||}{||B_i||} \tag{3}$$

$$Share\_common_{i,j} = \frac{||B_i \cap B_j||}{||B_j||} \tag{4}$$

The final firm pair dataset is very large and most of the firm pairs are not involved in a merger. To make this dataset tractable, I run the analysis on different random samples and the results are very stable between the different samples.

The new entrant, firm-level dataset is fairly straightforward to construct. I identify all the patent(s) filed in the first year of patenting for a company applicant with a US address. I use the first year and not simply the first patent because, depending on the industry, some products are composed of multiple patents. For these patents, I build their originality and patent-level proximity measures as described above, then I take an average to get the measures at the firm level.

I also include industry and year controls in this dataset. To identify the industry of the firm, I use the Nace code table in Patstat to convert to 2-digit SIC codes. The Nace code table includes a weighting of the codes the patent can be classified under which is calculated from its technology codes. I take the sum of all the patents a firm has at entry and their Nace code weightings and I consider the firm's primary industry to be the Nace code with the highest weight. Another issue with merging Patstat and SDC is that Patstat only provides NACE codes which are used primarily in Europe and SDC provides only codes used in the US, namely SIC and NAICS. I therefore had to use a concordance table to convert the applicant's NACE code to an SIC code. Since the classification between the two are quite different and uses different information content, I can only convert the NACE code to the broad 2-digit SIC codes.[21]

I also know the year the firm first applies for a patent. With this, I gather and merge data on the short and long term treasury rates, regulatory measures, house price index, the AAA and BAA spread, the implied volatility index (VIX), consumer confidence measure, stock market indices, unemployment rate, as well as other macroeconomic variables and sector measures such as the Herfindahl Index and the share of top 4 firms in a sector as defined by its 2-digit SIC code. These are controls for the financing, concentration and macroeconomic environment at the year of firm entry.

4 Empirical Strategy and Results

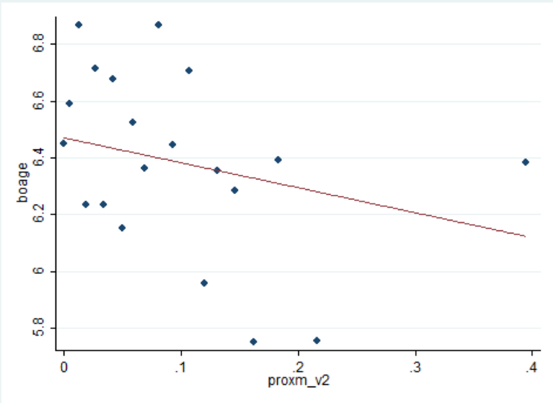
4.1 Firm pair proximity

Here we address the question of whether firm innovation positions can really affect their buyout likelihood. Figure 5 plots the buyout age of the target firm against its complementary proximity to the acquiring firm at buyout. The negative slope implies that firms are getting bought out faster when they are in closer proximity (more complementary) to their acquirers. The correlation is about -0.006.

Similarly, when we look at target firm originality we see the same result. On the left in figure 6, we have the maximum firm originality over its lifetime, while on the right in figure 7 I have the target firms' originality in the first year it enters. The effect is clearly positive for max originality but less clear for the originality of patents in the first year. The blue points represent the firm originality before 2008 and the red points represent the set after 2008. We see that the correlation becomes more positive after 2008 for firm initial originality. This implies that before 2008 new firms that entered didn't react to this strategy of closer positioning because they had less

[21]The conversion table is available upon request

Figure 5: Proximity with respect to firm buyout age



Binned scatter plot of firm complementary proximity at buyout vs target firm buyout age. This proximity measure is built from the set of patents the two firms have applied for up until the buyout year and uses the complementarity weighting described in the appendix. The buyout age is the difference between the buyout announcement year and the first year the target started patenting. The right most point seems to be an outlier but since this is a binned scatterplot, it aggregates multiple points. Most of the firms in this right-most group are pharmaceutical firms as their patents often consist of a similar set of technology codes.

reason to believe they would be bought out. However after 2008 they begin using this strategic channel as they believe their buyout likelihood has increased and that they can further influence their chances.

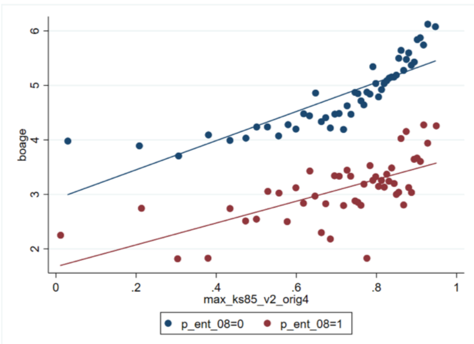


Figure 6: Maximum originality by firm buyout age

Binned scatter plot of firm age at buyout and the maximum firm originality achieved over its lifetime.

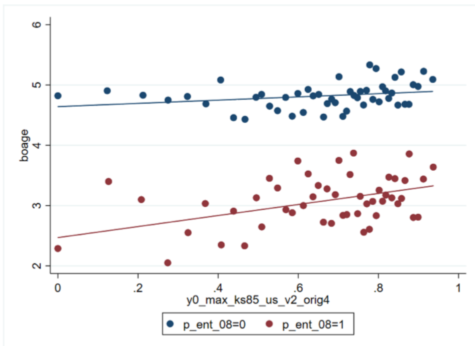


Figure 7: Initial firm originality by firm buyout age

Binned scatter plot of firm age at buyout and the firm originality in the first year it patents.

These figures are from the subset of deals that have been realized, we need to also consider the extensive margin and look at deals that have not happened yet. We also need to control for other factors. To do this I use the firm pair dataset described above on software firms.[22]

To address whether firms have reason to believe their innovation positioning has an effect on their buyout likelihood, I run a logit on a firm pair cross section.

$$\mathbb{1}[Firm\ i\ buys\ Firm\ j] = \alpha_0 + \alpha_1 Prox^0_{i,j} + \alpha_3 Controls^0_i + \alpha_4 Controls^0_j + \alpha_5 Controls^0_{i,j} + \varepsilon_{i,j} \quad (5)$$

Where firm i is the set of large firms in the top 10% for the firm size distribution and firm j are the other firms (the smaller firms in the bottom 90%). The firm size distribution is defined on the number of firm patent holdings. $Controls^0_i$ include the log knowledge stock of firm i which is defined as the aggregation of firm patents constructed using the usual inventory method with a depreciation rate of 15%. In a robustness check, the firm level controls also include financial measures such as total assets, number of employees and profits.

To avoid endogeneity I fix the right hand side variables over time intervals. Ideally I would take the pre-sample average of the proximity, originality, knowledge stock and firm pair spillover controls as my regressors. This means that I assume these values remained constant over the time period in my dataset. However my time period of 2000-2016 is quite long and firms are likely to have changed quite a bit over the time period. Instead I follow Prais (1958) and build two measures of each variable with each observation equally weighted. One average is constructed from the pre-sample 10 year period and one average is from the end of sample period. The end of sample average is taken from 2009 to 2016 to avoid a potential bias from the crisis in 2008.

Table 1 presents the results from the firm pair logit regression from equation 5. We indeed see that firm complementary proximity has a positive effect on likelihood of being bought out. This is consistent with figure 5 where we saw that firms get bought out faster when their complementary proximity is higher. The positive estimate on proximity is robust to different controls that are added. Table 1 also shows that the knowledge stock of a firm has a positive effect on the

[22] Software firms are defined as firms with a majority of software patents and software patents are defined following Bessen and Hunt (2007).

likelihood of getting bought out. If we included firm 2

	(1)	(2)	(3)	(4)	(5)
Proximity	0.9883***	0.9629***	0.9699***	0.9631***	0.9698***
	(0.0406)	(0.0420)	(0.0424)	(0.0422)	(0.0425)
Firm 1 knowledge stock	0.4724***	0.4773***	0.4765***	0.4773***	0.4765***
	(0.0148)	(0.0149)	(0.0149)	(0.0149)	(0.0149)
Firm 2 knowledge stock	0.3020***	0.2968***	0.2967***	0.2967***	0.2967***
	(0.0530)	(0.0527)	(0.0527)	(0.0527)	(0.0527)
Firm 1 originality			-1.3630***		-1.3631***
			(0.1025)		(0.1024)
Firm 2 originality				-0.0084	0.0039
				(0.0895)	(0.0895)
Collaborated		0.4475	0.4416	0.4475	0.4415
		(0.4741)	(0.4845)	(0.4741)	(0.4845)
Spill 2		60.2270**	61.3487**	60.2442**	61.3406**
		(28.5342)	(28.4264)	(28.5343)	(28.4268)
Spill 1		-58.1090***	-58.8728***	-58.1182***	-58.8684***
		(14.9743)	(15.0223)	(14.9740)	(15.0213)
Common cites		5.2639***	5.3635***	5.2646***	5.3631***
		(1.0970)	(1.0986)	(1.0968)	(1.0984)
Number of observations	168796	168793	168793	168793	168793

Table 1: Firm pair regressions with the complementarity proximity measure

This table contains the firm pair regressions with proximity calculated with the complementary weighting. The observations are on the software sector as defined following Bessen and Hunt (2007). Firm 1 is defined as large firms in the top 10% of the firm size distribution and firm 2 are the set of smaller firms in the bottom 90% of the firm size distribution. Both of the knowledge stock variables are logged with an adding 0.01 to avoid losing observations. A dummy variable is also added to control for those cases where the knowledge stock is zero.

From Table 1 we also see that the effect of the knowledge stock of firm i is large and significantly positive. Knowledge stock can be considered a proxy for firm size, so this suggests that buyout deals are more likely to come from larger firms. The coefficient on firm j originality is consistently negative albeit insignificant. This is consistent with the hypothesis that buyouts have a negative effect on target firm originality. Column (3) shows that the buyout likelihood is negatively affected by the larger firm's originality yet positively affect by its knowledge stock. This implies that the larger firm (the potential acquiror) is more likely to acquire another firm when it has historically invested a lot in R&D yet has a low degree of originality. The coefficient on firm 2 (the smaller firm and potential target) originality is insignificant. Since we expect proximity to capture the majority of the type of innovation choices, it is not surprising that the originality of the smaller firm is insignificant. When we include the other spillover measures, we see that Spill 2, the percent of firm 2 patents that are cited by firm 1 has a positive effect on buyout likelihood. However Spill 1, the percent of firm 1 patents cited by firm 2 has a negative and significant effect on buyout likelihood. When the two firms have a higher amount of common cited patents, they also increase their likelihood of buyout.

On the other hand, when we regress the likelihood of buyout on the Jaffe proximity (which measures proximity along the substitutability axis) we see that proximity has a negative and significant coefficient (see table 2. This result contrasts with the findings in Bena and Li (2014) and Hussinger (2010) who find a positive effect with the Jaffe measure. This discrepancy is likely due to my specification focusing on large-small firm pairs instead of all firm pairs. Small firms that are closely positioned along the substitutability axis to large firms have to compete more directly with the large firms. A buyout along substitutable firms primarily occurs in order to gain market share and eliminate a rival. My dataset here has specifically chosen firm 2 to be small firms and thus they are unlikely rivals to the large firms right away. In this case, large firms have different strategies they can use to deal with the competition from small firms and although an acquisition is an option, here we see that it is not likely. If the firm pairs also included large-large firm pairs, it is possible that a closer proximity will increase the likelihood of an M&A deal. The estimates on the other measures, namely knowledge stock, originality, and spillovers remain consistent.

The different result between these two tables highlight that the complementary proximity measure in table 1 captures a different interaction between the small firm and large firm. They are less likely to be rivals. And the complementary technologies may imply that synergies can be found with a buyout. In general, we have seen that there is some reason for a firm to expect its innovation positioning can affect its likelihood of getting bought out. A firm that is closely positioned in complementary technological areas increases its likelihood of buyout while a firm that is closely positioned in substitutable areas has a lower likelihood of getting acquired.

4.2 Firm entry innovation

To address my central question of how buyout beliefs affect new entrant innovation originality, I first develop a model for buyout expectations. New entrants do not expect to get acquired immediately after they enter the market. The average age of target firms when they are acquired is 9.5 years while the median is 7 years. Since my focus is on deals involving a start-up acquisition,

	(1)	(2)	(3)	(4)	(5)
Proximity	-2.3787***	-2.3825***	-2.4398***	-2.3825***	-2.4398***
	(0.0879)	(0.0882)	(0.0887)	(0.0882)	(0.0887)
Firm 1 knowledge stock	0.4408***	0.4434***	0.4379***	0.4434***	0.4379***
	(0.0126)	(0.0126)	(0.0127)	(0.0126)	(0.0127)
Firm 2 knowledge stock	0.2900***	0.2776***	0.2785***	0.2777***	0.2787***
	(0.0411)	(0.0413)	(0.0417)	(0.0413)	(0.0417)
Firm 1 originality			-1.8790***		-1.8794***
			(0.1111)		(0.1111)
Firm 2 originality				0.0154	0.0258
				(0.0938)	(0.0936)
Collaborated		1.5623***	1.5617***	1.5617***	1.5607***
		(0.4943)	(0.5060)	(0.4943)	(0.5059)
Spill 2		67.1130**	68.9348**	67.0820**	68.8833**
		(29.4007)	(29.3384)	(29.4015)	(29.3394)
Spill 1		-75.3877***	-76.3697***	-75.3653***	-76.3321***
		(12.7034)	(12.6561)	(12.7005)	(12.6503)
Common cites		9.2161***	9.3445***	9.2133***	9.3397***
		(1.5300)	(1.5255)	(1.5297)	(1.5248)
Number of observations	115506	115503	115503	115503	115503

Table 2: Firm pair regressions with the Jaffe proximity measure

This table contains the firm pair regressions on the software sector with the Jaffe proximity measure. Both of the knowledge stock variables are logged with an adding 0.01 to avoid losing observations. A dummy variable is also added to control for those cases where the knowledge stock is zero. Firm 1 is defined as large firms in the top 10% of the firm size distribution and firm 2 are the set of smaller firms in the bottom 90% of the firm size distribution. The observations are on the software sector as defined following Bessen and Hunt (2007).

I remove all deals with a target firm above 10 years old.[23] As such, the average buyout age is 4.7 and the median is 4. This implies that new entrants will base their entry decisions on their buyout beliefs at least a few years into the future.

There is, unfortunately, no consensus on how to model firm expectations. Landier et al. (2019) provide some discussion and experimental evidence comparing rational expectations with interpolation and extrapolation. They find that extrapolation is the most prominent while rational expectations is the least realistic. Kuchler and Zafar (2019) also find that extrapolation matches best with survey evidence. There is a discussion on models of expectations formation in macroeconomics as well. Although their models are usually focused on inflation expectations, they also find that the full information rational expectations model is often mismatched with reality.[24] There is also a financial economics and behavioral economics literature on expectations and learning with many different models put forward.[25]

Here I will build a simple reduced form expectations model based on the extrapolation concept where I define the information set of the entrant firms as a set of variables that characterize the 2008 crisis. I assume that the potential firm entrant already knows what industry it will enter in and the strategic innovation decision is made within that industry on the technological class. In particular, since I want to build an expectations measure for firms before they enter, I do not have any information on firm specific characteristics. As such, my expectations measure will be formed at the industry level. Specifically, I assume that prior to entering the market, all potential new firms within an industry have the same expectations and that their expectations are based on a common set of macroeconomic, financial, and industry specific data points.

Let $\mathcal{F}_t$ be the information set of all potential entrants at time t. This includes data such as past buyout deal details as well as historical short and long term interest rates, financial regulation changes and a house price index to proxy financing conditions plus macroeconomic measures and industry level concentration. To capture regulatory changes I use the number of restric-

[23] An older target firm also implies that the firm has an established market share and that it is more likely to be bought out for market share reasons rather than R&D reasons.

[24] See Woodford (2013), Coibion et al. (2018), Negro and Schorfheide (2004), Davila (2014), Bordalo et al. (2018), etc.

[25] See Fudenberg and Levine (2016), Heidhues et al. (2018), Gilboa et al. (2008), Gilboa (2014), Diecidue and de Ven (2008), etc.

tions in financial titles collected and parsed by RegData. The macroeconomics measures include the VIX as a volatility indicator, a measure of consumer sentiment from the OECD, the S&P 500 index as a measure of stock market sentiment, the inflation rate and the unemployment rate. As buyouts may be more likely to happen in different times in the industry life cycle, I control for this with the Herfindahl index and the share of sales by the top ten firms in each industry.

I assume that the expectations of buyout is a linear model of the number of buyouts in the same industry as the potential entrant. Let $Y_{s,t}$ be the number of buyouts in industry s in year t. What I want to predict is :

$$\mathbb{E}(Y_{s,t+\gamma}|\mathcal{F}_t;\beta^{(t,\gamma)}) \tag{6}$$

where γ is the number of years ahead predicted and $\beta^{(t,\gamma)}$ is the set of parameters at time t for γ years ahead. Since $\beta^{(t,\gamma)}$ is unobserved, I estimate it with the information available at t. Namely:

$$\hat{\beta}^{(t,\gamma)} = \min_{\beta}(Y_{s,t} - \mathbb{E}(Y_{s,t}|\mathcal{F}_{t-\gamma};\beta^{(t,\gamma)}))^2 \tag{7}$$

Assuming that $\mathbb{E}(Y_{s,t}|\mathcal{F}_{t-\gamma};\beta^{(t,\gamma)})$ is linear, equation (7) can be concretely rewritten as:

$$\textit{Num buyouts}_{s,t} = \beta_0^{(t,\gamma)} + \beta_f^{(t,\gamma)}\textit{Financing measures}_{t-\gamma} + \beta_m^{(t,\gamma)}\textit{Macro controls}_{t-\gamma} + \beta_s^{(t,\gamma)}\textit{Sector controls}_{s,t-\gamma} + \epsilon_{s,t}$$

Since the estimates from equation (7) are used in the main regression, I need a source of exogenous variation. The variables in the information set are mostly the same as the variables that I use in the main regression as controls. I gain some additional variation by including the lagged number of buyout deals in this first stage regression. The previous number of buyout deals should not have any direct effect on the entrant firm's innovation choice except through its buyout expectations. For another source of variation, I also run a robustness check where my first step estimate is calculated with an added second lag on the variables. Specifically, let $\mathcal{F}_t = \{f_t, f_{t-1}, f_{t-2}, ...\}$ where f_t is the information arriving in year t. Then my $\hat{\beta}^{t,\gamma}$ is estimated from :

$$\hat{\beta}^{(t,\gamma)} = \min_{\beta}(Y_{s,t} - \mathbb{E}[Y_{s,t}|f_{t-\gamma}, f_{t-1-\gamma};\beta^{(t,\gamma)}])^2 \tag{8}$$

This gives an estimate of $\hat{\beta}^{(t,\gamma)}$ which is plugged into eq (6) to give the predicted number of buyout deals in year $t+\gamma$. To be clear, my predicted number of buyouts at year t for year $t+\gamma$ is :

$$\widehat{Num\ buyouts}_{s,t+\gamma} = \hat{\beta}_0^{(t,\gamma)} + \hat{\beta}_f^{(t,\gamma)} Financing\ measures_t + \hat{\beta}_m^{(t,\gamma)} Macro\ controls_t + \hat{\beta}_s^{(t,\gamma)} Sector\ controls_{s,t} + \eta_{s,t}$$

This step is run multiple times with different γ lag years (ex. $0, 1, ..., 5$). Having obtained these predicted number of buyouts, I return to the main question of how expectations affect new firm entrants' innovation originality. My main specification is:

$$Originality_{i_{s,t}} = \mathbb{E}_{s,t}[i \text{ will be bought out}] + Financing\ measures_t + Sector\ Controls_{s,t} + Macro\ Controls_t + \upsilon_i \tag{9}$$

Where $Originality_{i_{s,t}}$ is the average originality of the firm entrant i in industry s in the first year it enters t. Financing measures include interest rates, regulatory restrictions and the house price index and the sector and macro controls are the same as the set in the estimation of β in eq (8). I assume: $\mathbb{E}_{s,t}[i \text{ will be bought out}] = \phi_0 + \phi_1 \widehat{Num\ of\ buyouts}_{s,t+\gamma} + \zeta_{s,t}$

Table 3 presents the results from the regression as described in equation (7). We expect that financing conditions should be a major predictor and we indeed see that the federal funds rate (a.k.a. the overnight borrowing rate) is highly negative and significant with the effect becoming slightly less significant in higher lead years. The treasury 10 year rate is weakly negative and significant here. The financial regulatory restrictions on the other hand, have an ambiguous effect across the lead years. We expected the regulatory restrictions to have more of an effect on the young firms and we will clearly see that this is the case later. The fact that the coefficient fluctuates here implies that the negative effect of more regulatory restrictions is entangled with the positive effect of the asymmetric pass-through.

The lagged number of buyouts is the most significant predictor of future buyouts. The coefficient stays positive and significant over the different lead years I use. This implies that momentum is an important cause of buyouts. When there are more buyouts one year, there is likely to be more next year as well. The M&A literature has suggested that when a deal happens between two firms in an industry, the competition landscape changes and spurs the other firms in that industry to also do deals to remain competitive. Note that this regression is over the entire time period (1980-2016) thus the R^2 is quite high. However the predicted number of buyouts measure I use in the second stage is from rerunning the regression each year with only data up until that year. Evidently the out-of-sample fit is worse although it gets better over time as more data becomes available.

	Number of Buyouts				
	t+1	t+2	t+3	t+4	t+5
Number of Buyouts t-1	9.924e-01***	9.804e-01***	9.740e-01***	9.647e-01***	9.535e-01***
	(3.997e-02)	(5.584e-02)	(6.336e-02)	(5.781e-02)	(4.999e-02)
Fed Funds rate t-1	-4.212e+00***	-8.865e+00***	-1.282e+01***	-1.112e+01***	-7.331e+00*
	(1.224e+00)	(1.935e+00)	(2.591e+00)	(3.301e+00)	(3.756e+00)
Treasury 10yr rate t-1	6.560e+00**	6.086e+00	1.377e+01**	1.034e+01	8.125e+00
	(2.871e+00)	(4.943e+00)	(6.791e+00)	(8.580e+00)	(1.031e+01)
Regulatory Restrictions t-1	4.630e+03	-1.814e+04	-2.810e+04	-3.588e+04*	-1.659e+04
	(8.538e+03)	(1.294e+04)	(1.774e+04)	(1.928e+04)	(2.148e+04)
Concentration t-1	7.621e-01	1.744e+00**	2.614e+00**	3.629e+00***	5.021e+00***
	(6.244e-01)	(8.790e-01)	(1.040e+00)	(1.069e+00)	(1.229e+00)
House Price t-1	-8.341e-02*	-2.825e-01***	-3.180e-01***	-2.194e-01*	5.949e-02
	(4.505e-02)	(7.941e-02)	(9.440e-02)	(1.176e-01)	(1.215e-01)
Nasdaq t-1	-1.351e-02***	-2.212e-02***	-1.224e-02	8.718e-03	2.781e-02***
	(4.061e-03)	(6.867e-03)	(9.050e-03)	(7.708e-03)	(7.708e-03)
Volatility t-1	-7.227e-01***	-7.771e-01***	-9.966e-01***	-1.181e+00***	-9.280e-01*
	(1.691e-01)	(2.199e-01)	(2.834e-01)	(3.786e-01)	(5.273e-01)
Consumer Confidence t-1	5.085e+00***	5.663e+00**	-9.173e-01	-1.456e+01***	-3.107e+01***
	(1.791e+00)	(2.840e+00)	(3.752e+00)	(4.829e+00)	(5.749e+00)
Inflation t-1	6.390e-01**	4.752e-01	3.613e-01	-4.288e-01	-1.088e+00*
	(2.568e-01)	(3.549e-01)	(5.227e-01)	(5.583e-01)	(5.764e-01)
Unemployment t-1	1.306e+00	-3.181e+00*	-3.716e+00*	-2.300e+00	-4.906e-01
	(1.521e+00)	(1.810e+00)	(2.095e+00)	(2.719e+00)	(3.170e+00)
Oil Price t-1	5.402e-03	3.742e-01***	2.000e-01	-2.866e-01	-1.013e+00***
	(9.329e-02)	(1.393e-01)	(1.575e-01)	(2.169e-01)	(2.787e-01)
N	1912	1849	1779	1714	1648

Table 3: Regression output from the first step based on OLS

Note: All the regressor variables are lagged by one year. This is a linear regression at the sector-year level and the number of buyouts are by sector year where a sector is a 2-digit SIC code. The concentration index is the Herfindahl index and the various indices that come in daily or monthly or quarterly frequencies have been averaged to the yearly frequency. The standard errors are in the parenthesis.

Since the dependent variable in the first stage is a count of deals, I run a robustness check with a negative binomial regression with the output in Table 4. The results are largely consistent with the OLS regression.

Table 4 presents the main regression described in equation 9. We see clearly that the expected number of buyouts in the firms' primary sector has a negative and significant effect on its originality confirming our hypothesis that increased beliefs of buyout likelihood lead new firms to conduct less original innovation. The effect is slightly weaker as the lead years increase but this

	Number of Buyouts				
	t+1	t+2	t+3	t+4	t+5
Number of Buyouts t-1	6.635e-03***	6.599e-03***	6.577e-03***	6.595e-03***	6.573e-03***
	(5.060e-04)	(5.330e-04)	(5.510e-04)	(5.760e-04)	(5.950e-04)
Fed Funds rate t-1	-3.060e-02	-7.932e-02***	-1.001e-01***	-8.369e-02***	-5.748e-02**
	(2.020e-02)	(2.160e-02)	(2.550e-02)	(2.730e-02)	(2.700e-02)
Treasury 10yr rate t-1	4.300e-02	5.380e-02	8.310e-02	5.440e-02	4.450E-02
	(4.490e-02)	(4.760e-02)	(5.430e-02)	(5.980e-02)	(6.270e-02)
Regulatory Restrictions t-1	-1.790e+01	-2.635e+02**	-4.222e+02***	-2.975e+02**	6.360e+00
	(1.100e+02)	(1.240e+02)	(1.310e+02)	(1.370e+02)	(1.520e+02)
Concentration t-1	9.299e-02***	9.338e-02***	9.271e-02***	9.240e-02***	9.347e-02***
	(6.120e-03)	(6.710e-03)	(7.110e-03)	(7.660e-03)	(8.420e-03)
House Price t-1	-1.620e-03**	-3.004e-03***	-3.233e-03***	-2.092e-03**	-2.370E-05
	(7.250e-04)	(8.230e-04)	(8.710e-04)	(8.990e-04)	(9.930e-04)
Nasdaq t-1	-1.854e-04***	-1.693e-04***	-1.670e-05	1.450e-04***	2.934e-04***
	(3.770e-05)	(4.620e-05)	(5.180e-05)	(5.500e-05)	(5.700e-05)
Volatility t-1	-6.749e-03**	-7.565e-03**	-9.916e-03***	-1.041e-02***	-7.791e-03*
	(2.690e-03)	(2.980e-03)	(3.160e-03)	(3.520e-03)	(4.160e-03)
Consumer confidence t-1	7.844e-02**	5.700e-02	-2.410e-02	-1.593e-01***	-3.037e-01***
	(3.320e-02)	(3.780e-02)	(4.000e-02)	(4.340e-02)	(4.520e-02)
Inflation t-1	1.089e-02***	6.069e-03*	1.770e-03	-3.790e-03	-8.679e-03**
	(3.390e-03)	(3.500e-03)	(3.800e-03)	(4.150e-03)	(4.120e-03)
Unemployment t-1	-2.010e-02	-5.394e-02**	-5.031e-02*	-3.510e-02	-2.540e-02
	(2.590e-02)	(2.670e-02)	(2.820e-02)	(2.960e-02)	(3.170e-02)
Oil Price t-1	-9.770e-04	1.530e-03	-1.070e-03	-5.425e-03**	-9.286e-03***
	(1.650e-03)	(2.160e-03)	(2.230e-03)	(2.500e-03)	(2.740e-03)
N	1.912e+03	1.849e+03	1.779e+03	1.714e+03	1.648e+03

Table 4: Regression output from the first step based on a negative binomial model

Note that all the regressor variables are lagged by one year. This is a negative binomial regression at the sector-year level and the number of buyouts are by sector year where a sector is a 2-digit SIC code. The concentration index is the Herfindahl index and the various indices that come in daily or monthly or quarterly frequencies have been averaged to the yearly frequency. The standard errors are in the parenthesis.

is likely due to the higher prediction error for higher lead years.

We also see that the fed funds rate, a principle measure of start-up access to financing, has a negative effect on innovation originality as expected. A higher interest rate means it is more costly to borrow and therefore makes R&D more difficult.The 10 year Treasury rate is also consistently negative albeit insignificant. Similarly more financial regulatory restrictions have a negative and significant effect on new firm innovation. Although the financial regulatory restrictions are only applied directly to financial intermediaries, we see evidence here that it is passed on to their borrowers as well.

Finally, the house price index is also a measure of young firm access to financing. Here we see that the coefficient on the house price index is positive which is inline with the intuition that higher house prices mean more collateral value which allows more access to debt capital and hence leads to more original innovation. The interest rates, regulatory restrictions and house price index have a push effect on innovation that I expect would also increase the rate of patenting while the buyout expectations measure captures the pull effect described earlier.

	Firm Entry Originality				
E[Num Buyouts$_{t+1}$]	-2.512e-05***				
	(1.064e-05)				
E[Num Buyouts$_{t+2}$]		-1.137e-05***			
		(3.985e-06)			
E[Num Buyouts$_{t+3}$]			-1.248e-05**		
			(4.979e-06)		
E[Num Buyouts$_{t+4}$]				-1.142e-05*	
				(6.661e-06)	
E[Num Buyouts$_{t+5}$]					3.120e-06
					(3.147e-06)
Fed Funds Rate	-2.762e-03***	-2.667e-03***	-2.947e-03***	-3.088e-03***	-2.710e-03***
	(8.035e-04)	(8.055e-04)	(8.048e-04)	(8.247e-04)	(8.106e-04)
Treasury 10 year rate	-4.503e-04	-1.804e-03	-8.723e-04	-3.355e-05	-1.066e-03
	(1.977e-03)	(1.984e-03)	(1.962e-03)	(2.057e-03)	(1.962e-03)
Regulatory Restrictions	-1.270e-06***	-1.365e-06***	-1.367e-06***	-1.235e-06***	-1.276e-06***
	(2.266e-07)	(2.308e-07)	(2.321e-07)	(2.262e-07)	(2.289e-07)
House Price Index	2.799e-04***	2.867e-04***	2.731e-04***	2.859e-04***	3.148e-04***
	(4.030e-05)	(3.931e-05)	(4.098e-05)	(4.093e-05)	(3.907e-05)
Nasdaq Avg	-4.190e-06**	-2.567e-06	-3.027e-06*	-2.590e-06	-1.609e-07
	(2.136e-06)	(1.703e-06)	(1.799e-06)	(1.852e-06)	(1.784e-06)
Volatility Avg	1.563e-04	1.788e-04	4.480e-05	-4.492e-06	1.496e-04
	(1.549e-04)	(1.552e-04)	(1.608e-04)	(1.818e-04)	(1.548e-04)
Consumer Confidence Avg	-4.360e-04	-2.686e-04	-2.166e-04	-1.491e-03	-2.038e-03
	(1.617e-03)	(1.593e-03)	(1.626e-03)	(1.493e-03)	(1.479e-03)
Inflation Index	6.065e-04***	4.231e-04**	4.889e-04**	5.159e-04**	3.001e-04
	(2.222e-04)	(1.948e-04)	(2.007e-04)	(2.175e-04)	(1.992e-04)
Unemployment Rate	-1.621e-03*	-1.083e-03	-1.044e-03	-1.530e-03	-8.274e-04
	(9.844e-04)	(9.674e-04)	(9.682e-04)	(9.936e-04)	(1.023e-03)
Oil Price	-1.749e-04*	-9.943e-05	-5.724e-05	-1.335e-04	-1.240e-04
	(9.912e-05)	(9.940e-05)	(1.037e-04)	(9.820e-05)	(9.949e-05)
N	154955	154955	154955	154955	154955

Table 5: Baseline regressions results

Stage 2 regression on 4-digit firm originality in the first year it patents. Also included in the regressors are 2-digit SIC controls as well as controls for whether the firm is in the ICT or software sector. Standard errors are clustered on 2-digit SIC sectors. Expected number of buyouts are measured in the first stage as described above. All time-varying RHS variables are lagged one year.

4.3 Robustness Check

Table 6 is a robustness check of the main result with the alternative proximity measure. Here the dependent variable is replaced by a measure built using the same methodology as proximity between firms. In this case, it is proximity between the technology codes in the firm's patents vs the cited patents. We expect proximity to have an inverse effect as compared to originality and indeed in Table 6 the sign of the coefficient on buyout expectations is now positive. Different constructions of the dependent variable are also tested and available in Appendix B. The results are generally very consistent.

	Firm Entry Proximity				
E[Num Buyouts t + 1]	1.709e-05***				
	(5.280e-06)				
E[Num Buyouts t + 2]		1.914e-05***			
		(2.589e-06)			
E[Num Buyouts t + 3]			1.005e-05***		
			(3.191e-06)		
E[Num Buyouts t + 4]				8.484e-06**	
				(3.688e-06)	
E[Num Buyouts t + 5]					1.491e-05***
					(4.013e-06)
Fed Funds Rate t - 1	7.317e-03***	7.265e-03***	7.356e-03***	7.315e-03***	7.396e-03***
	(8.568e-04)	(8.405e-04)	(8.468e-04)	(8.541e-04)	(8.636e-04)
Treasury 10yr Rate t - 1	-7.510e-03**	-6.919e-03**	-7.562e-03**	-7.677e-03**	-7.406e-03**
	(2.741e-03)	(2.724e-03)	(2.720e-03)	(2.710e-03)	(2.757e-03)
Regulatory Restrictions t - 1	4.242e-06***	4.283e-06***	4.271e-06***	4.247e-06***	4.193e-06***
	(3.075e-07)	(3.055e-07)	(3.049e-07)	(3.064e-07)	(3.084e-07)
House Price t - 1	-1.785e-04***	-1.791e-04***	-1.724e-04***	-1.715e-04***	-1.730e-04***
	(4.657e-05)	(4.527e-05)	(4.431e-05)	(4.405e-05)	(4.509e-05)
Oil Price t - 1	7.666e-04***	7.263e-04***	7.406e-04***	7.641e-04***	7.789e-04***
	(1.047e-04)	(1.043e-04)	(1.075e-04)	(1.044e-04)	(1.025e-04)
Nasdaq t - 1	5.175e-06***	5.314e-06***	5.228e-06***	5.324e-06***	5.423e-06***
	(8.026e-07)	(7.679e-07)	(8.113e-07)	(8.121e-07)	(7.271e-07)
Volatility t - 1	5.593e-04**	5.237e-04**	5.956e-04***	6.159e-04***	5.536e-04**
	(1.975e-04)	(1.937e-04)	(1.861e-04)	(1.809e-04)	(1.982e-04)
Consumer Confidence t - 1	6.763e-03***	6.052e-03***	6.680e-03***	7.068e-03***	6.888e-03***
	(2.014e-03)	(1.904e-03)	(1.948e-03)	(1.848e-03)	(1.876e-03)
Inflation t - 1	-1.163e-03***	-1.118e-03***	-1.165e-03***	-1.189e-03***	-1.180e-03***
	(2.077e-04)	(2.011e-04)	(2.010e-04)	(1.959e-04)	(2.047e-04)
Unemployment Rate t - 1	1.037e-02***	1.011e-02***	1.028e-02***	1.046e-02***	1.069e-02***
	(1.333e-03)	(1.346e-03)	(1.369e-03)	(1.386e-03)	(1.385e-03)
Sector controls	yes	yes	yes	yes	yes
N	125389	125389	125389	125389	125389

Table 6: Robustness check using the alternative patent proximity measure

Stage 2 regression on firm entry proximity in the first year it patents with 2-digit SIC controls as well as controls for whether the firm is in the ICT or software sector. The proximity measure is akin to an uncentered correlation measure of the technology codes in the patents held by the firm and the technology codes of the patents cited by the firm. Standard errors are clustered on 2-digit SIC sectors. Expected number of buyouts are measured in the first stage as described above. All time-varying RHS variables are lagged one year.

An alternative theory for why we might be seeing a drop in originality is that perhaps I am capturing some spurious effect due to changes in firm patenting strategies. There is some anecdotal evidence that some firms are choosing to protect their inventions by filing more patents of a smaller scope. This arguably increases the chances of at least one patent being granted. Since the Trajtenberg et al. (1997) originality measure is built at the patent level and then averaged to the firm level, this possibility would bias my results. In order to adjust for this, I simply build an additional originality measure at the firm level directly instead of at the patent-level firm. This means I aggregate all the firms patents in its first year and get the set of technology codes of the backwards citations, then I build the originality with the same formula as before. In this way, I group all the patents and their technology codes into a given firm-year so this measure should not be affected by the changes in patenting strategies just described. Tables 7 and 8 are the results for firm-level originality and firm-level proximity and again we find similar results.

	Firm Entry Average Firm-level Originality				
E[Num Buyouts$_{t+1}$]	-5.930e-05***				
	(1.008e-05)				
E[Num Buyouts$_{t+2}$]		-1.323e-05***			
		(3.759e-06)			
E[Num Buyouts$_{t+3}$]			-1.893e-05***		
			(4.725e-06)		
E[Num Buyouts$_{t+4}$]				-2.983e-05***	
				(6.296e-06)	
E[Num Buyouts$_{t+5}$]					-1.693e-06
					(2.956e-06)
Fed Funds Rate	-1.947e-03**	-1.897e-03**	-2.285e-03***	-2.882e-03***	-2.147e-03***
	(7.585e-04)	(7.602e-04)	(7.599e-04)	(7.784e-04)	(7.651e-04)
Treasury 10 year rate	-1.032e-03	-3.355e-03*	-2.216e-03	3.077e-04	-2.407e-03
	(1.851e-03)	(1.855e-03)	(1.836e-03)	(1.927e-03)	(1.836e-03)
Num Regulatory Restrictions	-1.077e-06***	-1.152e-06***	-1.202e-06***	-1.008e-06***	-9.853e-07***
	(2.094e-07)	(2.134e-07)	(2.148e-07)	(2.091e-07)	(2.117e-07)
House Price Index	1.163e-04***	1.600e-04***	1.312e-04***	1.223e-04***	1.825e-04***
	(3.809e-05)	(3.716e-05)	(3.881e-05)	(3.869e-05)	(3.691e-05)
Nasdaq Avg	-6.536e-06***	-5.843e-07	-1.823e-06	-2.928e-06*	9.094e-07
	(2.030e-06)	(1.610e-06)	(1.705e-06)	(1.754e-06)	(1.685e-06)
Volatility Avg	-2.730e-05	-8.790e-06	-2.064e-04	-4.654e-04***	-3.908e-05
	(1.466e-04)	(1.469e-04)	(1.521e-04)	(1.720e-04)	(1.465e-04)
Consumer Confidence Avg	2.293e-03	6.316e-04	1.259e-03	-2.695e-04	-1.263e-03
	(1.538e-03)	(1.511e-03)	(1.544e-03)	(1.416e-03)	(1.403e-03)
Inflation Index	1.022e-03***	4.853e-04***	6.124e-04***	8.527e-04***	4.258e-04**
	(2.104e-04)	(1.839e-04)	(1.897e-04)	(2.055e-04)	(1.880e-04)
Unemployment Rate	-3.031e-03***	-1.815e-03**	-1.746e-03*	-2.965e-03***	-2.099e-03**
	(9.349e-04)	(9.187e-04)	(9.192e-04)	(9.435e-04)	(9.699e-04)
Oil Price	-2.504e-04***	-1.185e-04	-4.195e-05	-1.546e-04*	-1.766e-04*
	(9.425e-05)	(9.443e-05)	(9.852e-05)	(9.334e-05)	(9.447e-05)
N	143915	143915	143915	143915	143915

Table 7: Robustness check using the firm-level originality measure

Stage 2 regression on 4-digit mean firm US patenting firm level originality in the first year it patents. Also included in the regressors are 2-digit SIC controls as well as controls for whether the firm is in the ICT or software sector. Standard errors are clustered on 2-digit SIC sectors. The expected number of deals are . All time-varying RHS variables are lagged one year.

	Firm Entry Average Firm-level Proximity				
E[Num Buyouts$_{t+1}$]	8.187e-05***				
	(1.154e-05)				
E[Num Buyouts$_{t+2}$]		1.672e-05***			
		(4.259e-06)			
E[Num Buyouts$_{t+3}$]			1.266e-05**		
			(5.445e-06)		
E[Num Buyouts$_{t+4}$]				8.124e-06	
				(7.081e-06)	
E[Num Deals$_{t+5}$]					-9.189e-06***
					(3.368e-06)
Fed Funds Rate	2.485e-03***	2.436e-03***	2.794e-03***	2.863e-03***	2.354e-03***
	(8.771e-04)	(8.782e-04)	(8.797e-04)	(8.908e-04)	(8.802e-04)
Treasury 10 year rate	-1.157e-03	1.943e-03	6.419e-04	7.097e-05	8.483e-04
	(2.073e-03)	(2.077e-03)	(2.057e-03)	(2.155e-03)	(2.055e-03)
Num Regulatory Restrictions	-1.363e-06***	-1.279e-06***	-1.334e-06***	-1.467e-06***	-1.356e-06***
	(2.495e-07)	(2.538e-07)	(2.557e-07)	(2.490e-07)	(2.530e-07)
House Price Index	-2.343e-04***	-2.983e-04***	-2.946e-04***	-3.143e-04***	-3.475e-04***
	(4.406e-05)	(4.269e-05)	(4.476e-05)	(4.442e-05)	(4.222e-05)
Nasdaq Avg	6.504e-06***	-1.938e-06	-2.230e-06	-3.178e-06	-6.643e-06***
	(2.282e-06)	(1.787e-06)	(1.913e-06)	(1.970e-06)	(1.894e-06)
Volatility Avg	-1.606e-04	-1.856e-04	-3.662e-05	-3.330e-05	-1.368e-04
	(1.657e-04)	(1.660e-04)	(1.724e-04)	(1.930e-04)	(1.659e-04)
Consumer Confidence Avg	-8.782e-03***	-6.264e-03***	-5.542e-03***	-4.108e-03**	-3.452e-03**
	(1.731e-03)	(1.702e-03)	(1.754e-03)	(1.596e-03)	(1.583e-03)
Inflation Index	-1.621e-04	5.896e-04***	5.567e-04***	5.781e-04**	8.353e-04***
	(2.366e-04)	(2.050e-04)	(2.124e-04)	(2.305e-04)	(2.090e-04)
Unemployment Rate	4.989e-05	-1.615e-03	-1.612e-03	-1.224e-03	-2.453e-03**
	(1.095e-03)	(1.074e-03)	(1.075e-03)	(1.099e-03)	(1.125e-03)
Oil Price	-2.297e-05	-1.989e-04*	-2.221e-04*	-1.419e-04	-1.838e-04*
	(1.091e-04)	(1.093e-04)	(1.145e-04)	(1.081e-04)	(1.096e-04)
is ICT	-1.066e-02***	-1.066e-02***	-1.062e-02***	-1.062e-02***	-1.063e-02***
	(1.858e-03)	(1.858e-03)	(1.858e-03)	(1.858e-03)	(1.858e-03)
is Software	8.419e-03**	8.463e-03**	8.505e-03**	8.577e-03**	8.534e-03**
	(3.727e-03)	(3.727e-03)	(3.728e-03)	(3.728e-03)	(3.728e-03)
N	147553	147553	147553	147553	147553

Table 8: Robustness check using the alternative proximity measure built at the firm-level

Stage 2 regression on 4-digit mean firm US patenting firm level patenting proximity to citations in the first year it patents. Also included in the regressors are 2-digit SIC controls. Standard errors are clustered on 2-digit SIC sectors. The expected number of deals are . All time-varying RHS variables are lagged one year.

5 Concluding remarks

We have seen that firm originality has been decreasing since 2008 and that young firms have become less original with respect to older firms. We have also established that the proximity of a firm to its potential acquirer has a positive effect on its likelihood of buyout and that indeed the expectations of being bought out have a robust negative effect on firm entry originality.

The innovation literature considers new firms to be an important source of radical innovation however this paper shows that due to changes in financing conditions, through buyout expectations, new firms are doing less original innovation. This study has been focused on innovation measures built with technology codes to quantify an innovation position. In chapter **??**, I will do a complimentary study on the effect of initial conditions, namely innovation position, on future firm development.

This study has also shed light on how startup innovation choices are affected by their exit options. I suggested that getting bought out has increasingly become a chief exit option since the financial crisis and this has consequently affected their initial entry innovation strategies. Namely, if a start-up believes that getting bought out is its primary exit option, then it will rationally choose to further increase its likelihood of getting bought out by innovating in closer complementary proximity to their potential acquirer.

The bigger picture is to consider the consequences of more consolidation and less original firms. If the objective of new firms is increasingly to be bought out then there will be less competition in the future, implying a stagnating economy. We did in fact see a prolonged and persistent period of low growth after the financial crisis and this paper suggests that changing firm innovation incentives due to firm interactions may be one mechanism.

In addition to proposing a part of the reason for declining business dynamism, our analysis also has implications for policy makers. The increase in alternative funds for new firms may offset some of the direct effect on entry however it may be skewing the innovation incentives on the new entrants that leave a longer term effect.

This paper also provides a new perspective on the push and pull effects of financing on innovation. Traditionally finance has been considered to have a push effect on innovation however here I suggest it can also have an indirect pull effect. Namely when the medium of financing is equity, there are a mix of motives for the firm and the investors. In the case of buyouts, the source of funds is the acquiring firm and that firm has a demand for certain kinds of technology and innovations. This demand from firms for types of innovation is what influences the initial decisions made by startups.

6 Bibliography

References

Aghion, P., Bergeaud, A., Lequien, M., and Melitz, M. (2019). The heterogeneous impact of market size on innovation: Evidence from french firm level exports. *Working Paper.*

Akcigit, U. and Ates, S. (2019). Ten facts on declining business dynamism and lessons from endogenous growth theory. *American Economic Journal: Macroeconomics*, forthcoming.

Akcigit, U., Hanley, D., and Serrano-Velarde, N. (2013). Back to basics: Basic research spillovers, innovation policy and growth. *NBER Working Paper*, 19473.

Arora, A., Fosfuri, A., and Ronde, T. (2018). Waiting for the payday? the market for startups and the timing of entrepreneurial exit. *Centre for Economy Policy Research, 35 p.*

Atanasova, C. and Chemla, G. (2020). Does familiarity breed alternative investment? evidence from corporate defined benefit pension plans. *Working paper.*

Ayyagari, M., Beck, T., Soledad, M., and Peria, M. (2018). The micro impact of macroprudential policies firm level evidence. *IMF Working Paper*, 18(267).

Bacchetta, P., Benhima, K., and Poilly, C. (2019). Corporate cash and employment. *American Economic Journal: Macroeconomics*, 11(3):30–66.

Bena, J. and Li, K. (2014). Corporate innovations and mergers and acquisitions. *Journal of Finance*, LXIX(5):1923–1960.

Bessen, J. and Hunt, R. M. (2007). An empirical look at software patents. *Journal of Economics and Management Strategy*, 16(1):157–189.

Bloom, N., Jones, C., Van Reenen, J., and Webb, M. (2017). Are ideas getting harder to find? *NBER Working Paper.*

Bloom, N., Schankerman, M., and Reenen, J. V. (2013). Identifying technology spillovers and product market rivalry. *Econometrica*, 81(4):1347–1393.

Bordalo, P., Gennaioli, N., and Shleifer, A. (2018). Diagnostic expectations and credit cycles. *Journal of Finance*, LXXIII(1):199–227.

Cahn, C., Girotti, M., and Landier, A. (2019). Entrepreneurship and information on past failures: A natural experiment. *Working Paper.*

Chemmanur, T. and Tian, X. (2018). Do anti-takeover provisions spur corporate innovation? a regression discontinuity analysis. *Journal of Financial and Quantitative Analysis.*

Choi, J., Goldschlag, N., Haltiwanger, J., and Kim, J. D. (2019). Founding teams and startup performance. *Working paper.*

Coibion, O., Gorodnichenko, Y., and Kumar, S. (2018). How do firms form their expectations? new survey evidence. *American Economic Review*, 108(9):2671–2713.

Davila, J. (2014). The rationality of expectations formation. *CORE Discussion Paper.*

Davis, S. and Haltiwanger, J. (2019). Dynamism diminished: The role of housing markets and credit conditions. *NBER Working Paper*.

Decker, R., Haltiwanger, J., Jarmin, R., and Miranda, J. (2014). The role of entrepreneurship in us job creation and econmic dynamism. *Journal of Economic Perspectives*, 28(3):3–24.

Decker, R. A., Haltiwanger, J., Jarmin, R. S., and Miranda, J. (2016). Declining business dynamism: What we know and the way forward. *American Economic Review*, 106(5):203–207.

Diecidue, E. and de Ven, J. V. (2008). Aspiration level, probability of success and failure and expected utility. *International Economic Review*, 49(2):683–700.

Eaton, J., Kortum, S., Neiman, B., and Romalis, J. (2016). Trade and the global recession. *American Economic Review*, 106(11):3401–3438.

Fudenberg, D. and Levine, D. (2016). Whither game theory? towards a theory of learning in games. *Journal of Economic Perspectives*, 30(4):151–170.

Gilboa, I. (2014). Rationality and the bayesian paradigm.

Gilboa, I., Postlewaite, A., and Schmeidler, D. (2008). Probability and uncertainty in economic modeling. *Journal of Economic Perspectives*, 22(3):173–188.

Greenstone, M., Mas, A., and Nguyen, H.-L. (2020). Do credit market shocks affect the real economy? quasi-experimental evidence from the great recession and normal economic times. *American Economic Journal: Economic Policy*, 12(1):200–225.

Hall, B. and Lerner, J. (2010). The financing of r&d and innovation. *Hall, B.H. and N. Rosenberg (eds.), Handbook of Economics of Innovation*.

Haucap, J., Rasch, A., and Stiebale, J. (2019). How mergers affect innovation: Theory and evidence. *International Journal of Industrial Organization*, 63:283–325.

Heidhues, P., Koszegi, B., and Strack, P. (2018). Unrealistic expectations and misguided learning. *Econometrica*, 86(4):1159–1214.

Horbach, J., Rammer, C., and Rennings, K. (2012). Determinants of eco-innovations by type of environmental impact - the role of regulatory push/pull, technology push and market pull. *Ecological Economics*, 78:112–122.

Hussinger, K. (2010). On the importance of technological relatedness: Smes versus large acquisition targets. *Technovation*, 30:57–64.

Jaffe, A. (1986). Technological opportunity and spillovers of r&d: Evidence from firms' patents, profits, and market value. *The American Economic Review*, 76(5):984–1001.

Jones, C. (2011). Misallocation, economic growth, and input-output economics. *NBER Working Paper*.

Kerr, W. and Nanda, R. (2009). Financing constraints and entrepreneurship. *NBER Working Paper*.

Kerr, W. and Nanda, R. (2015). Financing innovation. *Bank of Finland Research Discussion Papers*.

Kuchler, T. and Zafar, B. (2019). Personal experiences and expectations about aggregate outcomes. *Journal of Finance*, LXXIV(5):2491–2542.

Landier, A., Ma, Y., and Thesmar, D. (2019). Biases in expectations: Experimental evidence. *Working Paper*.

Lemley, M. A. and McCreary, A. (2021). Exit strategy. *Boston University Law Review*.

Nanda, R. and Rhodes-Kropf, M. (2013). Innovation and the financial guillotine. *NBER Working Paper*.

Negro, M. D. and Schorfheide, F. (2004). Priors from general equilibrium models for vars. *International Economic Review*, 45(2):643–673.

Nemet, G. (2009). Demand-pull, technology-push, and government-led incentives for non-incremental technical change. *Research Policy*, 38:700–709.

Ornaghi, C. (2009). Mergers and innovation in big pharma. *International Journal of Industrial Organization*, 27(1):70–79.

Ouimet, P. and Zarutskie, R. (2014). Who works for startups? the relation between firm age, employee age, and growth. *Journal of Financial Economics*, 112(3):386–407.

Prais, S. (1958). The statistical conditions for a change in business concentration. *Review of Economics and Statistics*, 40(3):268–272.

Seru, A. (2014). Firm boundaries matter: Evidence from conglomerates and r&d activity. *Journal of Financial Economics*, 111(2):381–405.

Sevilir, M. and Tian, X. (2015). Acquiring innovation. *Working Paper*.

Tian, X. and Wang, T. Y. (2014). Tolerance for failure and corporate innovation. *Review of Financial Studies*, 27(1):211–255.

Trajtenberg, M., Henderson, R., and Jaffe, A. (1997). University versus corporate patents: A window of the basicness of invention. *Economics of Innovation and New Technology*, 5:19–50.

Woodford, M. (2013). Macroeconomic analysis without the rational expectations hypothesis. *Annuel Review of Econmics*, 5.

Zhao, X. (2009). Technological innovation and acquisitions. *Management Science*, 55(7):1170–1183.

A Name matching process

The names we start off with are applicant names from Patstat that have already been cleaned (the variable psn_name), and target and acquiror names from Thomson SDC Platinum. Names in Patstat are particularly difficult to work with as there is no regularization nor tracking over time or between patent offices, they are not verified with official databases and they are prone to misspellings.

The names matching consists of first dealing the the name misspellings and disambiguation, then running a fuzzy string matching algorithm on the cleanned names, and then filtering out mismatches where names are short or consist of common words.

The name cleaning and disambiguation consists of:

- first converting all letters to uppercase,
- then dealing with symbols. Almost all are removed and replaced with a space. Except, we replace & with "AND" and $ with "S"
- then we group single letter words in the name together. This could be relevant for initials or country codes or incorporation status, etc.
- then we remove a list of words that do not define the company. To be clear, these are: CORPORATION, COMPANY, COMPANIES, COMP, CORP, INCORPORATED, INTERNATIONAL, HOLDING, SYSTEM(S), PRODUCT(S), KABUSHIKI KAISHA, THE, INC, SAS, GMBHDE, GMBH, MBH, LTDA, LTD, SRL, SARL, SA, SPA, SE, ABP, AB, BV, NV, PTE, PTY, LLC, PLC, AG, KG, OY, SL, AS.
- finally we build a dictionary for certain words that appear with different spellings but refer to the same thing. This also includes firms that are often referred to be their abbreviations. For example: IBM -¿ International Business Machines, 3M -¿ Minnesota Mining and Manufacturing, BMW -¿ Bayerische Motoren Werke, as well as Mgmt -¿ Management, Tech -¿ Technology, etc.

We do this initial cleaning with both Patstat applicant PSN names and SDC target and acquiror names.

Then we start the name matching. We assume that company names in SDC are more reliable and therefore base our definition of a *name* match as a threshold percentage of word matches in the SDC name. We ran this twice, once with a requirement that all words match and once with a requirement that 60 percent of the words match. To minimize potential false positives, we show only results from the requirement that all words match.

The algorithm then iterates the list of SDC names and compares with each PSN name. This comparison is done by splitting the name into a list of words then comparing each word for a match.

A *word* match is calculated from the levenshtein distance and the restriction is variable depending on the length of the word. If the word has less than or equal to 5 characters, we require an exact match. If there are between 6 and 9 characters and the levenshtein distance is less than or equal to 1, we consider that a matched Finally, if the word is over 9 characters, we say it is a match if the levenshtein distance is less than or equal to 2.

At this point we have a set of potential matches however there are a few checks to be made. Some company names use generic words and make a minor change (e.g. SOLUTIONS vs. eSOLUTIONS). "eSolutions" will match with any company name that has the word "solutions" in it and since it is a common word, there may be many mismatches. To check for common words, we build a list of common words from Patstat names by simply spliting the Patstat name into words and counting the occurance of each word. We consider common words to be words that are counted at least 100 times and are at least five characters long. We then run through the potential name matches and check if the SDC name contains a common word substring. If so, then we require that the matching name has a word that matches exactly with this.

Another source of error in our potential matches are in the length of names. After our cleaning step, we have a few SDC names that are only one word. As Patstat has a lot of applicants and many with long names, we get many erroneous matches for short SDC names. One check we do is to identify the one word SDC firm names and require that the matching name be at most 2 words with the one that doesn't match being a maximum of 3 letters. Another check we do is on both short names and common words. If the SDC firm name is less than or equal to 3 words and at least one is a common word, then we require an exact match on the uncommon word.

This is the extent of our name matching right now. To complete the merge between the two databases, we use additional data on zip code, state code, and country code to supplement the matching when available.

Patstat has address data on applicants however SDC only goes down to the granularity of zip code and the zip code field is poorly populated in Patstat. Sometimes it appears in the address field and needs to be parsed. To do so, I use the usaddress python package to extract the zip codes and state codes when available. The zip code and state code data in both Patstat and SDC are still relatively poorly populated but for the firms that we do have information for, we use as a filter to check for an address match.

B Additional Results

	Firm Entry Originality				
E[Num Deals$_{t+1}$]	-2.943e-05***				
	(5.533e-06)				
E[Num Deals$_{t+2}$]		-9.770e-06***			
		(1.954e-06)			
E[Num Deals$_{t+3}$]			-1.313e-05***		
			(2.439e-06)		
E[Num Deals$_{t+4}$]				-1.634e-05***	
				(3.400e-06)	
E[Num Deals$_{t+5}$]					-4.830e-06***
					(1.521e-06)
Fed Funds Rate	-2.771e-03***	-2.706e-03***	-2.973e-03***	-3.225e-03***	-2.986e-03***
	(3.963e-04)	(3.978e-04)	(3.964e-04)	(4.113e-04)	(4.010e-04)
Treasury 10 year rate	2.391e-03**	1.021e-03	1.857e-03*	3.126e-03***	1.684e-03*
	(1.022e-03)	(1.030e-03)	(1.015e-03)	(1.072e-03)	(1.015e-03)
Regulatory Restrictions	-1.768e-06***	-1.838e-06***	-1.865e-06***	-1.724e-06***	-1.670e-06***
	(1.184e-07)	(1.206e-07)	(1.211e-07)	(1.180e-07)	(1.190e-07)
House Price Index	3.175e-05	4.764e-05**	2.893e-05	3.349e-05	5.907e-05***
	(2.061e-05)	(1.986e-05)	(2.060e-05)	(2.087e-05)	(1.978e-05)
Nasdaq Avg	-4.978e-06***	-2.476e-06***	-3.273e-06***	-3.436e-06***	-2.291e-06**
	(1.107e-06)	(8.612e-07)	(9.047e-07)	(9.394e-07)	(8.952e-07)
Volatility Avg	-2.413e-04***	-2.226e-04***	-3.589e-04***	-4.709e-04***	-2.362e-04***
	(7.851e-05)	(7.871e-05)	(8.114e-05)	(9.336e-05)	(7.838e-05)
Consumer Confidence Avg	-1.749e-04	-5.332e-04	-1.620e-04	-1.345e-03*	-1.747e-03**
	(8.148e-04)	(8.036e-04)	(8.176e-04)	(7.473e-04)	(7.394e-04)
Inflation Index	1.038e-03***	7.913e-04***	8.756e-04***	9.687e-04***	7.939e-04***
	(1.161e-04)	(1.001e-04)	(1.028e-04)	(1.133e-04)	(1.029e-04)
Unemployment Rate	-3.592e-03***	-2.967e-03***	-2.914e-03***	-3.570e-03***	-3.518e-03***
	(4.859e-04)	(4.753e-04)	(4.758e-04)	(4.927e-04)	(5.062e-04)
Oil Price	-3.234e-04***	-2.464e-04***	-1.943e-04***	-2.724e-04***	-3.043e-04***
	(4.850e-05)	(4.880e-05)	(5.083e-05)	(4.795e-05)	(4.873e-05)
is ICT	3.177e-02***	3.178e-02***	3.177e-02***	3.178e-02***	3.175e-02***
	(7.285e-04)	(7.283e-04)	(7.283e-04)	(7.283e-04)	(7.282e-04)
is Software	-3.325e-03**	-3.308e-03**	-3.294e-03**	-3.369e-03**	-3.404e-03**
	(1.534e-03)	(1.534e-03)	(1.534e-03)	(1.534e-03)	(1.534e-03)
N	154955	154955	154955	154955	154955

Table 9: Robustness check using the N-digit originality measure

Stage 2 regression on n-digit firm originality in the first year it patents. Also included in the regressors are 2-digit SIC controls. Standard errors are clustered on 2-digit SIC sectors. Expected number of buyouts are measured in the first stage as described above. All time-varying RHS variables are lagged one year.

	Firm Entry Max Originality				
E[Num Deals$_{t+1}$]	-2.368e-05**				
	(1.066e-05)				
E[Num Deals$_{t+2}$]		-9.719e-06**			
		(3.992e-06)			
E[Num Deals$_{t+3}$]			-9.715e-06*		
			(4.984e-06)		
E[Num Deals$_{t+4}$]				-8.700e-06	
				(6.678e-06)	
E[Num Deals$_{t+5}$]					2.471e-06
					(3.151e-06)
Fed Funds Rate	-2.490e-03***	-2.414e-03***	-2.643e-03***	-2.748e-03***	-2.457e-03***
	(8.030e-04)	(8.051e-04)	(8.043e-04)	(8.247e-04)	(8.102e-04)
Treasury 10 year rate	-1.508e-03	-2.733e-03	-1.950e-03	-1.314e-03	-2.101e-03
	(1.978e-03)	(1.986e-03)	(1.964e-03)	(2.060e-03)	(1.964e-03)
Regulatory Restrictions	-1.137e-06***	-1.214e-06***	-1.206e-06***	-1.103e-06***	-1.135e-06***
	(2.258e-07)	(2.301e-07)	(2.315e-07)	(2.255e-07)	(2.281e-07)
House Price Index	2.926e-04***	3.017e-04***	2.928e-04***	3.032e-04***	3.254e-04***
	(4.031e-05)	(3.930e-05)	(4.098e-05)	(4.094e-05)	(3.907e-05)
Nasdaq Avg	-3.345e-06	-1.597e-06	-1.831e-06	-1.463e-06	4.111e-07
	(2.143e-06)	(1.705e-06)	(1.802e-06)	(1.856e-06)	(1.788e-06)
Volatility Avg	2.854e-04*	3.045e-04**	1.983e-04	1.626e-04	2.798e-04*
	(1.550e-04)	(1.553e-04)	(1.610e-04)	(1.821e-04)	(1.550e-04)
Consumer Confidence Avg	5.685e-04	5.481e-04	4.629e-04	-5.365e-04	-9.566e-04
	(1.619e-03)	(1.595e-03)	(1.628e-03)	(1.495e-03)	(1.481e-03)
Inflation Index	4.222e-04*	2.363e-04	2.816e-04	2.998e-04	1.340e-04
	(2.227e-04)	(1.951e-04)	(2.010e-04)	(2.178e-04)	(1.995e-04)
Unemployment Rate	-1.106e-03	-5.923e-04	-5.669e-04	-9.390e-04	-3.942e-04
	(9.843e-04)	(9.671e-04)	(9.679e-04)	(9.935e-04)	(1.023e-03)
Oil Price	-1.021e-04	-3.362e-05	-3.828e-06	-6.331e-05	-5.562e-05
	(9.903e-05)	(9.928e-05)	(1.036e-04)	(9.809e-05)	(9.937e-05)
is ICT	6.936e-02***	6.937e-02***	6.935e-02***	6.935e-02***	6.935e-02***
	(1.653e-03)	(1.653e-03)	(1.653e-03)	(1.653e-03)	(1.653e-03)
is Software	-3.189e-02***	-3.186e-02***	-3.187e-02***	-3.193e-02***	-3.192e-02***
	(3.402e-03)	(3.402e-03)	(3.402e-03)	(3.402e-03)	(3.403e-03)
N	154955	154955	154955	154955	154955

Table 10: Robustness check using the maximum 4-digit IPC code firm originality

Stage 2 regression on 4-digit max firm originality in the first year it patents. Also included in the regressors are 2-digit SIC controls. Standard errors are clustered on 2-digit SIC sectors. Expected number of buyouts are measured in the first stage as described above. All time-varying RHS variables are lagged one year.

	Firm Entry Max Originality				
E[Num Deals$_{t+1}$]	-2.771e-05***				
	(5.421e-06)				
E[Num Deals$_{t+2}$]		-8.364e-06***			
		(1.908e-06)			
E[Num Deals$_{t+3}$]			-1.094e-05***		
			(2.377e-06)		
E[Num Deals$_{t+4}$]				-1.468e-05***	
				(3.330e-06)	
E[Num Deals$_{t+5}$]					-4.601e-06***
					(1.485e-06)
Fed Funds Rate	-2.525e-03***	-2.474e-03***	-2.699e-03***	-2.935e-03***	-2.729e-03***
	(3.849e-04)	(3.863e-04)	(3.850e-04)	(4.004e-04)	(3.899e-04)
Treasury 10 year rate	1.917e-03*	6.832e-04	1.394e-03	2.547e-03**	1.253e-03
	(9.969e-04)	(1.003e-03)	(9.890e-04)	(1.046e-03)	(9.899e-04)
Regulatory Restrictions	-1.561e-06***	-1.618e-06***	-1.638e-06***	-1.521e-06***	-1.468e-06***
	(1.125e-07)	(1.146e-07)	(1.151e-07)	(1.121e-07)	(1.131e-07)
House Price Index	4.432e-05**	6.096e-05***	4.583e-05**	4.742e-05**	6.995e-05***
	(2.006e-05)	(1.934e-05)	(2.006e-05)	(2.035e-05)	(1.928e-05)
Nasdaq Avg	-4.284e-06***	-1.809e-06**	-2.441e-06***	-2.731e-06***	-1.768e-06**
	(1.083e-06)	(8.400e-07)	(8.822e-07)	(9.179e-07)	(8.741e-07)
Volatility Avg	-1.663e-04**	-1.505e-04*	-2.645e-04***	-3.726e-04***	-1.615e-04**
	(7.674e-05)	(7.694e-05)	(7.928e-05)	(9.138e-05)	(7.659e-05)
Consumer Confidence Avg	2.494e-04	-2.086e-04	6.765e-05	-8.784e-04	-1.228e-03*
	(7.961e-04)	(7.838e-04)	(7.971e-04)	(7.294e-04)	(7.213e-04)
Inflation Index	9.123e-04***	6.742e-04***	7.429e-04***	8.363e-04***	6.832e-04***
	(1.138e-04)	(9.802e-05)	(1.006e-04)	(1.110e-04)	(1.008e-04)
Unemployment Rate	-3.185e-03***	-2.601e-03***	-2.558e-03***	-3.140e-03***	-3.120e-03***
	(4.711e-04)	(4.604e-04)	(4.609e-04)	(4.779e-04)	(4.914e-04)
Oil Price	-2.876e-04***	-2.180e-04***	-1.754e-04***	-2.399e-04***	-2.698e-04***
	(4.679e-05)	(4.704e-05)	(4.900e-05)	(4.623e-05)	(4.699e-05)
is ICT	3.103e-02***	3.104e-02***	3.103e-02***	3.103e-02***	3.101e-02***
	(6.888e-04)	(6.887e-04)	(6.887e-04)	(6.887e-04)	(6.886e-04)
is Software	-4.291e-03***	-4.280e-03***	-4.270e-03***	-4.332e-03***	-4.365e-03***
	(1.445e-03)	(1.445e-03)	(1.446e-03)	(1.446e-03)	(1.446e-03)
N	154955	154955	154955	154955	154955

Table 11: Robustness check using the maximum n-digit IPC code firm originality.

Stage 2 regression on n-digit max firm originality in the first year it patents. Also included in the regressors are 2-digit SIC controls. Standard errors are clustered on 2-digit SIC sectors. Expected number of buyouts are measured in the first stage as described above. All time-varying RHS variables are lagged one year.