Investigating Correlation Between Intellectual Property (IP) Filings and Financial Statement Data

Ram Soundararajan
IP Authority, LLC

ABSTRACT

Correlation between financial data extracted from publicly available financial statements (such as EBITDA, revenue, gross profit, net income, and R&D) associated with various companies and patent data also related with the same companies were investigated. It is shown that in 71.43% of the companies investigated, there appears to be a positive correlation between the R&D data obtained from such financial statements and that of the patent data extracted via Thomson's Innovation database.

Keywords: Intellectual Property, Patents, Financial Statements

1. INTRODUCTION

Attempts have been made in the past (see Van Triest, S. & Vis, W. (2007)) to link financial data with the valuation process associated with patents. Particularly, they investigate cash flow due to competitive advantage (i.e., competitive advantage from practicing the technology at a lower cost than their competitors who would have to license such technology), licensing income, and costs to maintain existing, enforceable patents (commonly referred to as maintenance fees).

Prior attempts generally looked at financial data and financial ratios, and linking such data/ratios to stock returns, without any analysis of associated patents or licensing revenue associated with a given company (Bukit, I. N. H. & Anggono, I. A. H. (2013)). Other such (minimal) attempts have been made from a financial perspective mainly with regards to a company's licensing fees (both fees that it pays as a licensing agreement and fees it collects as part of licensing agreements with others) (see, for example, Arora, A. (1997), Eswaran, M. (1994), Gallini, N. T. (2002), and Goh, P. C. & Lim, K. P. (2004)) an interdisciplinary analysis of IP filings as related to financial statements has not yet been done in the prior art.

Such an investigation could potentially uncover correlation and trend data between data points in such financial statements and a company's IP filings. This paper addresses such an investigation.

2. METHODOLOGY

To limit the current investigation to identify any existing correlation between IP filings and financial statements, only the top 10 patent filers were considered at the beginning of the investigation. Accordingly, the following top 10 patent filers in the Unites States in the last full calendar year, i.e., 2015, were identified (Source: http://www.ificlaims.com/index.php?page=misc_top_50_2015):

Rank	Company	No. of Patents Granted
1	International Business Machine Corp (IBM)	7,355
2	Samsung Electronics Co. Ltd.	5,072
3	Canon KK	4,134
4	Qualcomm	2,900
5	Google Inc.	2,835
6	Toshiba Corp.[1]	2,627
7	Sony Corp.	2,455
8	LG Electronics Inc.[1]	2,242
9	Intel Corp.	2,048
10	Microsoft Technology Licensing LLC	1,956

TABLE 1: Top 10 patent filers in the US for 2015

Next, data was collected in a two-prong fashion. First, Thomson's Innovation Database (http://info.thomsoninnovation.com/) was used to collect patent filing data for each of the companies identified in the listing above for years 2006 through 2013. Particularly, data was collected with regards to how many U.S. patents and pre-grant publications were filed for each of the entities in the above-listing for each of the years from 2006 through 2013.

It should be noted that while data was collected for years 2014 through 2016, such data was found to be unreliable and incomplete (as it takes up to eighteen (18) months for a patent pre-grant publication to publish, a patent application filed in late 2015 would not have published to be properly counted in the data set). Therefore, these years were not included in the data set. It should also be noted that corporate tree data was used within Thomson's Innovation database to include subsidiaries/acquisitions of each company to get a more detailed listing of IP assets.

Second, financial data from Morningstar® (http://beta.morningstar.com/) was collected for each of the companies listed above for the same years the above-noted patent data was collected, i.e., years 2006 through 2013. Particularly, the following information was collected with regards to each of the entities listed above for each of the above-noted years: EBITDA, revenue, gross profit, net income, and research & development.

---

[1] Analysis with regards to these companies was not conducted due to a lack of publicly available R&D data

It should be noted that Morningstar® did not provide a breakdown for research & development data for Toshiba, Sony and LG. Therefore, in an effort to provide even data analysis, Toshiba, Sony and LG were removed from the list of companies to be investigated.

A sample of both the financial data and patent filing data collected for Google is provided below (numbers in millions, except for patent data, which is the raw number):

Fiscal Year	EBITDA	Revenue	Gross profit	Net income	R&D	Patent/Patent Applications Filed
2006	4583	10605	6380	3077	1229	234
2007	6643	16594	9945	4204	2120	406
2008	8132	21796	13174	4227	2793	468
2009	9836	23651	14806	6520	2843	511
2010	11777	29321	18904	8505	3762	750
2011	14235	37905	24717	9737	5162	1,882
2012	16432	50175	29541	10737	6793	2,970
2013	18518	59825	33967	12920	7952	2,402

TABLE 2: Sample financial and patent filing data

For ease of data analysis, logarithmic values of the financial data collected were used. A sample of such logarithmic data generated for the data collected for Google is provided below:

Fiscal Year	Log (EBITDA)	Log (Revenue)	Log (Gross Profit)	Log (Net Income)	Log (R&D)	Log (Patents/Patent Applications Filed)
2006	3.6611	4.0255	3.8048	3.4881	3.0896	2.3692
2007	3.8224	4.2200	3.9976	3.6237	3.3263	2.6085
2008	3.9102	4.3384	4.1197	3.6260	3.4461	2.6702
2009	3.9928	4.3738	4.1704	3.8142	3.4538	2.7084
2010	4.0710	4.4672	4.2766	3.9297	3.5754	2.8751
2011	4.1534	4.5787	4.3930	3.9884	3.7128	3.2746
2012	4.2157	4.7005	4.4704	4.0309	3.8321	3.4728
2013	4.2676	4.7769	4.5311	4.1113	3.9005	3.3806

TABLE 3: Log of sample financial and patent filing data

Next, the collected data was normalized by offsetting the first data point in each column to 0, and offsetting other data points by the same value. A sample of such normalized data generated for the data collected for Google is provided below:

Fiscal Year	Norm Log (EBITDA)	Norm Log (Revenue)	Norm Log (Gross Profit)	Norm Log (Net Income)	Norm Log (R&D)	Norm Log (Patents/Patent Applications Filed)
2006	0	0	0	0	0	0
2007	0.1612	0.1944	0.1928	0.1355	0.2368	0.2393
2008	0.2490	0.3129	0.3149	0.1379	0.3565	0.3010
2009	0.3317	0.3483	0.3656	0.3261	0.3642	0.3392
2010	0.4099	0.4417	0.4717	0.4415	0.4859	0.5058
2011	0.4922	0.5532	0.5882	0.5003	0.6233	0.9054
2012	0.5545	0.6750	0.6656	0.5428	0.7425	1.1035
2013	0.6064	0.7514	0.7262	0.6231	0.8109	1.0114

TABLE 4: Normalized sample financial and patent filing data

Next, graphs were generated for the financial data sets for comparison against the patent data for each entity.

Lastly, correlation was determined between each normalized financial data array, *F*, (e.g., a first financial data array formed by the column for Norm Log(EBITDA), a second financial data array formed by the column for Norm Log(Revenue), etc.) and the normalized patent data array, *P*, (i.e., the last array labeled Norm Log(Patents/Patent Applications Filed), as follows:

$$Correlation(F,P) = \frac{\sum (f - \bar{f})(p - \bar{p})}{\sqrt{\sum (f - \bar{f})^2 \sum (p - \bar{p})^2}},$$

where $\bar{f}$ and $\bar{p}$ are sample means.

This correlation for each financial data array associated with EBITDA, Revenue, Gross Profit, Net Income, and R&D when compared to the patent data array should give an indication of any correlation between the financial and patent data for each entity.

3. RESULTS AND ANALYSIS

The computed correlation of normalized data generated for the data collected for IBM, Samsung, Canon, Qualcomm, Google, Intel and Microsoft is provided below:

Company	Correlation of Norm Log (EBITDA) with Patent Data	Correlation of Norm Log (Revenue) with Patent Data	Correlation of Norm Log (Gross Profit) with Patent Data	Correlation of Norm Log (Net Income) with Patent Data	Correlation of Norm Log (R&D) with Patent Data
IBM	-0.0192	0.3693	0.1252	-0.1006	0.6415
Samsung	0.4883	0.3400	0.4010	0.5135	0.4824
Canon	-0.7896	-0.6851	-0.6179	-0.6644	-0.4930
Qualcomm	0.4671	0.7426	0.7483	0.3103	0.8297
Google	0.9513	0.9637	0.9606	0.9355	0.9685
Intel	0.1965	0.3191	0.1714	0.0984	0.5772
Microsoft	-0.4374	-0.6799	-0.5956	-0.2648	-0.8291

TABLE 5: Correlation data of financial and patent filing data for top patent filers

From this correlation table, it was noted that except for Canon and Microsoft, there seems to be a positive correlation between the normalized log of research and development data with that of the patent data. In other words, 5 out of the 7 companies, or 71.43% of the companies investigated, showed a positive correlation between the research and development data and the patent data. There does not appear to be a common trend like this in the other financial array data, when looked at in a consolidated manner.

To further visualize such correlation and specifically analyze the correlation between R&D data and patent data, graphical results for each of the seven entities noted in the correlation table above are summarized by way of two graphs – a first consolidated correlation graph and a second, more focused, correlation graph.

3.1 CORRELATION ANALYSIS FOR IBM

FIG. 1 depicts IBM's consolidated graph showing trends in each of the following log-normalized data for IBM: EBITDA, revenue, gross profit, net income, and R&D, as compared with IBM's log-normalized patent filing data. FIG. 2 depicts a more focused graph showing the trend line for IBM's log-normalized R&D data as compared to IBM's log-normalized patent filing data, along with a numerical correlation value (i.e., *correlation=0.642*) computed between the same data sets.
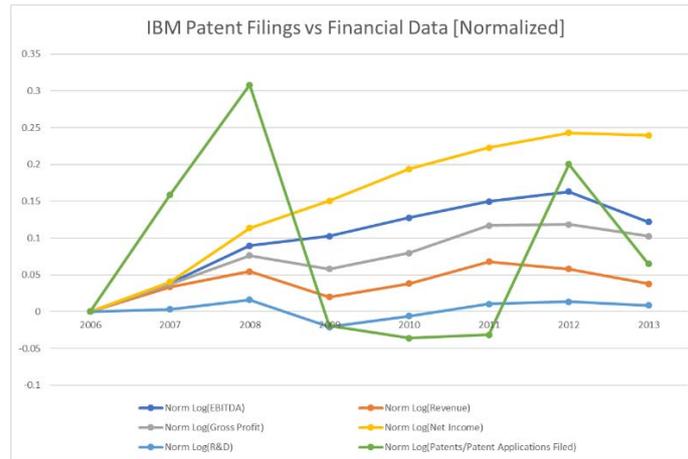
FIG. 1: IBM's consolidated correlation trend
graph
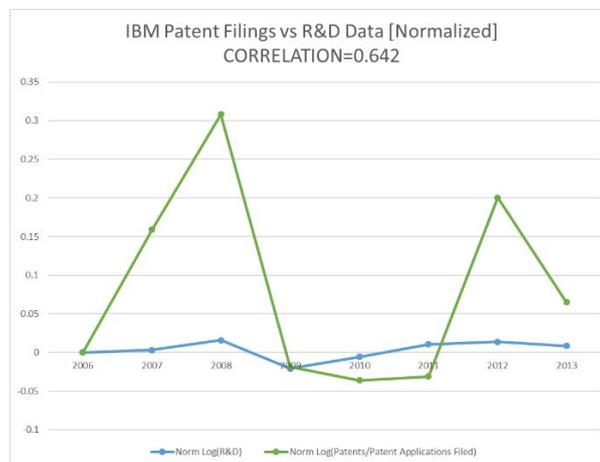
FIG. 2: IBM's correlation between R&D and
patent filing

3.2 CORRELATION ANALYSIS FOR SAMSUNG

FIG. 1 depicts Samsung's consolidated graph showing trends in each of the following log-normalized data for Samsung: EBITDA, revenue, gross profit, net income, and R&D, as compared with Samsung's log-normalized patent filing data. FIG. 2 depicts a more focused graph showing the trend line for Samsung's log-normalized R&D data as compared to Samsung's log-normalized patent filing data, along with a numerical correlation value (i.e., *correlation=0.482*) computed between the same data sets.
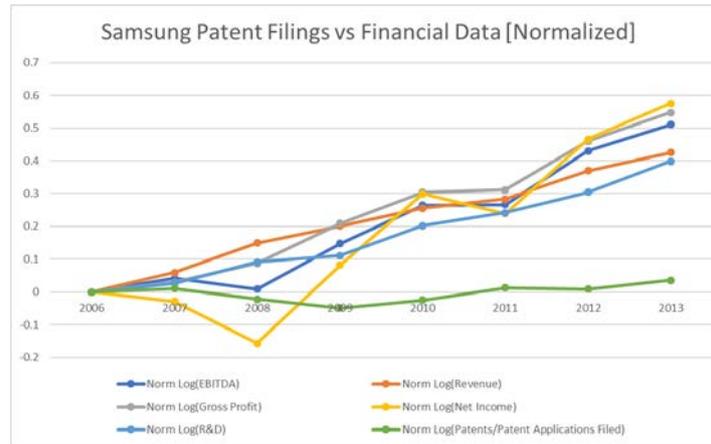
FIG. 3: Samsung's consolidated correlation
trend graph
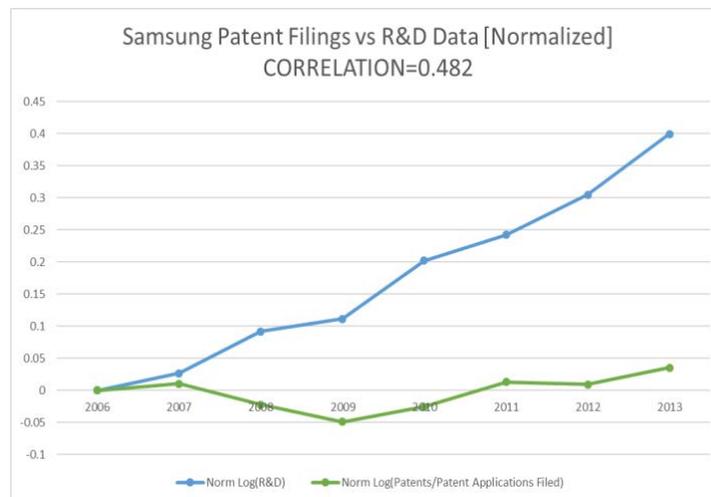
FIG. 4: Samsung's correlation between R&D
and patent filing

3.3 CORRELATION ANALYSIS FOR CANON

FIG. 1 depicts Canon's consolidated graph showing trends in each of the following log-normalized data for Canon: EBITDA, revenue, gross profit, net income, and R&D, as compared with Canon's log-normalized patent filing data. FIG. 2 depicts a more focused graph showing the trend line for Canon's log-normalized R&D data as compared to Canon's log-normalized patent filing data, along with a numerical correlation value (i.e., *correlation=0.493*) computed between the same data sets.
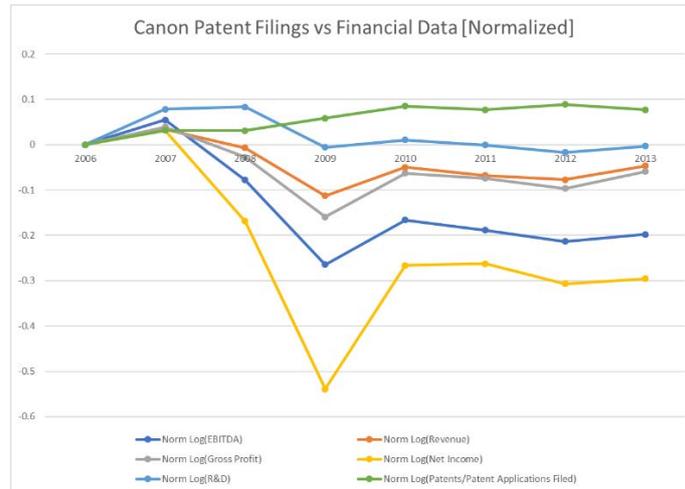
FIG. 5: Canon's consolidated correlation
trend graph
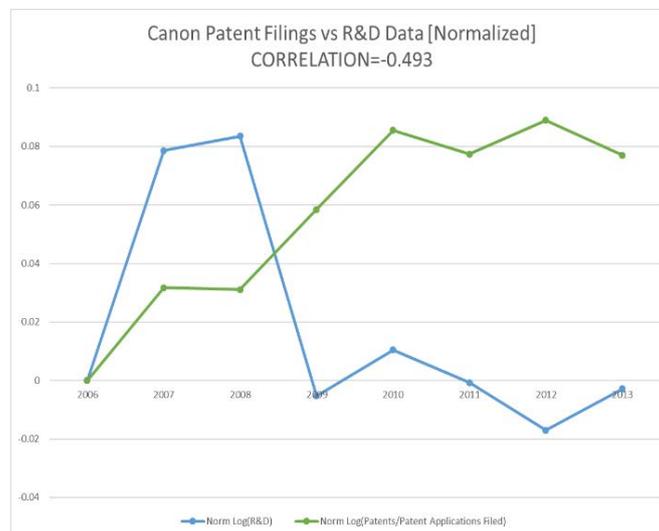
FIG. 6: Canon's correlation between R&D
and patent filing

3.4 CORRELATION ANALYSIS FOR QUALCOMM

FIG. 1 depicts Qualcomm's consolidated graph showing trends in each of the following log-normalized data for Qualcomm: EBITDA, revenue, gross profit, net income, and R&D, as compared with Qualcomm's log-normalized patent filing data. FIG. 2 depicts a more focused graph showing the trend line for Qualcomm's log-normalized R&D data as compared to Qualcomm's log-normalized patent filing data, along with a numerical correlation value (i.e., *correlation=0.830*) computed between the same data sets.
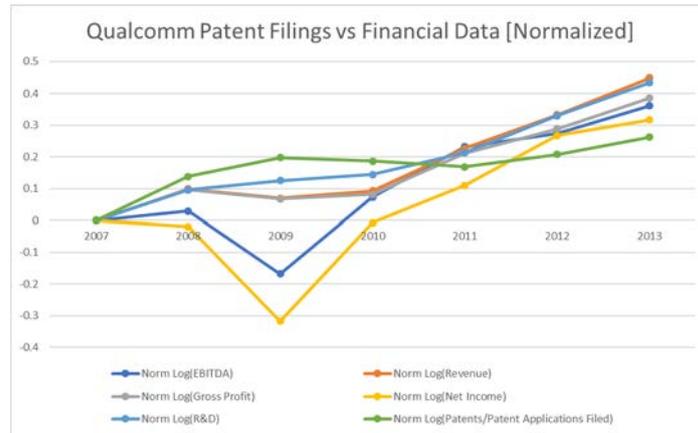
FIG. 7: Qualcomm's consolidated correlation
trend graph
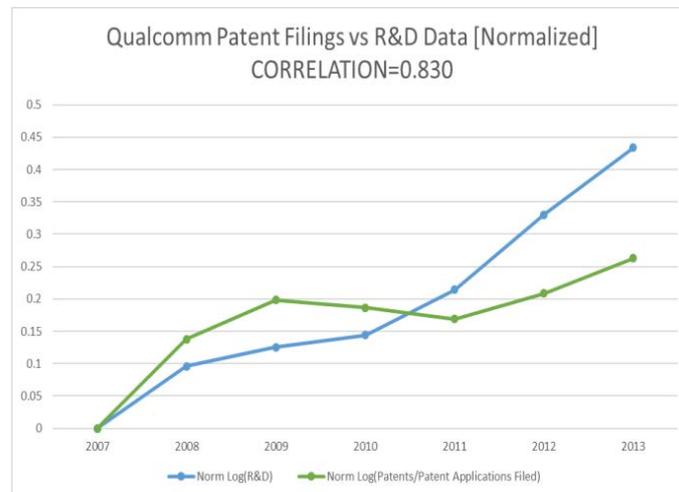
FIG. 8: Qualcomm's correlation between
R&D and patent filing

3.5 CORRELATION ANALYSIS FOR GOOGLE

FIG. 1 depicts Google's consolidated graph showing trends in each of the following log-normalized data for Google: EBITDA, revenue, gross profit, net income, and R&D, as compared with Google's log-normalized patent filing data. FIG. 2 depicts a more focused graph showing the trend line for Google's log-normalized R&D data as compared to Google's log-normalized patent filing data, along with a numerical correlation value (i.e., *correlation=0.969*) computed between the same data sets.
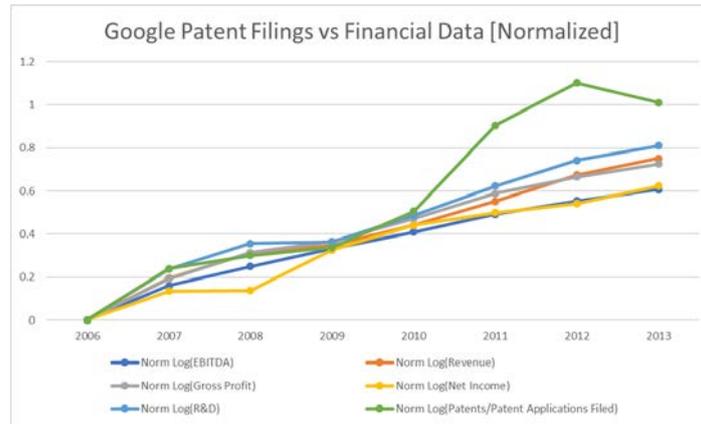
FIG. 9: Google's consolidated correlation trend graph
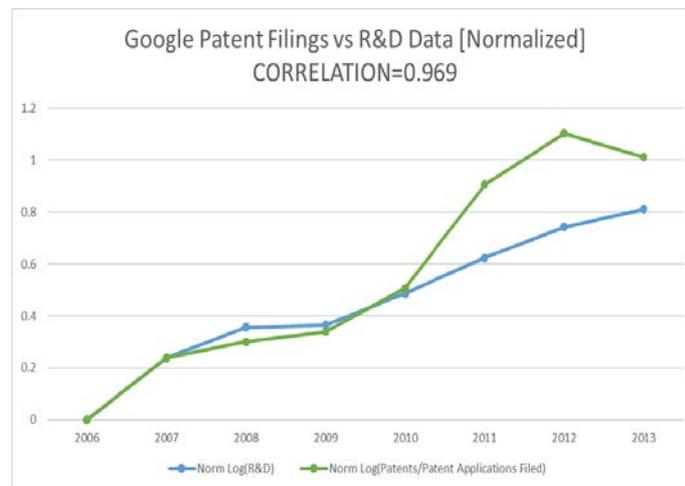
FIG. 10: Google's correlation between R&D and patent filing

3.6 CORRELATION ANALYSIS FOR INTEL

FIG. 1 depicts Intel's consolidated graph showing trends in each of the following log-normalized data for Intel: EBITDA, revenue, gross profit, net income, and R&D, as compared with Intel's log-normalized patent filing data. FIG. 2 depicts a more focused graph showing the trend line for Intel's log-normalized R&D data as compared to Intel's log-normalized patent filing data, along with a numerical correlation value (i.e., *correlation=0.577*) computed between the same data sets.
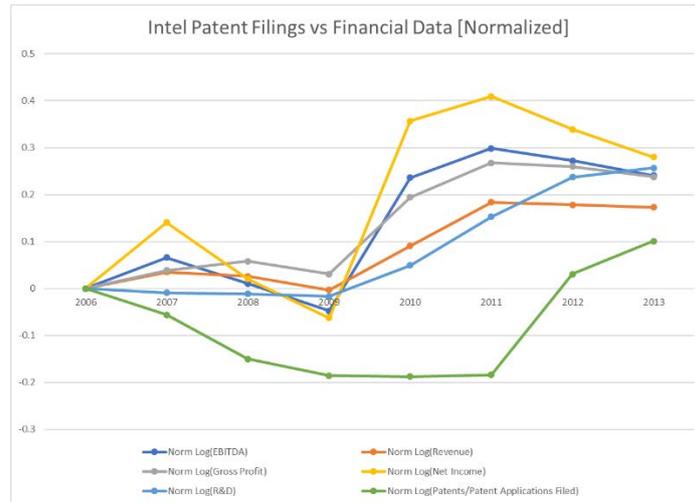
FIG. 11: Intel's consolidated correlation trend
graph
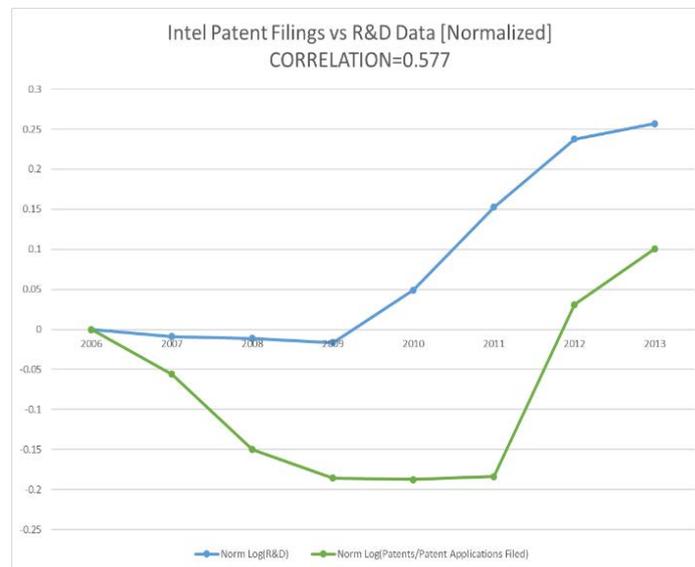
FIG. 12: Intel's correlation between R&D and
patent filing

3.7 CORRELATION ANALYSIS FOR MICROSOFT

FIG. 1 depicts Microsoft's consolidated graph showing trends in each of the following log-normalized data for Microsoft: EBITDA, revenue, gross profit, net income, and R&D, as compared with Microsoft's log-normalized patent filing data.   FIG. 2 depicts a more focused graph showing the trend line for Microsoft's log-normalized R&D data as compared to Microsoft's log-normalized patent filing data, along with a numerical correlation value (i.e., *correlation=-0.829*) computed between the same data sets.
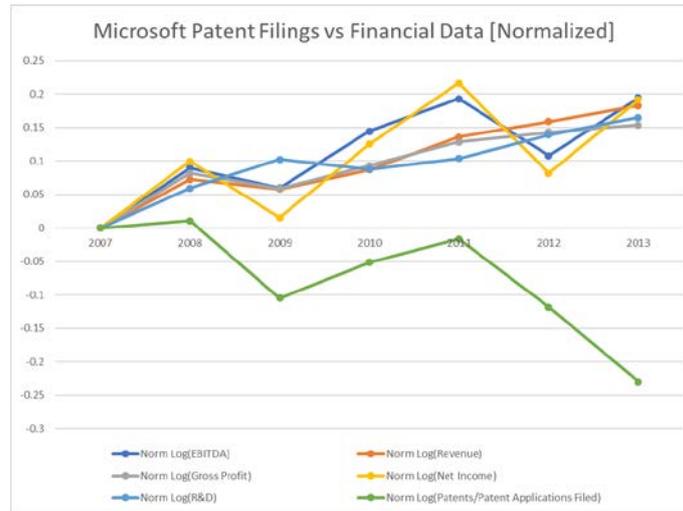
FIG. 13: Microsoft's consolidated correlation trend graph
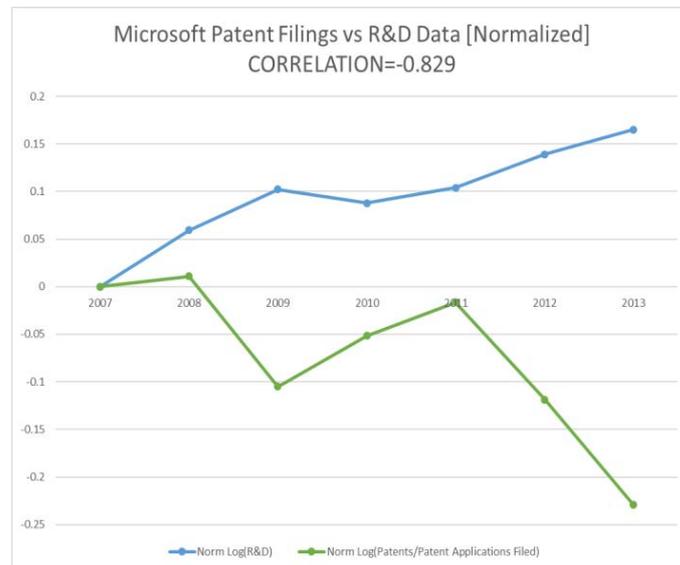
FIG. 14: Microsoft's correlation between R&D and patent filing

Between the various financial data investigated (EBITDA, Revenue, Gross Profit, Net Income, R&D), the one most promising in terms of a trend as compared to patent data appears to be research and development data, which makes sense since R&D is generally tied to IP patent portfolios associated with a company. As noted above, 71.43% of the companies investigated showed a positive correlation amongst the R&D data obtained from financial statements and that of their respective patent data extracted via Thomson's Innovation database. Interesting patterns also emerge when some of these companies are looked at individually. For example, Google shows a very strong across-the-board positive correlation (with numbers greater than or equal to 0.9355) between all examined financial data (i.e., EBITDA, revenue, gross profit, net income) and examined patent filing data (see table below with bolded data items). As another example, in addition to its R&D data, Qualcomm shows

strong correlation between its revenue and gross profit data when compared to its patent filing data (see table below with bolded data items).

Correlation Data for Qualcomm & Google	Qualcomm	Google
Correlation of Norm Log (EBITDA) with Patent Data	0.4671	**0.9513**
Correlation of Norm Log (Revenue) with Patent Data	**0.7426**	**0.9637**
Correlation of Norm Log (Gross Profit) with Patent Data	**0.7483**	**0.9606**
Correlation of Norm Log (Net Income) with Patent Data	0.3103	**0.9355**
Correlation of Norm Log (R&D) with Patent Data	**0.8297**	**0.9685**

TABLE 6: Correlation data for Qualcomm's and Google's

Identifying such positive correlation between aspects of financial data and a company's patent filings could yield valuable insight into the company's intellectual property filings. For example, knowing that IBM's correlation of R&D data with its patent filing data stands at 0.642 and knowing that Qualcomm's correlation of R&D data with its patent filing data stands at 0.830, one can look at future 10-k filings of IBM or Qualcomm, examine their R&D data from such filings and predict, in a reverse fashion, with a reasonable certainty what their patent filing data could be for that year (even though such data might not be available in the public domain yet, as it takes up to 18 months for any patent filings to publish in the public domain). Similarly, knowing that Google's correlation of R&D data with its patent filings stands at a very strong 0.969, one can examine Google's future 10-k filings for a given year and predict, in a reverse fashion, with very good accuracy, what its patent filings could be for that year. Competitors could benefit immensely from knowledge of such predicted patent filing data.

4. CONCLUSION AND FUTURE WORK

The above results of the preliminary investigation appear to show promising correlation between at least one aspect of the financial data obtained from financial statements (i.e., the R&D data) and that of the patent filing data, other data points derived from such financial statement data could be of interest as well. For example, it may be worthwhile to investigate key financial ratios for these companies and see if there is any positive correlation between such key ratios and the patent filing data.

Another interesting possibility involves further analyzing financial data associated with Qualcomm, Google and Intel, since they exhibit across-the-board positive correlation with all sectors of financial data investigated (see table in section 3), with Google particularly showing a very strong correlation (i.e., all correlations greater than or equal to 0.9355). Further investigations are needed to identify the reasons behind such strong across-the-board correlation

numbers for companies like Google. Particularly, further research is needed to see why such strong correlation exists between data associated with Google's EBITDA, revenue, gross profit, and net income and that of Google's patent filings. In addition, such an investigation should also address what such across-the-board correlation means with regards to the financial strength of a company (i.e., does such an across-the-board correlation mean the company is financially stronger when compared to others where such across-the-board correlation doesn't exist).

REFERENCES

[1] Arora, A. (1997). Patents, Licensing, and Market Structure in the Chemical Industry. *Research Policy*, 26(4-5), 391-403.

[2] Bukit, I. N. H. & Anggonno, I. A. H. (2013). The Effect of Price to Book Value (PBV), Dividend Payout Ratio (DPR), Return on Equity (ROE), Return on Asset (ROA), and Earning per Share (EPS) toward Stock Return of LQ 45 Index. *Review of Integrative Business and Economics Research*, 2(2), 22-43.

[3] Eswaran, M. (1994). Cross-Licensing of Competing Patents as a Facilitating Device. *The Canadian Journal of Economics*, 27(3), 689-708.

[4] Gallini, N. T. (2002). The Economics of Patents: Lessons from Recent U.S. Patent Reform. *Journal of Economic Perspectives*, 16(2), 131-154.

[5] Goh, P. C. & Lim, K. P. (2004). Disclosing Intellectual Capital in Company Annual Reports: Evidence from Malaysia. *Journal of Intellectual Capital*, 5(3), 500-510.

[6] Van Triest, S. & Vis, W. (2007). Valuing Patents on Cost-Reducing Technology: A Case Study. *International Journal of Production Economics*, 105(1), 282-292.