

# **Patent citation analysis for company valuation in the field of electric vehicles**

**Yi Qian**

## **Abstract**

The electric vehicle industry is one of the hottest technology sections where patents are essential for companies to survive and compete. The goal of my work is to present EV startups with higher market value locate at more central positions in the patent citation network and to prove forward patent citation has statistical significance to EV startups' market value, but backward patent citation does not. This study is conducted in four steps: data collection and process, network visualization, network description, and quadratic assignment procedures (QAP). Because this study contributes to practical use in the business field, to ease the use by investors from any backgrounds and industries, I chose social network analysis software NodeXL, which is an easy-to-use and widely adopted add-in to Microsoft Excel. The network visualization presents that, the top two most valued U.S.-based EV startups, Tesla and Atieva, stay at the most pivotal position in the patent citation network. QAP shows strong evidence to support that the number of forward citations positively relates to EV startups' market value but does not exhibit evidence to prove the statistical significance of backward citation. The results of my study are expected to shed some insights for investors to find investment targets.

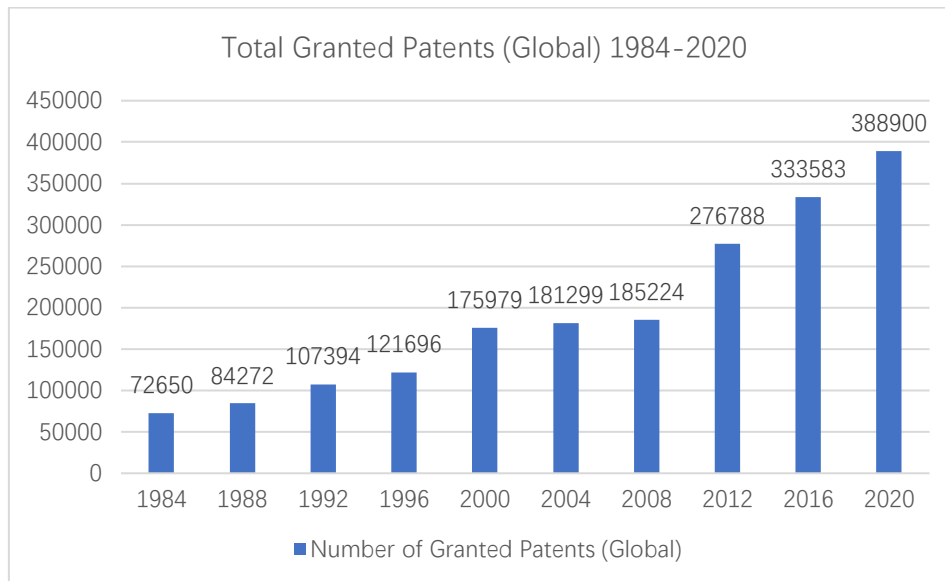
## **Introduction**

Today, the core competitiveness of companies depends less on physical assets but more on knowledge, technologies, patents, and other intangible assets (Madhani, 2012). Patents grant

inventors the monopoly right about invention (“Elon Musk refuses to file patents at his high-tech private space company”, 2013). Such a monopoly right can generate revenues and sometimes create competitive market advantages for companies (Hagelin, 2005). Ernst et al. (2003) showed how important patents are for running a company: patents could support R&D investment decisions, protect products, and maximize intangible asset value via the patent portfolio. For investors, patents also play a crucial role. Lin et al. (2006) mentioned that a company’s patent portfolio could indicate: (1) technology diversity; (2) product diversity, both of which can impact company market value. In sum, patent portfolio functions as an essential source of competitive advantage either for mature companies to maintain and expand market share or for new entries to distinguish themselves and compete with others. Companies have been actively committed to filing patents for the sake of higher evaluation (Chen, 2010; Pratty, 2022). However, one weakness of adopting intangible assets to evaluate company value is the lack of a systematical quantitative approach to examine, evaluate, and compare the quality of patents, such as leading position in technology flow, influence on later inventions, innovation value, etc. (Chen, 2010; “Patent valuation indicators”, n.d.). Financial reports, the most significant source for investors to access information about public companies, only provide primary information about intangible assets or details of R&D (Breitzman, 2002). For example, the China Security Regulatory only requires public companies to include the name and a brief introduction for each patent in financial reports. For the R&D section, total expenditure and R&D department personnel with an education background are the only information required to be reported. Such information is not sufficient to delineate a panorama of a company’s technology strength, including core

techniques, comparative advantages, as well as management strategies. As a result, financial reports are short of substantiated information for investors to evaluate the technology strength of a company and then make wise investment decisions. Fortunately, patent applications and grants require detailed company value-relevant information (Martens, 2019), which contains the complete technical details of the invention such that a person with ordinary skill in the field can understand the description of the invention and if necessary, can implement the invention accessible to the public. Investors can grasp information from these patent files. Therefore, patents contribute an irreplaceable source of information for investors to evaluate the company value. Patents usually directly relate to products, core technologies, and inventors (source of a company's technology). Additionally, changes in patent information can reflect a company's fluctuation in R&D activity and inflow and outflow of technology resources. An approach to include patent quantitative analysis in the evaluation of company market value can help investors to develop better knowledge about company technology strength and make wiser investment decisions. Admittedly, to patent an idea means to publish it. Some companies would rather keep valuable ideas secret to prevent intellectual property from leaking, but this should not impede us from evaluating those having patents.

**Table 1: number of total granted patents globally every 4 years from 1984 to 2020**



In table 1, a clear increasing pattern of the total number of granted patents globally each year can be observed.

Noticeably, since 2010's, the speed of increasing has been further accelerated.

In my study, I crawled data from US Patent and Trademark Office's (USPTO) online database, cleaned data, and kept patent number, assignee organization, cited patent number, and cited-by patent number. Then, I constructed networks with the software NodeXL and described and analyzed network graphical characteristics. I summarized the overall network pattern and compared individual companies. I found that companies had distinctive patent citation situations. Some companies have patents citing a lot, while some others have patents being cited frequently. So, I hypothesized that different patent citation situations have different impacts on company market value. For the last step, I conducted QAP twice to examine the relationship between EV startups' market value and forward citation and backward citation, respectively. In sum, my research question has two components: (1) how EV startups with different market values locate in different positions in the patent citation network; (2) how the forward/backward patent citation affects EV startups' market value. My

hypothesis for (1) is that EV startups with higher market values will locate in more central positions in the patent citation network; for (2) is that patents being cited more (more forward citations) are more statistically significant to EV startups' market value than patents citing more (more backward citations).

A network is composed of a set of nodes connected by relations. Social network analysis (SNA) graphically maps such a network and studies the pattern it exhibited. SNA is a popular application in the field of patent citation study because it is a useful tool to illustrate both the overall structure of a network and specific interactions within it. I chose SNA because it works well to present both the holistic technology flow and the role that each node plays within this flow. This feature makes SNA compatible with my research question: to present that more valuable EV startups locate at more central positions in the patent citation network. SNA can statistically show how central the position is, and has specific indexes to measure the extent of the centrality, such as density and betweenness. Density is how many connections every EV startups have with other companies. Betweenness measures how many times these startups stay on a path linking another pair of companies. With these data, my study is more accurate than observation only.

I chose the QAP regression to prove the statistical significance of the forward citation to EV startups' market value. For network dyadic data, data in the format of the matrix and for the measurement of a pair of entities, it is hard to apply ordinary least squares (OLS) regression because OLS assumes observations to be independent and identically distributed (Lee, 2016).

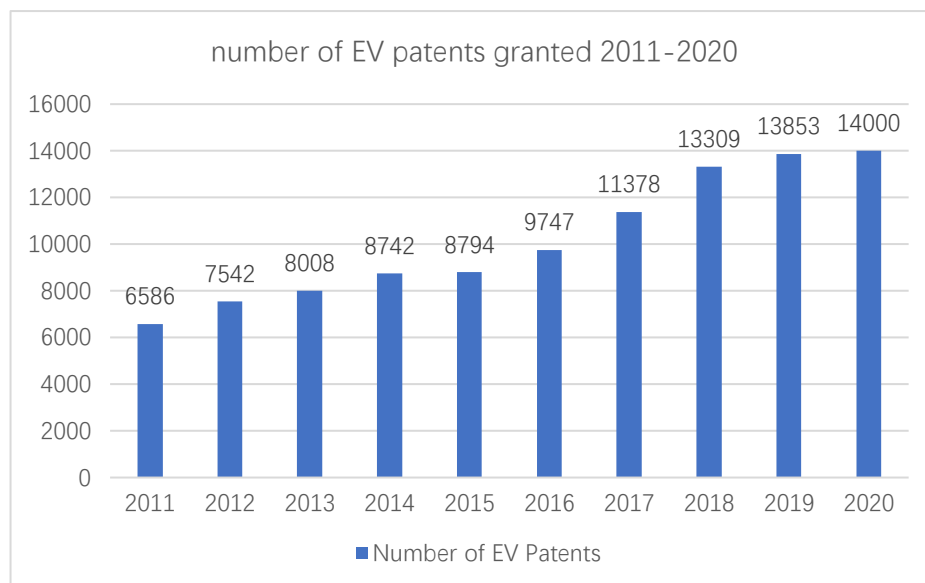
However, in the network, nodes are connected by edges, implying a potentially dependent relationship. Therefore, the assumption of independence for the OLS regression could hardly be satisfied. Another issue caused by using OLS to deal with dyadic data is the correlation. Correlated sample data can inflate or deflate the standard error, resulting in an invalid OLS model. In the contrast to OLS, the QAP regression uses permutation (Krackhardt, 1988). During QAP, rows and columns of the dependent matrix are permuted multiple times. Such a procedure generates multiple random datasets with the dependent matrix while removing the relationship between the dependent and independent matrices at the same time. As a result, coefficients from all permuted datasets can be analyzed and compared. Thus, QAP turns out the correlation between the independent matrix and dependent matrix, which can be used to test the null hypothesis of the regression. The null hypothesis is that there does not exist a relationship between dependent and independent matrices. However, when the p-value is small, QAP provides strong evidence to prove the existence of such a relationship. In my study, I need to prove the positive statistical relationship between the number of forward citations and EV startups' market value. In my patent citation network, EV startups are linked by patent citations, so they may not be independent: if a startup, A, shares some technologies with both startups B and C, and cites both a lot, B and C might cite each other a lot as well. Thus, I chose the QAP regression to solve the issue of dependence and correlation.

### **Basics of Patent and patent citation**

The patent protects inventors by granting a 20-year exclusive right on the invention (Chakraborty, 2020). Potential profits accompanied by the monopoly right encourage

innovation and technological development and prompt companies to commit to obtaining patents (Caviggioli, 2020). In 2019, 391,103 patents were granted by the US Patent and Trademark Office, hitting a new record (“U.S. Patent Statistics Chart Calendar”, 2020). However, even though patents are so common, measuring the value of patents is complicated (Og, 2020). Patent value evaluation has to be multifaceted. Og et al. (2020) found a positive relationship between patent family size, number of claims, forward citations, and patent value. The Tech Line database built by CHI Research, Inc. adopts seven measurements to evaluate the technology competent of companies: the number of patents, cites per patent, current impact index, technology strength, technology cycle time, science linkage, and science strength (Huang, 2012).

**Table 2: number of total granted patents in the field of electric vehicles from 2011 to 2020**



The total number of granted EV patents doubled during the latest decade. Since 2016, the number of EV patents has been increasing at a quicker speed.

Patent citation is an important channel by which technology flows (Duguet, 2003). Patent citation is categorized by backward citation and forward citation. According to the definition by USPTO, the forward citation is a citation made to a published patent by a patent issued later; the backward citation is a citation made by a patent to another patent as prior art. Citation is the bridge connecting generations of inventors and building technology flow (“Citation of Prior Art and Written Statements by USPTO”, 2022). Patents being cited more is more likely to be the core technology within their industries (Sternitzke, 2008). They are expected to lead progress in industry development (Wang, 2018). Patent citing more might combine prior works and have a synergistic effect. Indeed, citation is not the only factor that determines the patent value. Highly cited patents do not necessarily bring the most profits or make the most important innovation. However, citation is still the most indicator to measure the impact of patents (Choe, 2013).

### **Literature review of SNA for patent citation and QAP**

SNA has been popular in patent citation studies. Networks can be constructed in various ways according to patent assignees, inventors, countries, companies, etc. Sternitzke et al. (2008) constructed patent citation networks in unit of patent applicants. Patents with a high frequency of citing or being cited were identified as important players in technology fields. Then, the network presented the extent of patents building upon other knowledge. Closeness between nodes was regarded as technological relation. Sternitzke et al. considered a pair of closely related patents as engaging in either cooperation or competition. The technology competence between companies is the core finding of this study. In addition, the network

result exhibited that highly active patent applicants are responsible for over 80% of all patent families. Also, Sternitzke et al.'s patent citation networks revealed further information including technology similarities and distinctions via citation paths; the situation in their competitive environment via the number of citations; active technology areas via clustering. One conclusion reached by Sternitzke et al.'s study is that the highlight of the citation network not only frequently cited previous patents but also was extensively cited. This finding enlightened me to make my main hypothesis: patents highly cited should have a significant impact on their assignees' market value. In my study, I will explore the statistical relationship between patent citations and EV startups' market value.

Networks can be constructed in different subgroups. Choe et al. (2013) contributed to patent citation network analysis in the organic photovoltaic cells industry in the unit of countries, institutions, and technology fields respectively. Choe et al. ranked the extent of the importance of players in the citation network and determined the degree of knowledge transfer efficiency by comparing networks they constructed and random networks of the same size. Take the institution citation network for instance. Choe et al. found that Konarka Technologies, JX Crystals, and Universal Display were the top 3 most important brokers in the citation network in the unit of the institution since they had the highest in-degree centrality. Graphically, these three institutions formed three hubs. Konarka Technologies ranked top 1 in both in-degree and out-degree centralities, indicating that it was the most important player in the institution citation network. In comparison to a random network, the institution network exhibited a smaller average path, larger clustering coefficient, and higher

degree centralization index, indicating that, in the institution network, knowledge was transferred more effectively; players gathered more closely.

Patent citation analysis can also turn out clues about long-term industry development history.

Wang et al. (2018) constructed patent citation networks by using companies as nodes and cumulative citations as links to compare technologies trajectories of solar photovoltaic (PV) and solar water heaters (SWH) in China from 1995 to 2016. Both PV and SWH networks showed an increasing density, shorter average path, and less fragmentation after 2008, meaning that knowledge diffused faster. Overall, the PV network presented a much higher density than the SWH network. Such a difference illustrated the divergent development pace between PV and SWH industries. The former is rising and fast developing while the latter is becoming obsolete.

There are previous works using patent citation for company market value evaluation.

Breitzman et al. (2002) used patent citation analysis to evaluate merger and acquisition targets. Breitzman et al. defined the technology score by several metrics, including the number of patents, the number of patents granted in latest three years, and the number of highly cited patents. The number of patents determined whether a company relied on a single patent or a series of patents. The latter might suggest better protection for company technology. The condition of the latest three years examined the ability to continue innovating. The number of highly cited patents aimed to weigh high-quality patents more.

The regression showed the technology score positively correlated with company's market-to-book value.

QAP has been widely adopted in studies involving social network analysis. Lee et al. (2016) studied the relationship between various robot technologies to examine the convergence of robot technology. They mapped the convergence network of robot technologies and then used QAP to determine what factors affect the convergence between two robot technologies. Their dependent matrix was the co-occurrence matrix, in which the rows and columns represented robot technologies, and the elements were the number of patents that contained the keywords matching corresponding technologies. The independent matrix was the network centrality index. For instance, the betweenness centrality of a node measured the number of shortest paths between nodes of other pairs that passed through the node. A node acting as a bridge between many groups would have a high betweenness centrality. QAP results exhibited the standardized coefficient and p-value corresponding to each robot technology. Lee et al. derived the conclusion from the sign of coefficients. For example, the positive coefficient of betweenness centrality indicated that technologies connecting more pairs of other technologies tend to converge to each other. The study by Lee et al. has an important implication for mine. They found network centrality indexes essential for robot technology convergence by the hypothesis test conducted by QAP. In addition, they determined the statistical positive/negative relationship by the sign of QAP coefficients. I will adopt a similar approach: to use QAP to conduct the hypothesis test and examine the sign of QAP

coefficients to prove my main hypothesis: patent forward citation positively relates to EV startups' market value.

These previous studies come from distinctive fields but share similarities. First, they proved that patent citation is closely related to the technological impact of patents. Second, patents could represent the technological strength of a company. Third, the patent citation network worked well in presenting the position of each player (company, institution, etc.) in technology flow, and, more importantly, by comparing network centrality indexes, we are able to know whether a node locates at a central position and functions as an important player, or at a peripheral position. However, previous works focused more on finding key nodes or graphically comparing networks. In addition to graphical analysis, I plan to explore the statistical relationship between patent citation and company market value by using companies as nodes and linking them via patent citations in a social network. I will observe graphical characteristics and compare network centrality indexes, such as betweenness, closeness, and average path length. My hypothesis is that if a company has a higher market value, it should locate at a more central position in the patent citation network, indicated by higher betweenness or closeness centrality, or higher density. My work also suggests that two different kinds of citation records, patents with more forward citations (been cited more), and patents with more backward citations (cite more), impact company market value differently. Because some companies with fewer backward citations but more forward citations stay at a more central position in the graphics I obtained, I hypothesize that forward citation has a more significant impact on market value than backward citation. My main contributions are

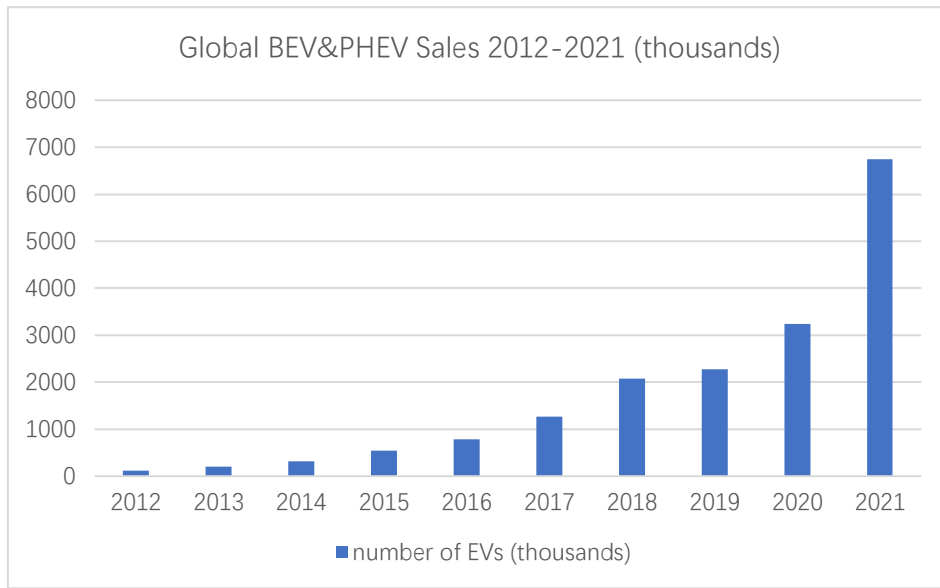
as follows: (1) using social network analysis to present that EV startups with higher market value locate at more pivotal positions in the patent citation network; (2) proving that forward patent citation has a more significant positive impact on company market value than backward patent citation.

### **Electric vehicle industry and patents**

The prosperity of the EV industry originates from EV technology development. During the last century, the invention of electric mobility and rechargeable battery built the foundation of EV technology (“Timeline: History of the electric car”, 2022). From the earliest cumbersome and unsustainable electric mobility to today’s substitute of internal combustion engine (ICE) vehicles, the EV industry has finally become one of the hottest sectors and formed a strong market, where traditional automakers and new startups have been actively competing. As a result, patents are so important in protecting players to distinguish themselves and survive. To compete with traditional automakers who have already patented their technologies is not easy (Thomas, 2019). Startups are pushed to adopt more aggressive business strategies. New entries invent the carbon fiber body to build lighter-weight cars, and find new materials like Lithium cobalt oxide battery and nickel cobalt manganese Lithium-ion battery to increase battery life (Liu, 2022). There have been massive activities related to batteries in an attempt to address problems like how much energy batteries can charge, how to improve battery management, etc. In sum, patents cannot be ignored when studying the relationship between EV company market value and technology strength. Since it takes time for patents to accumulate forward citations (Aristodemou, 2018), patents granted earlier should have more

forward citations than those granted later. Without weighing the factor of time, the value of cutting-edge patents can be undervalued. To avoid inaccuracy caused by the time gap, this study will focus on EV startups and their patents, which have similar ages.

**Table 3: number of electric vehicles sold annually from 2012 to 2021 (in thousands)**



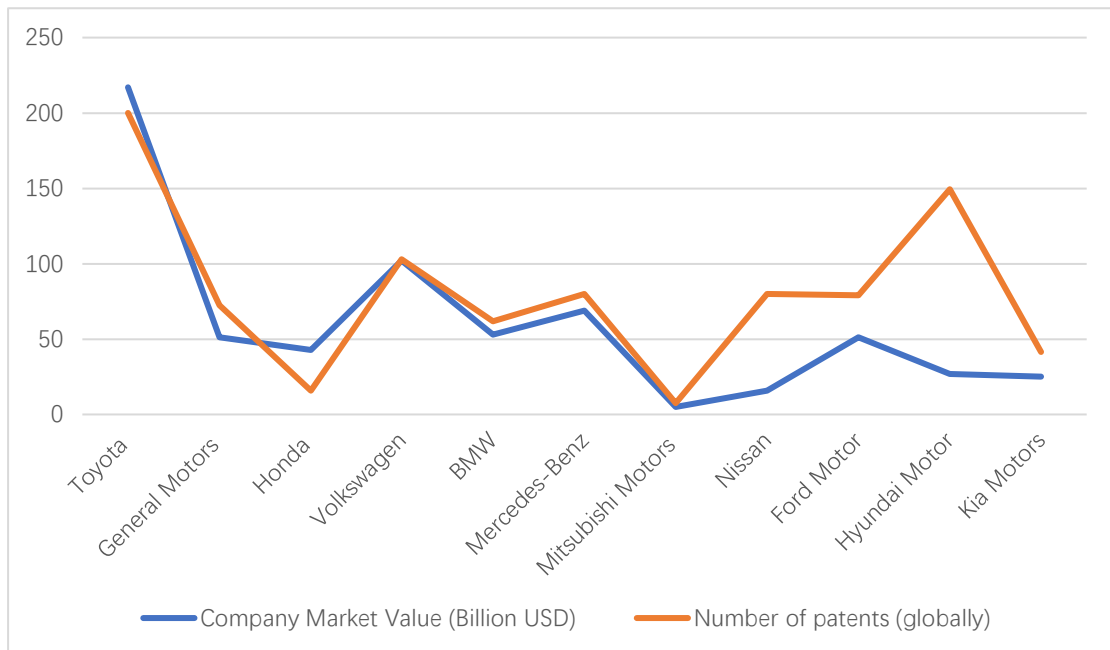
The number of annually sold EVs has exponentially increased since 2014 when policies supporting the manufacture and sales of EVs were introduced by many governments. Policies are one of the most important factors contributing to the prosperity of the EV industry. Many successful EV startups were established in 2014.

**Table 4: traditional automaker market value**

Traditional Vehicle Manufacture Company	Company Market Value (Billion USD)
Toyota	217
General Motors	51
Honda	43
Volkswagen	102
Volvo	43
BMW	53

Mercedes-Benz	69
Mitsubishi Motors	5
Nissan	16
Ford Motor	51
Hyundai Motor	27
Kia Motors	25

**Table 5: the number of patents owned by traditional automakers and their market value**



The market value of most traditional automakers stays around 50 billion US dollars. The data comes from Google Finance in June 2022. The two lines in Table 5 are highly similar. Toyota is the most valued traditional automaker and has the most granted patents as well. Such a consistency also can be observed in Volkswagen, BMW, Mercedes-Benz, Mitsubishi Motors, and Kia motors. For some of the rest, such as Honda, Nissan, and Hyundai, there is a mismatch between company market value and the number of patents owned.

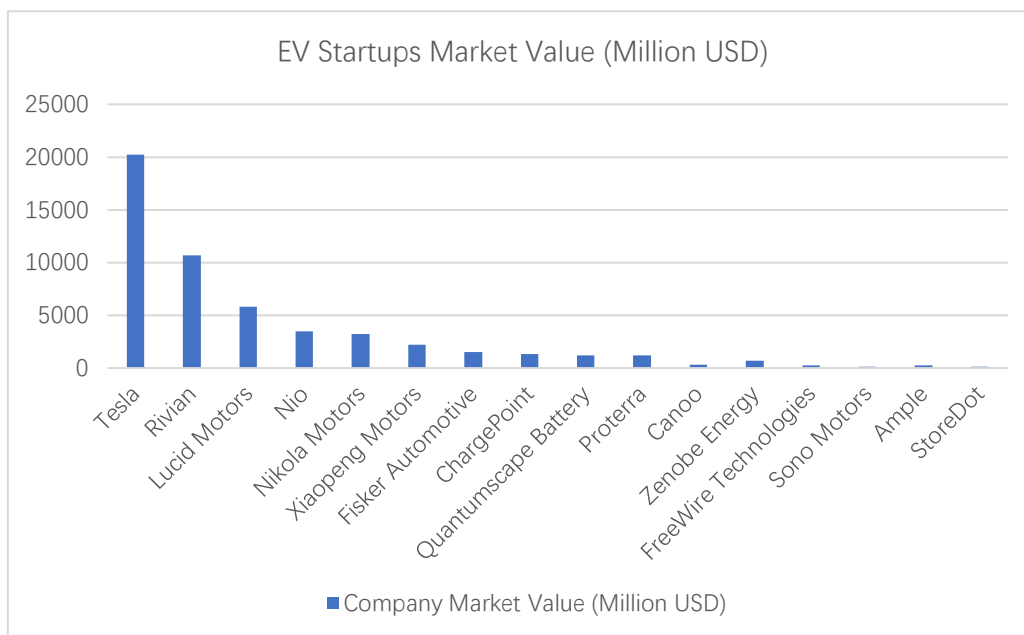
**Table 6: EV startups market value**

Startups in EV industry	Startups Market Value (Million USD)
Tesla	20,200

Rivian	10,700
Atieva	5,800
ChargePoint	1,300
Quantumscape Battery	1,200
Ample	275
FreeWire Technologies	230
StoreDot	146

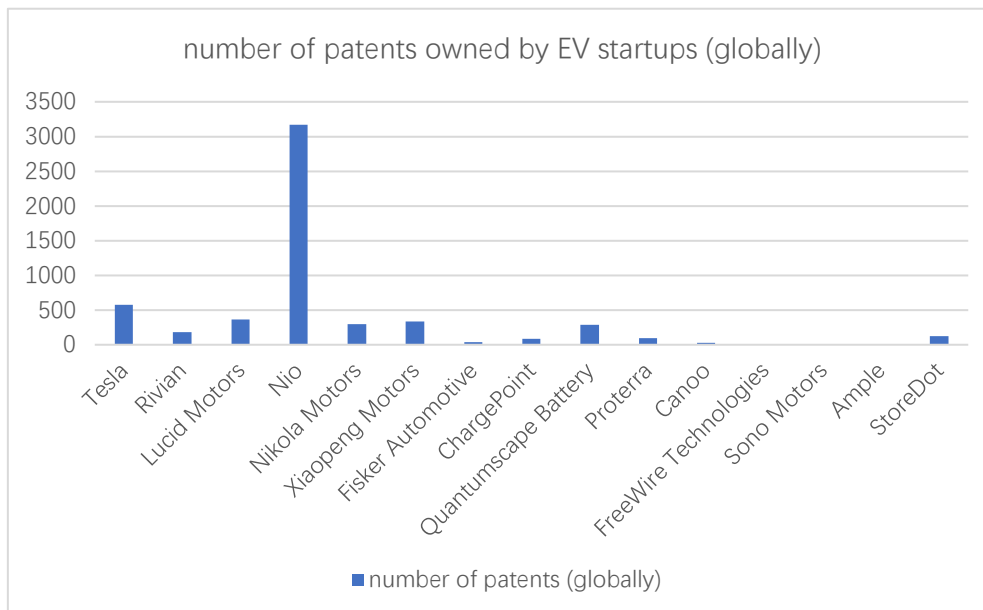
The EV startups market value evaluation comes from the website EnergyStartups (energystartups.org), retrieved on June 17, 2022

**Table 7: EV startups market value**



The EV startups' market value has a salient right-skewed distribution. Tesla and Rivian are the top 2 most valued EV startups. Notably, there is no obvious relationship between startups' market value and the number of patents they owned. This is different from the pattern found exhibited among traditional automakers.

**Table 8: the number of patents owned by EV startups**



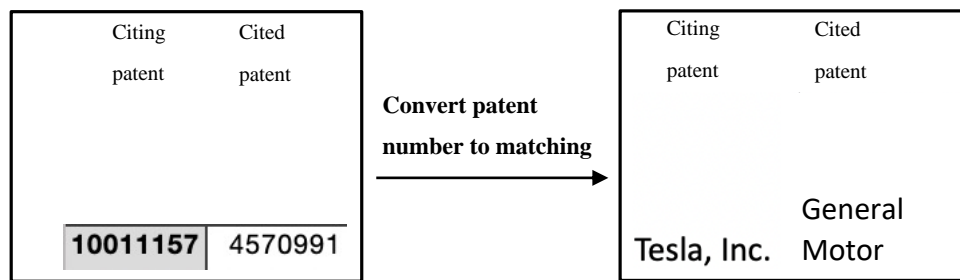
Overall, EV startups own much fewer patents in comparison to traditional automakers. One primary reason is that patent application is a time-consuming process. Also, being more capitalized, traditional automakers are capable of acquiring more patents. While companies have different strategies for patent portfolio management. Tesla made its patents free for everyone's use to prevent technology monopoly as well as influence market standards. More companies use patents to protect their products and the embodiment of technology strength.

### Data collection and process

I collected 152 patents under 9 EV startups with 880 citation records in total. The 9 EV startups were randomly selected by BINOM.DIST formula in Excel among all the total 22 U.S.-based EV startups from the EV startups list retrieved from [energystartups.org/top/electric-cars](http://energystartups.org/top/electric-cars). Patent data was collected from USPTO online database and saved as csv files ([patentsview.org/apis](https://patentsview.org/apis)). Each patent file contains information such as patent name, abstract, granted year, assignee, assignee country, inventor, inventory country, and technology field. In my study, patent data were crawled if their abstract included any one of the EV key technologies glossaries, which were referred from an EV introductory file by

General Motors [7]. Only patent number, assignee organization, cited patent number, and cited-by patent number were kept and saved in the unit of companies. Then, I created patent cite files and cited-by files for each company. For instance, Tesla had two files, one recording forward citation data and the other one recording backward citation data. Both files had the same format: two columns of patent numbers. In each row, the patent in the left column cited the one in the right column. If a patent has 3 citation records, it will have 3 rows. Next, for each cell in all these files, I converted the patent number to its assignee. Table 9 shows this process:

**Table 9: data processing**



In table 9 example, patent 10011157 cites patent 4570991, which is crawled from the USPTO database. Next, I converted the two patent numbers to corresponding assignee names. Patent 10011157 is assigned to Tesla and patent 4570991 is assigned to General Motor.

There is no missing data in this study. For all selected companies, their patent data is accessible from the USPTO dataset because they are all U.S. based. One limitation of the dataset is that some companies might have patents not collected by the USPTO database. But no data is deliberately ignored. If there is any missing data, it would be uniformly distributed among the 9 companies. Thus, the final result should not be impacted.

Finally, I obtained cite and cite-by files in the unit of companies. Every file has two columns of patent assignees. I saved all data in csv files so that they could be directly dealt with by NodeXL, an add-in to the Microsoft Excel. NodeXL is a widely adopted and easy-to-use tool for visualizing and analyzing social network data across a range of fields for both novices and experts (Smith, 2009). In comparison to other network analysis tools like Pajek, UCINet, and NetDraw, NodeXL has much easier data handling and more flexible graphing features while keeping visualization high-quality at the same time. Furthermore, the most outstanding advantage of NodeXL is that it integrates the whole process of social network analysis. By extending existing features of Excel spreadsheets, NodeXL allows users to collect and store data, visualize the network, analyze data by various algorithms and obtain network metrics, and publish reports. Users can do these steps either manually or automatically. Therefore, NodeXL largely lowers the barrier of social network analysis and achieves wider adoption. By virtue of these advantages, my study can be easily replicated by investors to evaluate the market value of companies in any other field as long as patent citation data is accessible. More importantly, NodeXL is able to create readable network graphs with a highlight of key nodes ("Patents: Citing patents", n.d.), meanwhile, provide statistics such as indegree, outdegree, clustering coefficient, closeness, and betweenness. Unlike UCINet, NodeXL clearly presents these network metrics by statistics reports, graphing in bar charts. These reports contain a summary of the quantitative information associated with the overall structure of the network. These network centrality indexes and reports are crucial to solve my research question and can help to prove the main hypothesis: to present that more valuable companies locate at more central positions. I can compare these indexes of EV startups with

different market values to examine if more valuable startups are associated with higher network centrality.

With data processed, I imported data and created network graphs. In graphs, patent assignees are nodes and patent citation records are edges. In a directed network, edges can be inward or outward. Assume we have a node M, any edge pointing out from M to other nodes is outward; any edge pointing to M from other nodes is inward. As mentioned previously, the patent citation is categorized by forward citation and backward citation. In this study, if an edge represents a forward citation, it is outward; if it represents a backward citation, it is inward. In terms of network centrality indexes, outward and inward edges are measured by as outdegree and indegree. In a directed network, the indegree of a node is the number of inward connections it has with neighbor nodes; the outdegree is the number of outward connections. In this study, a company that has more highly cited patents will have larger outdegree; a company that has more patents with many backward citations will have larger indegree.

**Table 10: EV key technologies (electric vehicles glossaries)**

<u>Technology Keywords</u>	<u>Functions</u>
Ampere	a unit used to measure electric current, usually used in the context of EV charging
Anode	The part of a battery that releases electrons during discharge
Battery	The most essential part of EV
Cathode	The part of a battery from which the current leaves the cell
Drive Units	The part giving EVs horsepower and torque

Electric Vehicle	A vehicle that is entirely or partially powered by electricity
Fast Charging	Accelerating charging, usually available at public stations
Range	The total distance an electric vehicle can travel on one full charge before the battery needs to be recharged

**Table 11: network analysis index and function**

Network Analysis Index	Definition
The number of nodes and links	Nodes are individual actors within the network; ties represent the flow of relationships between the actors. In this study, nodes are patent assignee organizations, including EV startups and other companies; links are patent citation records.
Density	Density measures the proportion of possible lines that are present in the graph. In this study, the density equals to actual citations divided by total possible citations.
Degree	Degree is a measure of activity of the actor it represents, which can be found by counting the number of lines incident with it. In this study, the degree measures the number of citations.
Indegree of a node	Indegree counts the number of nodes that are adjacent to any certain node $n(i)$ . In this study, indegree counts the number of forward citations.
Outdegree of a node	Outdegree counts the number of nodes adjacent from any certain node $n(i)$ . In this study, outdegree counts the number of backward citations.
path length	The path length measures the average number of steps along the shortest paths for all possible pairs of network nodes, measuring connectivity of the network. In this study, the path length indicates the efficiency of technology flow in the EV industry. The shorter the path length, the technology spread more efficiently.
Diameter	The diameter is defined as the length of the largest geodesic between any pair of nodes.

In this study, the diameter relates to the technology flow efficiency as well. The longer the diameter, the harder and slower for knowledge to spread.

Components

Component is a subgraph in which there is a path between all pairs of nodes in the subgraph. Component illustrates if nodes are isolated or not. In this study, if the network is constituted by few of components, patent assignees are technologically connected. Conversely, assignees barely have connections through patent citation.

Actor betweenness centrality

An actor (nodes) is central if it lies between other actors. The actor betweenness centrality is high if the node locates at a central position in the network. The betweenness centrality is essential in this study, because it shows whether a patent assignee is either central or peripheral in the technology flow. The main hypothesis is that if the patent assignee takes a more central position in the patent citation network (higher betweenness centrality), it will be valued more by the financial market.

Actor closeness centrality

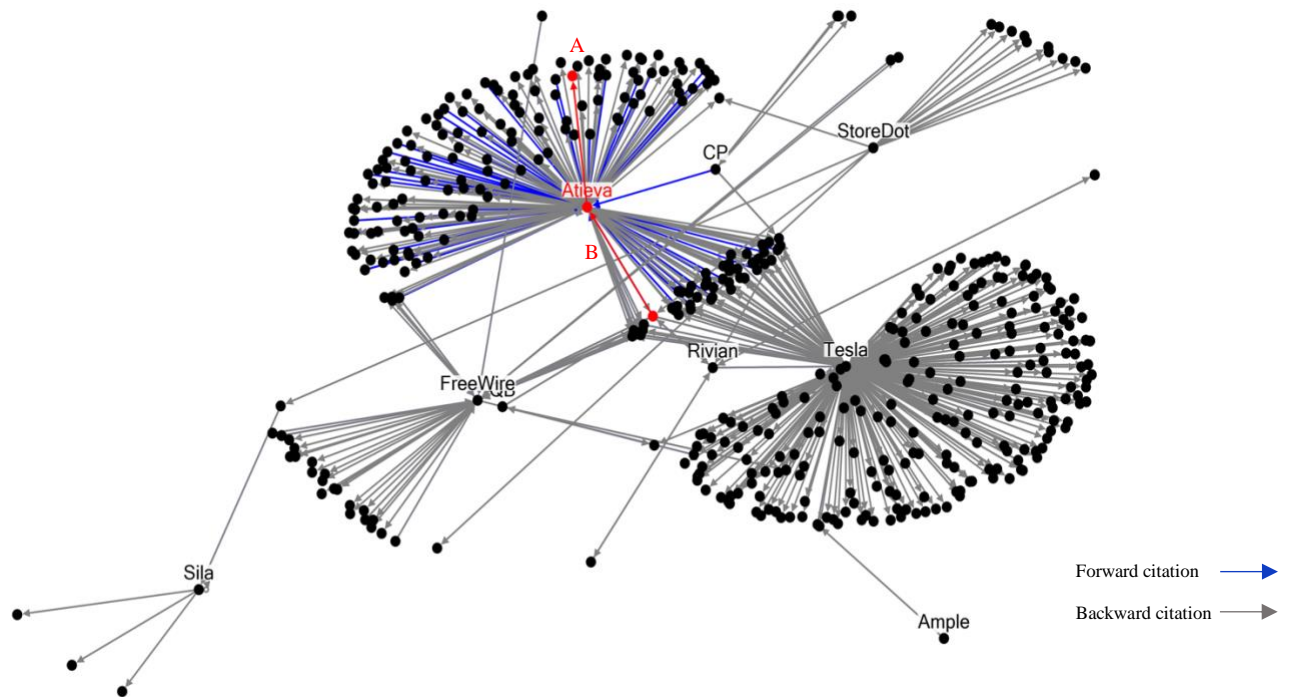
An actor is central if it can quickly interact with all others. Centrality is inversely related to distance. Actor closeness centrality equals 1 when the actor is adjacent to all other actors. In this study, the actor closeness centrality reflects the importance of patent assignee in the technology flow as well. If a patent assignee has an actor closeness centrality 1, it is actively interacting with other assignees and thus plays an important role in technology spread.

All definitions of network centrality indexes come from the book Social Network Analysis by Wasserman and

Faust, (1994) [24, 25].

## Network visualization and network topological analysis results

Table 12: forward and backward citations network graph



In this study, all networks were graphed by NodeXL and laid out by the Harel-Koren fast multiscale algorithm. Table 10 presents a directed network. In this network, nodes are patent assignees, most of which are EV startups and technology companies. The focus of this study, the 9 EV startups were labeled by company names. Some were labeled by the abbreviation: Quantumscape Battery was shortened by QB; ChargePoint was tagged by CP. The rest of the unlabeled nodes are miscellaneous companies which have cited or been cited by these 9 EV startups but are not the focus of this study. Again, forward and backward citations determined the direction of edges. Forward citations correspond to inward edges (blue edges); backward citations match outward citations (grey edges). For example, Atieva has patent 9896093, which cited patents 6604042, 6607255, 7109854, 7881865, 7957874, and 8452464. So, patent 9896093 would create 6 outward citations from the node Atieva like the red line segment A. At the same time, patent 9896093 was cited by patents 10507829, 10820182, 10820349,

10939471, 11116010, 11160111. Thus, patent 9896093 creates 6 inward citations to node

Atieva like the red line segment B.

**Table 13: forward and backward citations network graph (structure) statistical results**

Network	No. of nodes	No. of unique edges	Density	Average degree	Average path length	Diameter	No. of components
forward and backward citations network	454	472	0.003	57.89	2.98	8	1

Findings from the forward and backward patent citation network (table 8):

(1) There are 404 unique backward citations and 68 unique forward citations in total. We can

observe that, overall, companies have more backward citations than forward citations.

This finding enlightens us that forward citation and backward citation should be

differently weighted for company valuation. It might also cause the right-skewed EV

startup's market value. Specifically, startups having patents with a large number of

forward citations might be very popular in the financial market. I will attempt to prove

this in the discussion section.

(2) From the network graph, it can be observed that two nodes, Tesla and Atieva, locate at

very central positions and have very high density. Tesla has the highest normalized

betweenness centrality (2.13) and closeness centrality (1.85), indicating it stays at the

most pivotal position in the network. Atieva, which has the second most valuable startup,

has the second highest normalized betweenness (3.76) and closeness (0.89). As the top 2

valued EV startups, Tesla and Atieva both have developed mature technologies and

achieved mass production. Tesla has an annual production capacity of 600,000 vehicles.

Atieva can produce 12,000 to 14,000 vehicles every year. For the rest of the startups,

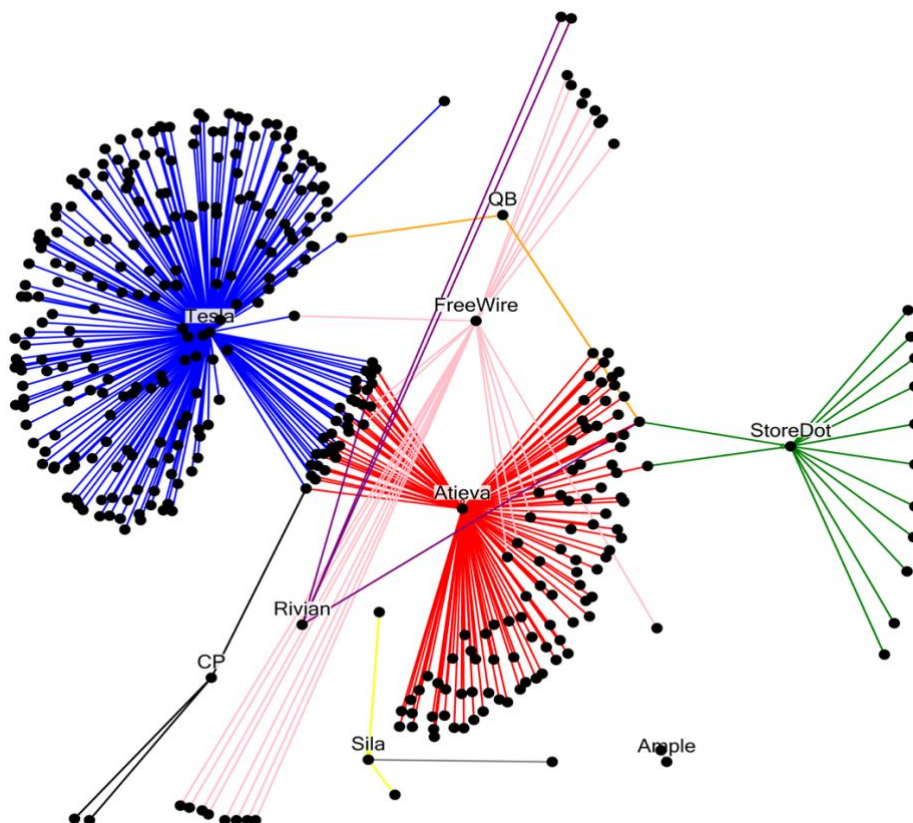
neither backward nor forward citations exceed 30 times. Locating on the central position in the network, Tesla and Atieva have the most connections with other companies. Tesla EV relative patents have cited previous patents from 246 companies and been cited by 30 companies. Companies having cited Tesla come from countries all over the world, mostly from the U.S., Japan, and China. It indicates that Tesla has been actively participating in the global EV technology flow. Atieva has cited 130 companies and has been cited by 73 companies. Notably, Atieva has more forward citations. It is cited by both automakers such as Toyota, NIO, Ford, Mitsubishi, and companies from other fields, like Shape Corporation, which is a supplier of lightweight materials. Companies having citation relationships with Atieva are mostly from the U.S., Japan, and China, which is similar to findings about Tesla.

- (3) Among the rest of the EV startups, connections established by patent forward and backward citations exhibit much lower density. Notably, even having a similar number of granted patents, companies can have very different forward and backward citation situations. FreeWire Technologies, Inc. was established in 2014. It has 9 granted patents and 3 of them have EV glossaries in the abstract. These 3 patents have 24 backward citations and 13 forward citations in total. Both exceed the average forward and backward citations of companies with a similar scale, which brings FreeWire to a more central position in the patent citation network. FreeWire has higher betweenness (0.59) and closeness (0.92) centralities than companies with a similar scale. Ample, Inc. is a company close to FreeWire in size and age. Ample Inc. was also founded in 2014 and awarded as the TIME best invention in 2021. Ample has 7 patents, 6 among which

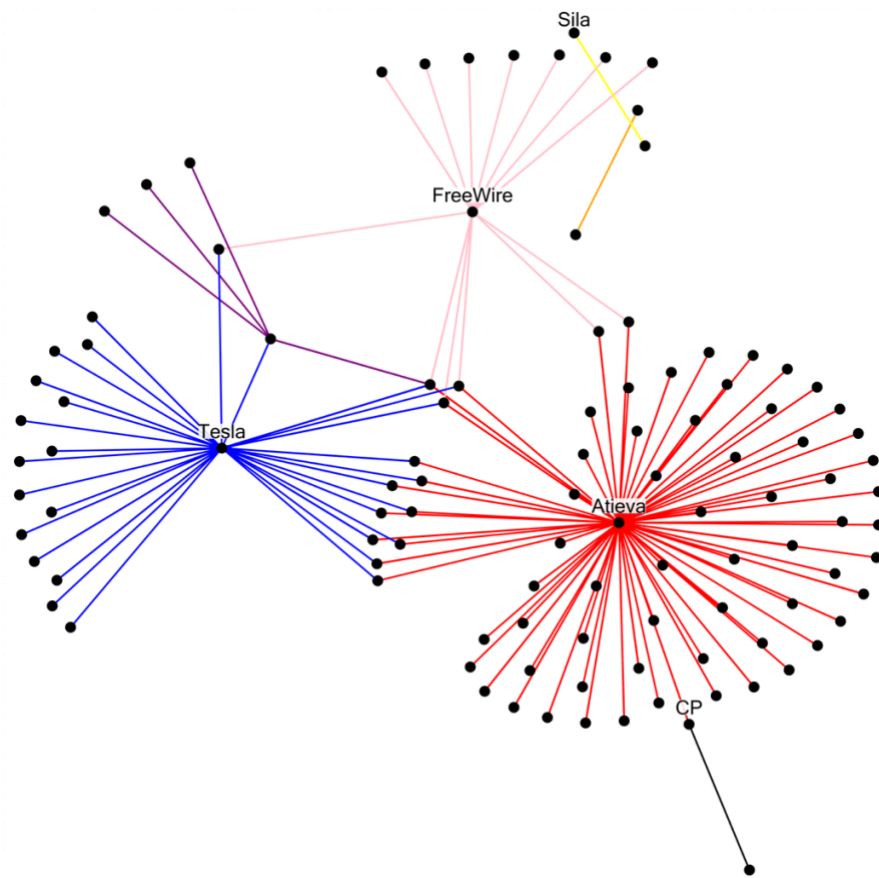
include EV glossaries in abstracts. But only 2 patents owned by Ample have forward citation records, both of which cited Apple, Inc. As a result, Ample is very distant from other companies and almost isolated in the network. It has low betweenness (0.21) as well as closeness (0.06).

- (4) There is only 1 component in the network. So, all companies are directly or indirectly connected. This indicates that, in the EV industry, companies are technologically connected. Such a high connectivity results in knowledge transforming easily in the EV industry, thus associating with a high risk of being plagiarized. Thus, patents are significant in protecting intangible assets and making companies differentiated from others [33]. This finding is consistent with the main hypothesis that companies with higher market value locate at more central positions in the network.

**Table 14: Backward citations network graph**



**Table 15: Forward citations network graph**



**Table 16: backward citations and forward citations network graph statistical information**

Network	No. of nodes	No. of edges	Density	Average degree	Average path length	Diameter	No. of components
Backward citation	388	484	0.01	2.20	2.98	6	3
Forward citation	107	396	0.02	2.30	2.70	5	3

In both table 14 and table 15, networks are undirected. Table 14 presents the backward citation network graph and table 15 exhibits the forward citation network. In both charts, edges were colored according to the company owning the corresponding patent. In table 14, if patent A owned by Tesla and A has a backward citation from patent B, the edge from B to A

is blue. In table 15, if patent C owned by Atieva and C has a forward citation to patent D, the edge from C to D is red. Table 17 reports the color separation.

**Table 17: edge color separation**

EV startups	Tesla	StoreDot	Sila	Quantumscape Battery	ChargePoint	Rivian	FreeWire	Ample	Atieva
Color	Blue	Green	Yellow	Orange	Black	Purple	Pink	Grey	Red

Findings from table 14 and table 15:

- (1) In table 15, The minimum degree is 0, which means that some companies have never been cited. The maximum degree is 73. So the most cited company has 73 times forward citations. The average degree is 13.67 and the median is 2. The number of forward citations is right-skewed and only a few of patents have been highly cited.
- (2) In table 14, the minimum degree is 1, which illustrates that each company has cited some prior work at least once. The maximum degree is 246, which is larger than the maximal degree; the average degree is 47.67; the median degree is 5. Overall, backward citation is much more common than forward citation, which gives an insight that forward citation should be weighed more than backward citation when quantifying patent value. I will test this finding in the discussion section.
- (3) In the forward citation network, the density is 0.01; the average path length is 2.98; the diameter is 6. In the backward citation network, the density is 0.02; the average path length is 2.70; the diameter is 5. The backward citation network exhibits a higher density, shorter average path, and shorter diameter than the forward citation network. Thus, more

players participate in the backward citation network. Players are more connected, and knowledge can be transferred more efficiently via backward citation.

### Quadratic Assignment Procedure (QAP)

I implemented QAP to further examine the statistical relationship between EV startups' market value and the forward and backward citation. I did QAP twice respectively. To examine the statistical relationship between EV startups' market value and the forward citation, in QAP, the market value gap matrix is the dependent matrix, where the rows and columns represent selected EV startups in the same order and the elements are the corresponding market value gap for each pair of startups; the forward citation gap matrix is the independent matrix, where the rows and columns still represent selected EV startups and the elements are the corresponding number of forward citation gap of each pair of startups. For proving the statistical relationship between EV startups' market value and the backward citation, in QAP, the dependent matrix is the same, but the independent matrix is the backward citation gap matrix, where the elements are the corresponding number of backward citation gap of each pair of startups. All values are absolute values. Table 18 shows examples to illustrate dependent and independent matrices.

**Table 18: dependent and independent matrices example**

	Startup1	Startup2	Startup3
Startup1	-	The gap between startup1's market value and startup2's market value	The gap between startup1's market value and startup3's market value
Startup2	The gap between startup2's market value and startup1's market value	-	The gap between startup2's market value and startup3's market value

Startup3	The gap between startup3's market value and startup1's market value	The gap between startup3's market value and startup2's market value	-
----------	---	---	---

	Startup1	Startup2	Startup3
Startup1	-	The gap between startup1's forward citations and startup2's forward citations	The gap between startup1's forward citations and startup3's forward citations
Startup2	The gap between startup2's forward citations and startup1's forward citations	-	The gap between startup2's forward citations and startup3's forward citations
Startup3	The gap between startup3's forward citations and startup1's forward citations	The gap between startup3's forward citations and startup2's forward citations	-

	Startup1	Startup2	Startup3
Startup1	-	The gap between startup1's backward citations and startup2's backward citations	The gap between startup1's backward citations and startup3's backward citations
Startup2	The gap between startup2's backward citations and startup1's backward citations	-	The gap between startup2's backward citations and startup3's backward citations
Startup3	The gap between startup3's backward citations and startup1's backward citations	The gap between startup3's backward citations and startup2's backward citations	-

The rows and columns of the dependent variable matrix were permuted 2000 times. QAP

coefficients and p-values are as follows:

**Table 19: QAP results**

Dependent variable	Standardized coefficient	Standardized error	P-value
Backward citation	0.07	0.21	0.17
Forward citation	0.33	0.03	0.01

For backward citation, though it has a positive coefficient (0.07), the p-value (0.17) is too large. So, there is no strong evidence to conclude that the number of backward citations is statistically significant to EV startup's market value. Conversely, the number of forward citations has both a positive coefficient (0.33) and a small p-value (0.01). Thus, our main hypothesis is proved to be true: The number of forward citation is statistical significant to EV startups' market value, but the backward citation does not. Specifically, EV startups with more forward citations tend to have a higher market value. This result enlightens that forward citation, the value of patents as prior art to future inventions, should be taken into consideration for company valuation.

## **Conclusion**

In previous studies, patent citation network analysis usually focused on topologically describing and comparing networks. My study took one additional step to show how more valuable companies stay at more pivotal positions in the patent citation network. This finding contributes to practical use in the business field. If investors create a similar patent citation network, however, find a company with relatively low market value locate in the central position. It is possible that this company is undervalued and could be a wise investment target. One advantage of the social network analysis software NodeXL is that it is easy-to-use and not limited by fields of study. Investors from different backgrounds and industries can try to reproduce my study. One limitation of my study is that patent data might be incomplete. I chose U.S.-based EV startups and crawled patent data from the USPTO dataset. However, companies participating in international business, such as Tesla, might have applied for

patents in some specific nations and regions other than the U.S. Thus, the patent data set can be incomplete. However, this limit affects all EV startups, and no data is missed deliberately, so it should not significantly affect the final result.

This study used social network analysis to exhibit that more valuable EV startups locate at more central positions in the patent citation network. In addition, this study proved that forward patent citation has statistical significance to EV startups' market value, but backward patent citation does not. I conducted the study in four steps: (1) crawled and processed patent citation data; (2) created social network graphs by using EV startups as nodes and patent citation records as edges via NodeXL; (3) described graphical characteristics and compared them among startups; (4) used QAP to prove the significant impact of forward citation.

Detailed findings are as follow:

(1) The top 2 most valued EV startups, Tesla and Atieva, locate at the most central positions.

They also have much higher betweenness and closeness centrality than other startups.

(2) EV startups are highly connected. The backward and forward citation network has only one component. So, all EV startups are linked by either forward or backward citations.

Technologies easily spread in the EV industry. So, patents are significant in protecting intangible assets and making companies differentiated from others.

(3) Backward citation is more common than forward citation. Startups with more forward citations have larger network centrality indexes and thus locate at more central positions.

The backward citation network also exhibits a higher density than the forward one. This

is consistent with the QAP finding such that forward citation has a statistical significance on EV startups' market value but backward citation does not.

(4) Since the backward citation network is more connected, knowledge flows faster through backward citation.

(5) The number of either forward or backward citations is highly right-skewed, which affects the patent value distribution. As a result, the EV startups' market value is right skewed as well. For top valued startups like Tesla and Atieva, their market value is comparable to the market value of some traditional automakers. While for a majority of EV startups, their market value is much lower.

## **Reference**

Aristodemou, L., Tietze, F. (2018). Citations as a measure of Technological Impact: A review of forward citation-based measures. *World Patent Information*, 53, 39–44.

<https://doi.org/10.1016/j.wpi.2018.05.001>

Breitzman, A., Thomas, P. (2002). Using patent citation analysis to target/value M&A candidates. *Research-Technology Management*, 45(5), 28–36.

<https://doi.org/10.1080/08956308.2002.11671519>

Caviggioli, F., De Marco, A., Montobbio, F., Ughetto, E. (2020). The licensing and selling of inventions by US universities. *Technological Forecasting and Social Change*, 159, 120189.

<https://doi.org/10.1016/j.techfore.2020.120189>

Chakraborty, M., Byshkin, M., Crestani, F. (2020). Patent citation network analysis: A perspective from descriptive statistics and ERGMs. PLOS ONE, 15(12).

<https://doi.org/10.1371/journal.pone.0241797>

Chen, Y.-S., Chang, K.-C. (2010). The relationship between a firm's patent quality and its market value -- the case of US Pharmaceutical Industry. Technological Forecasting and Social Change, 77(1), 20–33. <https://doi.org/10.1016/j.techfore.2009.06.003>

Choe, H., Lee, D. H., Seo, I. W., Kim, H. D. (2013). Patent citation network analysis for the domain of organic photovoltaic cells: Country, institution, and technology field. Renewable and Sustainable Energy Reviews, 26, 492–505. <https://doi.org/10.1016/j.rser.2013.05.037>

Duguet, E., MacGarvie, M. (2003). How well do patent citations measure flows of technology? evidence from french innovation surveys. SSRN Electronic Journal.

<https://doi.org/10.2139/ssrn.452000>

Economist, T. (2013, March 15). Elon Musk refuses to file patents at his high-tech private space company. Business Insider. Retrieved July 25, 2022, from

<https://www.businessinsider.com/why-some-companies-are-not-filing-patents-2013-3>

Electric vehicle glossary: Terms definitions. GM. (n.d.). Retrieved July 5, 2022, from

<https://www.gm.com/electric-vehicles/glossary>

Ernst, H. (2003). Patent information for strategic technology management. *World Patent Information*, 25(3), 233–242. [https://doi.org/10.1016/s0172-2190\(03\)00077-2](https://doi.org/10.1016/s0172-2190(03)00077-2)

Hagelin, Ted. (2005). Valuation of Patent Licenses. *Texas Intellectual Property Law Journal*. 12.

Huang, M.-H., Sung, H.-Y., Wang, C.-C., Chen, D.-Z. (2012). Exploring patent performance and technology interactions of universities, industries, governments and individuals. *Scientometrics*, 96(1), 11–26. <https://doi.org/10.1007/s11192-012-0878-3>

Krackhardt D. Predicting with Networks: Nonparametric Multiple Regression Analysis of Dyadic Data. *Social Networks*. 1988; 10: 359–381

Lee WJ, Lee WK, Sohn SY (2016) Patent Network Analysis and Quadratic Assignment Procedures to Identify the Convergence of Robot Technologies. *PLoS ONE* 11(10): e0165091. <https://doi.org/10.1371/journal.pone.0165091>

Liu, W., Placke, T., Chau, K. T. (2022). Overview of batteries and battery management for Electric Vehicles. *Energy Reports*, 8, 4058–4084. <https://doi.org/10.1016/j.egy.2022.03.016>

Lin, B.-W., Chen, C.-J., Wu, H.-L. (2006). Patent portfolio diversity, technology strategy, and firm value. *IEEE Transactions on Engineering Management*, 53(1), 17–26.

<https://doi.org/10.1109/tem.2005.861813>

Madhani, P. M. (2012). Intangible assets: Value drivers for competitive advantage. *Best Practices in Management Accounting*, 146–165. [https://doi.org/10.1057/9780230361553\\_10](https://doi.org/10.1057/9780230361553_10)

Martens, T. (2019). The disclosure function of the U.S. patent system: Evidence from the PTDL program and extreme snowfall. *SSRN Electronic Journal*.

<https://doi.org/10.2139/ssrn.3347868>

Og, J. Y., Pawelec, K., Kim, B.-K., Paprocki, R., Jeong, E. S. (2020). Measuring patent value indicators with patent renewal information. *Journal of Open Innovation: Technology, Market, and Complexity*, 6(1), 16. <https://doi.org/10.3390/joitmc6010016>

Patent valuation indicators. Intellectual Property Expert Group. (n.d.). Retrieved July 23, 2022, from <https://www.ipeg.com/patent-valuation-indicators/>

Pratty, F. (2022, June 6). What do the EV patent wars mean for startups? Sifted. Retrieved July 25, 2022, from <https://sifted.eu/articles/electric-vehicle-patents-tesla/>

Resources. (n.d.). 2202 Citation of Prior Art and Written Statements [R-08.2017]. United States Patent and Trademark Office. Retrieved July 23, 2022, from <https://www.uspto.gov/web/offices/pac/mpep/s2202.html>

Smith, M. A., Shneiderman, B., Milic-Frayling, N., Mendes Rodrigues, E., Barash, V., Dunne, C., ... & Gleave, E. (2009, June). Analyzing (social media) networks with NodeXL. In Proceedings of the fourth international conference on Communities and technologies (pp. 255-264).

Sternitzke, C., Bartkowski, A., Schramm, R. (2008). Visualizing patent statistics by means of social network analysis tools. *World Patent Information*, 30(2), 115–131.  
<https://doi.org/10.1016/j.wpi.2007.08.003>

Thomas, V. J., Maine, E. (2019). Market entry strategies for electric vehicle start-ups in the automotive industry – lessons from Tesla Motors. *Journal of Cleaner Production*, 235, 653–663. <https://doi.org/10.1016/j.jclepro.2019.06.284>

Timeline: History of the electric car. Energy.gov. (n.d.). Retrieved July 23, 2022, from <https://www.energy.gov/timeline/timeline-history-electric-car>

U.S. Patent Statistics Chart Calendar Years 1963 - 2020. U.S. Patent Statistics Summary Table, Calendar Years 1963 to 2020, 05/2021 update. (n.d.). Retrieved July 16, 2022, from [https://www.uspto.gov/web/offices/ac/ido/oeip/taf/us\\_stat.htm](https://www.uspto.gov/web/offices/ac/ido/oeip/taf/us_stat.htm)

Wang, Y., Urban, F., Zhou, Y., Chen, L. (2018). Comparing the technology trajectories of solar PV and solar water heaters in China: Using a patent lens. *Sustainability*, 10(11), 4166. <https://doi.org/10.3390/su10114166>

Wasserman, S., Milton, F. K. L. (1994). Chapter 3: Notation for Social Network Data. In *Social network analysis: Methods and applications*. essay, Cambridge University Press.

Wasserman, S., Milton, F. K. L. (1994). Chapter 5: Centrality and Prestige. In *Social network analysis: Methods and applications*. essay, Cambridge University Press.