

The Technological Core of Apple: Using Artificial Neural Networks and Econometrics to Value Apple’s Patents

Daniel K.N. Johnson¹ and Matthew Whitehead²

¹Department of Economics and Business, Colorado College, 14 East Cache la Poudre Street, Colorado Springs, CO 80903, United States, djohnson@ColoradoCollege.edu

²Department of Mathematics and Computer Science, Colorado College, 14 East Cache la Poudre Street, Colorado Springs, CO 80903, United States, matthew.whitehead@ColoradoCollege.edu

This paper evaluates the importance of patent-related events to Apple’s stock price, testing three nuances that economics has omitted. First, patents vary widely in quality so we use artificial neural networks to generate word embeddings to estimate patent creativity and breadth as complements to traditional measures of patent generality, originality and significance. Second, we test the possibility that information about patent applications leaks during the long and uncertain time until patent grant, so evaluate the stock price effect at four different dates in the life of each patent. Third, we test whether the signaling value of Steve Jobs’s name on a patent document has an effect on Apple’s financial valuation.

1. Introduction

Firms know that patents represent valuable assets, as evidenced by the expenses to develop and subsequently protect innovations. However, valuations of intellectual property are performed on an ad hoc basis at times of merger, acquisition or other pivotal decision moments; this is not only because of the complex and dynamic technological relationships between firms, but because economic analysis has failed to provide a method that a priori can quantify the value of a patent or IP portfolio. We propose that this is due to two fundamental reasons: a) patents vary widely in value, and b) have a long and uncertain timeline to market (including multiple dates at which information might leak out to the market, and complicated final documents that take time for the market to interpret. This analysis follows close on the footsteps of Johnson and Scowcroft (2016), investigating whether Apple’s patent events are correlated with stock price fluctuations once we account for patent quality, the timing of news distribution, and potential star power signaling effects exerted by the name “Steve Jobs” on a patent document.

This paper aims to improve our modeling of the stock market valuation of patents, using a nuanced version of event analysis. We model the timing of four distinct stages in the disclosure of information about the potential value of a patent, two of them private and two of them public, two of them preliminary and two of them legally binding. We test this insight using patents granted to Apple, Inc. over a 13-year period in order to determine whether Apple can potentially predict, and even manipulate, their own stock prices by the strategic release of patent news. Our reasoning is that if Apple, a company closely watched by technology specialists and based almost entirely on a reputation for (mostly patented) technological innovation, cannot affect its market price via patent events, then either the events are too small to matter or our measurement tools are too crude to discern the impact. A secondary motivation ties into a larger literature on executive compensation (e.g. Adams et al. (2005)), but we aim here to quantify Steve Jobs’s impact as a co-inventor translated into a market signal of innovation quality.

2. Literature

There is a long literature on event analysis in the finance discipline, and we emulate that established methodology in the amended style of Johnson and Scowcroft (2016). Fama and French (1996) outline the four-factor model which sets the expected stock value in the absence of market-relevant news, with abnormal returns calculated as deviations from this

Paper submitted to:

R&D Management Conference 2016 “*From Science to Society: Innovation and Value Creation*” 3-6 July 2016, Cambridge, UK
expected value. Other studies (e.g., those reviewed in Hall, 2000; Hall et al., 2007) have found that patents do not explain as much market value as anticipated, faring worse than research expenditures in that regard but adding explanatory power beyond those expenditures alone.

The literature on patent valuation spans industrial and commercial analysis but also has a vibrant home in economics. Johnson and Popp (2003) document the empirical correlation evidence of many factors with patent value, building on previous work by Pakes and Simpson (1989), Trajtenberg (1990), Lanjouw et al. (1996), and Lanjouw and Schankerman (1999). All of these are predicated on characteristics of the patent document itself that conveys or implies value, whether subsequent citations (a measure of scholarly or scientific value), patent families (a measure of how broadly international applicants protect their rights, a measure of potential market value), or renewal rates (a measure of how often the applicant pays to retain patent rights, a measure of potential duration of market value). Direct measures of market value are rare, usually relying on auction prices (e.g., Sneed and Johnson, 2009) but point to other measures of patent value such as originality and generality which we will adopt and augment here.

Most importantly, Johnson and Scowcroft (2016) just recently published work on Apple stock price valuation, work which this current paper extends. We draw heavily on their data and methodology, with our primary additional contribution coming in the form of alternative valuation techniques from artificial neural networks: we use machine learning to create word embeddings from on the underlying patent documents, generate potential measures of creativity and patent breadth, and then re-implement many of the same tests that Johnson and Scowcroft (2016) used.

3. Data and Methodology

We consider 3,440 market days between 1998 and 2011 as potential event dates, tracking 395 separate patent-related events for 182 distinct patent documents. Those patent events occurred on 212 distinct dates, leaving 3,228 dates without a patent-related event concerning Apple (days which we use as a control group). For events which occurred on days when stock markets were closed, or for which the event window would have occurred on closure days, we use the market closing price from the previous day (if the closure occurred before the patent event) or subsequent day (if the closure occurred after the patent event).

3.1. Dependent variable

Relying on the work of Johnson and Scowcroft (2016), we used their (somewhat longer) time series of data as a baseline for defining average and abnormal returns. That paper used the corporate history of the firm and the personal history of Steve Jobs (Time, 2013; Telegraph, 2013; Apple Museum, 2013), identifying 64 events of potential significance which were classified as personal fame (e.g. Pixar winning an Academy Award, Steve Jobs being named “Entrepreneur of the Decade”), health (e.g. Steve Jobs retires citing health reasons), product launches (e.g. the iPhone is announced), collaborations (e.g. Disney buys Pixar), and cuts (e.g. personnel cutbacks). They included indicator variables for each type of news event within the relevant event window, and we emulate that approach to capture the flavor of the effect that each type of news event might have.

Using the Johnson and Scowcroft (2016) control group of observations, we regress each event window’s observed rate of return on French’s (2013) data for the Carhart (1997) four-factor model as extended from Fama and French (1996) and augmented with indicator variables for news events within the window:

$$R_t = \alpha_0 + \alpha_T T_t + \alpha_P P_t + \alpha_S S_t + \alpha_B B_t + \alpha_M M_t + \sum_{i=1}^5 \alpha_i N_i + e_t \quad (1)$$

where R_t is the rate of return on Apple stock;
 T_t is the implicit rate of return on one-month Treasury bonds;
 P_t is the difference between the rate of return on the market portfolio and T_t ;
 S_t is the difference between the rate of return on a portfolio of small-cap stocks and the rate of return on a portfolio of large-cap stocks;
 B_t is the difference between the rate of return on a portfolio of high book-to-market value stocks and a portfolio of low book-to-market value stocks;
 M_t is the momentum factor;
 N_1 through N_5 are indicator variables for the five news categories: fame, health, launches, collaborations and cuts;
 e_t is the residual.

The primary analysis of this control group uses a two-day rate of return, to compare with pre-event-to-post-event closing prices in our experimental group, but we also perform sensitivity analysis on week-long and two-week-long rates of return. The estimated coefficients of this model, White-corrected to address heteroskedasticity, are presented in Table 1 for the four time event windows we consider.

Table 1: Estimated coefficients for Four-Factor Model on Apple Stock Returns

Coeff.	Short -- Quick (1 day before until 1 day after)		Short -- Slow (1 day before until 7 days after)		Long -- Quick (7 days before until 1 day after)		Long -- Slow (7 days before until 7 days after)	
α_0	0.347	(3.87)***	1.104	(7.19)***	1.128	(7.13)***	1.870	(9.14)***
α_T	-3.069	(1.88)*	15.843	(3.87)***	-3.850	(4.04)***	-14.731	(5.42)***
α_P	13.829	(3.54)***	8.614	(2.37)**	-2.987	(1.05)	2.361	(0.85)
α_S	1.845	(0.28)	-27.603	(4.14)***	-12.672	(2.72)***	-20.610	(3.94)***
α_B	-17.245	(3.08)***	-10.425	(1.52)	-14.852	(3.50)***	-14.488	(2.85)***
α_M	-2.596	(0.53)	10.945	(2.72)***	-9.701	(2.84)***	-4.340	(1.29)
α_{Fame}	1.112	(1.16)	-0.844	(0.97)	2.521	(2.42)**	3.065	(3.07)***
α_{Health}	-0.421	(0.49)	-0.007	(0.01)	0.405	(1.32)	-0.682	(0.74)
$\alpha_{Launches}$	-0.749	(1.49)	0.064	(0.13)	-0.991	(1.69)*	-1.031	(1.80)*
$\alpha_{Collabs}$	0.856	(0.72)	-1.638	(1.46)	2.634	(3.46)***	1.691	(1.36)
α_{Cuts}	-0.694	(0.69)	-2.953	(3.40)***	5.036	(4.47)***	1.991	(1.23)
F-stat.	3.62***		6.90***		8.05***		6.85***	

Source: Johnson and Scowcroft (2016). Significance is indicated by *** for 99%, ** for 95%, * for 90%, and t-statistics are included in parentheses.

Notice that the simple four-factor model is satisfactory according to standard diagnostics (t-statistics and F-statistic), although it explains only a very small fraction of the variation in returns, a result which was largely anticipated as we model a single firm over time. Sensitivity tests which exclude all news categories show virtually no change in any coefficient values or significance levels among the traditional four-factor variables.

The constructed variables for categories of news events are (unsurprisingly) largely insignificant, with one notable exception. For the long-quick window aperture, four of the five news event indicators showed up with statistically significant coefficients. In that one case, news about fame-related events, collaborations and personnel cuts were all associated with stock price increases, while product and service launches were (inexplicably) associated with decreases. In order to give our model the best possible chance of success, we select this long-quick (7 days before until one day after the event) window for our primary results below but have verified that the alternatives show similar (although less statistically significant) results.

Based on these results, we calculate the expected rate of return \hat{R}_t for every date in our sample. Abnormal returns are therefore calculated as the standardized difference between actual and expected returns, $AR_t = [R_t - \hat{R}_t]/s$ where s is the standard deviation of the expected rates of return in the control group. The Dickey-Fuller test rejects the null hypothesis of a unit root in this series well above the 99% level for each event window.

3.2. Timing

We consider four events in the life of a patent document--- application, publication, status acceptance and grant--- and only consider successful patents (i.e., those which we know to have been granted by the end of 2012). While this definition serves to eliminate some heterogeneity from our sample, by eliminating unsuccessful and abandoned applications from consideration, it opens our analysis to criticisms of survival bias. Further, investors do not know in advance which applications will be successful, so may react to applications not represented in our sample. However, details of unsuccessful patents were not public information for more than half of our sample period, as explained below, so we choose to omit them entirely from consideration.

When an inventor (or their employer) applies for a patent, their potential protection period begins at the date of application. Prior to June 9 of 1995, when the United States began a transition to conform to World Trade Organization standards for patent law, the information in a patent application would be kept secret until the date of grant: if a patent was never granted, the information was never public. Since that date patent applications have been published 18 months after their application date, regardless of when and even whether they are granted. This date of public revelation we call the publication date, representing the first moment at which technical details about the patentable information are known to be public knowledge.

After consideration by patent examiners, and an average of three to five years of communication to clarify and justify the legal claims in the application, a status acceptance letter is sent privately to the applicant, indicating that the legal claim to a patent will be granted. While the public announcement usually follows within two to three weeks, this stage of the process is still private, and may have little to no stock market impact.

Finally, in a very predictable manner after the close of business each Monday, the United States Patent and Trademark Office publishes the details of newly granted patents, presenting the first public announcement of the legally enforceable intellectual asset. Patent defendants win roughly ninety percent of infringement cases brought to court, so this granting date represents a serious financial claim of ownership or at least of the legal right to exclude others from production (Kortum and Lerner, 1998).

When granted, each patent includes a set of required information, which we have found useful to analyze here. Every document includes the names of all applicants (inventors, not firms) so we easily identify patents which show Steve Jobs as a contributor. We compile not only the four dates of interest, two private (application and status acceptance) and two public (publication and grant) but also characteristics of each patent document which permit us to assess its potential value.

3.3. Valuation

Since patents vary widely in value, we search for methods of quantifying those variations. For example, each patent lists the technological claims made to ownership over intellectual space, in general listing more claims on a more broad or wide-reaching patent document. To justify those claims, patents must include a bibliography, other documents and patents of which the inventors were aware and which are clearly outside the scope of the current claims. We not only identify all citations listed in each Apple patent, but identify all subsequent patent documents which cite them in return. Using the technology type (technically, the U.S. Patent Class assigned by the patent examiner) of each forward and backward citation, we reproduce and extend the work of Hall et al. (2001), calculating originality and generality measures for each Apple patent. Originality of each patent is calculated as

$$originality = 1 - \sum_{i=1}^n s_i^2 \quad (2)$$

or a measure of how broadly the references span across technologies, where s_i is the share of all references within a given patent that hail from a particular technology class. Notice that the summation is a version of the Herfindahl Index of concentration, exponentially emphasizing classes with high shares of the total. Generality is measured symmetrically as

$$generality = 1 - \sum_{i=1}^n s_i^2 \quad (3)$$

where s_i here is the share of all subsequent citations within each technology class, or a measure of how broadly the given patent is cited, across technologies.

But potentially more important than individual indicators coded by the patent office is the entirety of the text within each patent document; presumably, that text reveals great insight into the nature of the technology and its possible value. Unfortunately, each document is far too long and technical to be assessed manually, so we rely on machine learning to perform that analysis instead. We propose the use of word embeddings, from the computational science field of natural language processing, to calculate two other metrics for patent evaluation: creativity and breadth.

Word embeddings are vector representations of words that can capture meanings in numeric ways. These representations can then be used in a wide variety of ways, including performing direct comparisons of the semantic relationships between individual words, sentences, and full documents. They are obtained from large corpora of plain-text data by finding patterns in existing written language and then extracting those patterns in a compressed way as numeric semantic features. There are many common ways to perform this type of feature extraction including latent semantic analysis (Deerwester, 1990), neural-probabilistic language models (Bengio, 2003), deep learning models (Collobert, 2008), and the skip-gram algorithm with negative sampling algorithm (Mikolov, 2013). The skip-gram model has been shown to have state-of-the-art performance on a number of prominent natural language processing tasks, so we use it for this work. The model uses artificial neural networks to try to predict individual word context, which consists of the words surrounding each single word in a document. As the model sees more training text, it gradually adjusts its parameters so that the error is minimized when performing the context prediction. Once training is complete, the network is presented with a single word and the resulting internal activations of the network are used as that word's semantic embedding vector. Figure 1 shows a skip-gram network diagram.

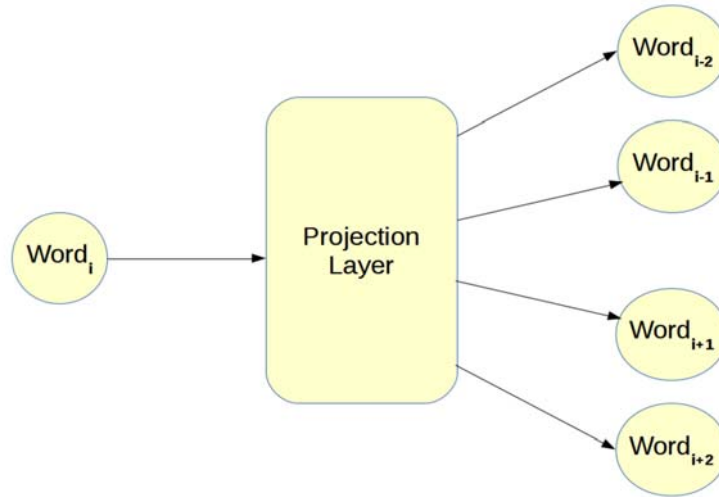


Figure 1: Skip-Gram Model: For a given word at position i in a document, the skip-gram algorithm's artificial neural network predicts the words that form the surrounding context ($i-2$, $i-1$, $i+1$, $i+2$) using a projection layer. The projection forms the word embedding.

For the purposes of this work, we use word embeddings directly with the text of patent documents to generate full document embeddings that lie in the same semantic space as the words themselves. These document embeddings can be formed in various simple ways including averaging or summing the word embedding vectors of the words contained within the document. Once document embeddings are obtained, then we can reason about the relationships between documents in computational ways.

Given the semantic document vectors for the patents under consideration, we can then calculate new measures of patent value. Our first proposed metric is patent creativity and it is calculated by measuring the distance between a patent's document vector, v , and the centroid of all the other patent document vectors in v 's technology class category. We reason that the greater the distance measured, the more original and creative that v is because the semantics of its textual information are separate from other patents in its category. We use the standard Euclidean measure to calculate distances between data points in the high-dimensional space. Figure 2 shows an example plot of document vectors in a single technology class in 3-dimensional space for visualization purposes, though the actual distances are calculated in 300-dimension space. Documents vectors that are far from the center of mass of the dataset (such as the red triangle point) have higher creativity scores than those nearer the center.

In addition to a patent's creativity measure, we would also like to determine a patent's scope of semantic coverage, so we propose the use of a patent breadth metric. The intuition for this metric is that more valuable patents with larger intellectual property coverage will also span larger parts of the semantic vector space of the word embeddings. We can then calculate the volume of the semantic space spanned by a document's word embeddings as an estimation of a patent's breadth. The volume spanned by our documents in 300-dimensional space could be interpreted in various ways but the most straightforward way is the use of a convex hull surrounding the word vector data points. Because of the computational difficulty of finding convex hulls and their volumes efficiently in 300 dimensions, we use an approximation volume using a simpler high-dimensional bounding-box method. This value is calculated by taking each of the word embedding vectors for a patent document and first finding the extrema along each axis. The document's approximate hypervolume and breadth is then given simply by the product of each of the i axis ranges: $\prod(\max_i - \min_i)$.

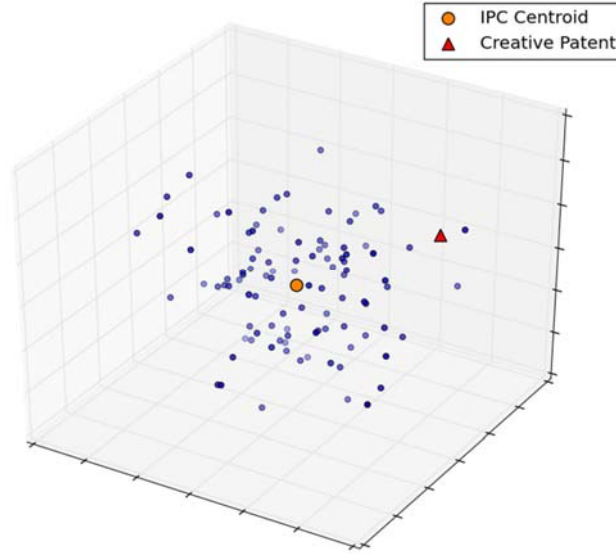


Figure 2: Patent Creativity: All the plotted points are patent document embeddings for a single technology class. The red triangle represents a patent document with a high level of creativity since it is far from its technology class centroid, which is depicted by the orange sphere.

Figure 3 shows two example bounding boxes of different patent word clusters in three dimensions (again for the purpose of visualization only). The left document (blue) contains words with more extreme values along its axes resulting in a larger bounding box and corresponding volume. The right document (red) shows a much tighter packing of word vectors which leads to a smaller bounding box and volume.

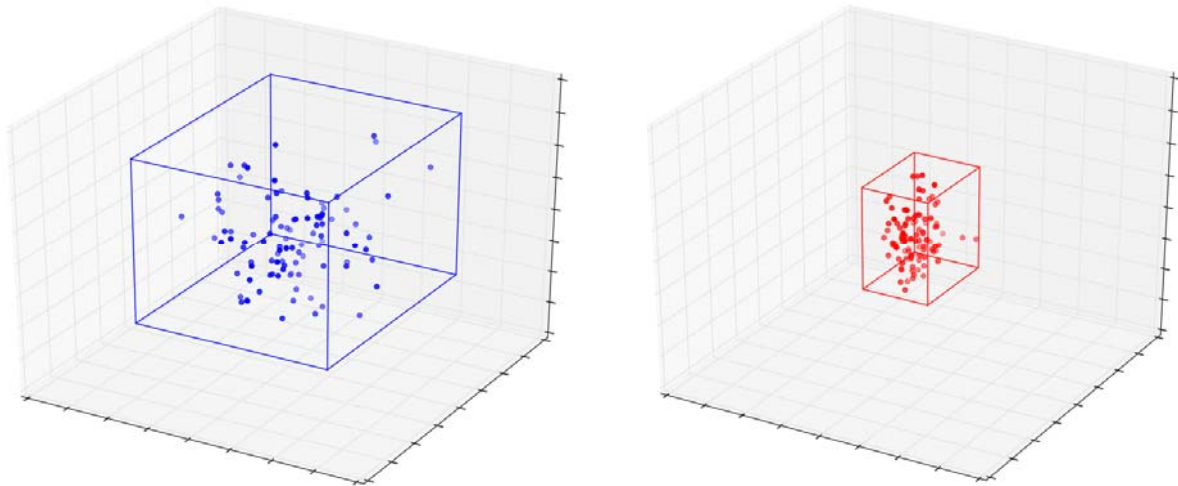


Figure 3: Patent Breadth: Both images show scatterplots of patent word embedding vectors for a single document each. The volumes of the surrounding boxes indicate the measures of patent breadth (high for the patent represented by the blue plot on the left and low for the patent represented by the red plot on the right).

Although machine learning has been used to evaluate patent documents before (e.g. latent semantic analysis of content to find adjacencies, as in Magerman et al., 2011), to our knowledge this is the first work to use these two new proposed metrics to score patents. We then test them against other more traditional indicators of patent value. Even before a formal multivariate regression test, there is interesting evidence that these proposed indicators of patent value are not measuring the same thing. Table 2 shows the correlation between various proposed indicators; notice that the

correlation coefficients are low and frequently even negative. While Johnson and Scowcroft (2016) found that many of these indicators took on higher values for event-days in the top quartile of the sample, they also found a shocking diversity among the alternative proxies for patent value.

Table 2: Correlations between alternative measures of patent value

Variable	Inventors	Citations	Claims	Creativity	Breadth	Generality	Originality	Steve Jobs
Inventors	1							
Citations	0.08	1						
Claims	0.03	-0.04	1					
Creativity	-0.06	-0.10	-0.18	1				
Breadth	-0.07	-0.10	-0.06	0.40	1			
Generality	0.06	-0.11	0.05	-0.10	-0.19	1		
Originality	-0.01	0.32	-0.02	-0.20	-0.09	-0.07	1	
Steve Jobs	0.03	0.13	0.03	-0.07	-0.02	-0.01	-0.03	1

As we proceed with the model, we must aggregate measures to represent daily events rather than patent document events. Since multiple patents may be granted on the same date, we sum their characteristics on any given day in order to determine their aggregate effects (if any) on Apple stock prices. In many cases, different types of events occur on a given day, with some patent applications being filed as others are granted, so we summarize the characteristics of patent documents in each event-group (application, publication, status, issue) separately.

3.4. Model

Now we test the ability of patent-related events to explain AR_t in the standard fashion, permitting each estimated coefficient β to vary by event type (application, publication, status or issue), and so grouping all patents which share a similar event on a given date:

$$AR_t = \beta_0 + \beta_1 Jobs_t + \beta_2 patentvalue_t + \delta_1 NASDAQ_t + u \quad (4)$$

where $Jobs$ is an indicator of how many patents on that date show Steve Jobs name;
 $patentvalue$ is a measure of the patent events on that date, either a count of patents, or the total number of subsequent citations, or the total number of claims, or the total originality or generality or creativity or breadth score for patents, or the total number of inventors listed on patents;
 $NASDAQ$ is the rate of return of the NASDAQ composite index over the same period;
and u is a residual or error term.

We elect not to include time-indicating variables, as the dependent variable has already been de-trended using the first step of the model.

Clearly, if patent events signal an asset's acquisition, then the estimated coefficients of patents should be positive (at least for public event types). We include the NASDAQ composite index to capture other market events which might confound our identified events but which have not already been netted out by the Fama-French method.

4. Results

Table 3 presents our primary results using the “long-quick” event window (from the preceding week until the following day), with alternative measures of patent value used in each column: patent counts in the first column, creativity as constructed using word embeddings in the second column, breadth as constructed using word embeddings in the third column. Other indicators of value show similar results, so are not reported here.

Each regression was performed with time control variables (e.g. day-of-the-week variables, NASDAQ variable, Steve Jobs effect, constant) and each type of event (application, publication, status and issue) was separated out as a separate effect. All results are presented here corrected for both heteroskedasticity.

First, notice that on average, patent-related events do not correlate strongly with abnormal returns. Instead, it is simply industry-wide returns that (inversely) correlate with Apple price movements. This result accords with the previous

Table 3: Regression results of “long-quick” AR_t on patent events

Variable	Patent counts		Creativity		Breadth	
	Coefficient	t-statistic	Coefficient	t-statistic	Coefficient	t-statistic
Application						
Jobs	6.873	(9.21)***	6.917	(12.33)***	6.611	(10.13)***
Patent	-0.063	(0.09)	-0.470	(0.24)	0.700	(0.35)
Publication						
Jobs	0.056	(0.03)	0.374	(0.19)	0.079	(0.04)
Patent	-0.183	(0.36)	-1.131	(0.48)	-0.710	(0.32)
Status						
Jobs	-0.435	(0.16)	-0.538	(0.19)	-0.366	(0.13)
Patent	-0.061	(0.39)	-0.103	(0.23)	-0.379	(0.77)
Issue						
Jobs	0.798	(0.87)	0.690	(0.80)	0.907	(1.00)
Patent	-0.083	(0.62)	-0.259	(0.63)	-0.443	(0.97)
NASDAQ	0.939	(23.04)***	0.939	(23.03)***	0.938	(23.03)***
Constant	-0.484	(0.47)	-0.498	(0.48)	-0.501	(0.49)
Obs		3440		3440		3440
F-statistic						
R ²		0.22		0.22		0.22

Significance is indicated in the table as *** for 99%, ** for 95%, * for 90%. Patent events are represented by the value measure identified in the column heading (counts, creativity or breadth).

literature which shows limited explanatory power from patents (Hall, 2000; Hall et al., 2007). This is also true across quintiles of the data, results available from the authors upon request.

Second, note that the only consistently significant patent-related event appears to be the presence of Steve Jobs’s name on a patent document, and that is only true at the initial date of application. Post-test investigation revealed that to be entirely due to a single patent application; removal of that observation removed all impact of patent-related events on stock market returns in our sample.

Furthermore, while it is comforting that the patent-quality variables produced via word embeddings result in virtually the same result as using simple counts of patent documents, this result does not inspire us to recommend the use of complex mathematical and computational algorithms in future analyses of this type. For full disclosure, none of the other alternative indicators of patent quality (number of claims, number of citations, number of inventors, generality or originality) displayed notable differences from a simple document count either.

Robustness tests performed using alternative event windows (which might represent a more immediate response to news or an overall longer time for reflection) displayed no significant differences. In the interest of space we elect not to present them here, but results are available from the authors upon request.

Lest the reader suspect that the insignificance was caused by a swamping of all other variables by the powerful explanatory effect of the NASDAQ or other news events, consider this: news events have already been netted out in the preceding stage, this stage of results changes trivially with the omission of the NASDAQ variable, and even including the NASDAQ variable leaves us with roughly 3/4 of the variation in abnormal returns unexplained ($R^2 < 0.25$ in most robustness tests).

5. Conclusion

The goal of this paper was to examine the role that patent-related events have on share prices, given the historical difficulties in identifying their effects. We proposed that two factors account for that quantitative difficulty: the variability in patent value, and the long and uncertain timeframe during which information leaks to the market. This paper has modeled both of those factors for Apple Inc., while incorporating a quantified personality effect and novel potential measures of patent value constructed using artificial neural networks to generate word embeddings.

Our findings are depressingly limited in usefulness. We confirm that (for the most part), patent events are not discernibly important to Apple’s stock price, presumably because stock prices do not respond to technical information within a week, or because they respond so gradually as to be indistinguishable from other noisy variations. That result

Paper submitted to:

R&D Management Conference 2016 “From Science to Society: Innovation and Value Creation” 3-6 July 2016, Cambridge, UK holds for all manner of events from application through patent grant, regardless of the “size” of the intellectual property asset under consideration. The result also holds across a wide variety of uncorrelated measures of patent value.

The advice to firms, and specifically to innovation managers, can be taken from a famous quote alternatively attributed to John Wooden and C.S. Lewis: “The true test of a man’s character/integrity is what you do when no one is watching” (Goodreads.com, 2016). Innovators and managers of creative teams must act as though no investor notices what they are doing, because apparently in the short-term, no one does. This poses some thorny asymmetric information problems, as presumably innovators and innovation managers have better information about the potential value of their intellectual property than the market can discern.

Alternatively, and perhaps more optimistically, the value of the intellectual property may lie entirely in the practice, in the application, in the market creation that occurs after the legal-technological event. It is unquestionable that the stock market values innovative technological firms, so our conclusion can be an inspiration for small (and large) firms everywhere: the market isn’t valuing what you own when you first acquire it, but is waiting to see what you do with it, whether you wield your tools to advantage, how you proceed with the legal rights given to you by the patent system. For that valuation exercise, there is no *a priori* calculation that will predict success or failure, just good business practices to follow.

References

- Adams, R., H. Almeida, and D. Ferreira. 2005. “Powerful CEOs and their impact on corporate performance.” *Review of Financial Studies* 18, 1403-1432.
- Apple Museum. 2013. <http://applemuseum.bott.org/sections/history.html> accessed in June of 2013.
- Bengio, Y., R. Ducharme, R. Vincent, and C. Jauvin. 2003. “A neural probabilistic language model.” *Journal of Machine Learning Research*, 3:1137–1155, 2003.
- Carhart, M.M. 1997. “On Persistence in Mutual Fund Performance”. *Journal of Finance* 52 (1), 57–82.
- Collobert, R. and J. Weston. 2008. “A unified architecture for natural language processing: Deep neural networks with multitask learning.” In ICML.
- Deerwester, S., S. Dumais, G. Furnas, T. Landauer, and R. Harshman. 1990. “Indexing by Latent Semantic Analysis.” *Journal of the American Society for Information Science* 41 (6): 391–407.
- Fama, E., and K. French. 1996. “Multifactor explanations of asset pricing anomalies”. *Journal of Finance* 51, 55-84.
- French, K. 2013. Data retrieved on February 14, 2013 from http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html
- [Google Patents. 2016. https://patents.google.com/ Retrieved on April 29, 2016.](https://patents.google.com/)
- Hall, B. H., A. B. Jaffe, and M. Trajtenberg. 2001. “The NBER Patent Citation Data File: Lessons, Insights and Methodological Tools.” *NBER Working Paper* 8498.
- Hall, B. H. 2000. “Innovation and Market Value,” In R. Barrell, G. Mason, and M. O'Mahoney, eds., *Productivity, Innovation and Economic Performance*, Cambridge, England: Cambridge University Press, 2000.
- Hall, B. H., G. Thoma, and S. Torrisi. 2007. “The market value of patents and R&D: Evidence from European firms,” *NBER Working Paper* 13426.
- Johnson, D.K.N. and D. Popp. 2003. “Forced Out of the Closet: The Impact of the American Inventors Protection Act on the Timing of Patent Disclosure”. *Rand Journal of Economics* 34(1), Spring 2003, 96-112.
- Johnson, D.K.N. and S. Scowcroft. 2016. “The Importance of Being Steve: An Econometric Analysis of the Contribution of Steve Jobs’s Patents to Apple’s Market Valuation”. *International Journal of Financial Research*, 7(2), April 2016, 30-38.
- Kortum, S. and J. Lerner. 1998. “Stronger protection or technological revolution: what is behind the recent surge in patenting?” *Carnegie-Mellon Conference Series on Public Policy* 48: 247-304.
- Lanjouw, J.O., A. Pakes and J. Putnam. 1996. “How to Count Patents and Value Intellectual Property: Uses of Patent Renewal and Application Data,” *NBER Working Paper* #5741.
- Lanjouw, J.O. and M. Schankerman. 1999. “The Quality of Ideas: Measuring Innovation with Multiple Indicators,” *NBER Working Paper* #7345.
- Magerman, T., B. Van Looy, B. Baesens, and K Debackere. 2011. “Assessment of Latent Semantic Analysis (LSA) text mining algorithms for large-scale mapping of patent and scientific publication documents,” Katholieke Universiteit Leuven Department of Managerial Economics Strategy and Innovation Working Paper 1114.
- Mikolov, T., I. Sutskever, K. Chen, G. Corrado, and J. Dean, J. 2013. “Distributed representations of words and phrases and their compositionality.” *Neural Information Processing Systems Conference Proceedings*, Paper 5021.
- Pakes, A. and M. Simpson. 1989. “Patent Renewal Data,” *Brookings Papers:Microeconomics*, 331-410.

Paper submitted to:

R&D Management Conference 2016 “*From Science to Society: Innovation and Value Creation*” 3-6 July 2016, Cambridge, UK

Sneed, K.A. and D.K.N. Johnson. 2009. “Selling Ideas: The Determinants of Patent Value in an Auction Environment,” *R&D Management*, 39(1), 87-94.

Time. 2013. History of Apple.

http://www.time.com/time/specials/packages/article/0,28804,1873486_1873491_1873530,00.html accessed June of 2013.

Telegraph. 2013. Steve Jobs Timeline. <http://www.telegraph.co.uk/technology/steve-jobs/8810045/Steve-Jobs-timeline.html> accessed June of 2013.

Trajtenberg, M. 1990. “A Penny for Your Quotes: Patent Citations and the Value of Innovations,” *Rand Journal of Economics*, 21, 172-187.