Career Analytics: data-driven analysis of turnover and career paths in knowledge intensive firms: Google, Facebook and others

ABSTRACT

Purpose: The purpose of this study is to demonstrate that by utilizing data-driven tools, based on HR analytics and cluster analysis, management can gain a nuanced perspective on employees’ career path initiatives. We thus lay the groundwork for further theory developments and contribute to the ongoing extant examination of career analytics.

Methods: This study utilizes two data sources. First, a uniquely large data set of over 970,000 curriculum vitas (CV’s). Second, benchmark market data retrieved from annual surveys which measure employee job satisfaction and compensation data retrieved from high technology organizations. These combined data sources enable a robust examination of the relationship between variables by utilizing data-driven methods, e.g., HR analytics, machine-learning and pattern recognition tools.

Results: We found hidden patterns in turnover and career moves which are not necessarily aligned with existing theories from the literature. Moreover, we uncover some disruptive aspects of the turnover phenomenon which are sometimes counterintuitive. Perhaps the most striking finding is an inverse relation between job satisfaction level, average employment period and job seeking patterns in specific high technology companies. We also found that while higher compensation packages often lead to higher job satisfaction, they do not translate into longer employment periods for
some employees. This counterintuitive effect was found to be strongest in global high technology companies, such as Google and Facebook.

**Contribution:** Our study presents three notable contributions. First, from a *theoretical* perspective, our findings advance the motivation literature by highlighting the counterintuitive effect of intrinsic and extrinsic motivation. Second, from a *practical* perspective, our findings contribute to the ongoing extant examination of macro level career analytics. Moreover, our findings lay the groundwork for the adoption of practical career planning management tools, in the form of career analytics, which are at the forefront of the investigation of the organizations of the future. Finally, from a *methodological* perspective, this study contributes to the increasingly growing use of data-driven tools for organizational and human resources investigation.

**Keywords:** Turnover, Knowledge intensive firms, Career analytics, HR analytics, Cluster analysis.
I. Introduction

Leading a successful career, which may result in higher levels of job satisfaction, are important components of individual’s well-being. In recent decades, one’s professional career plays an important role in individual’s self-actualization and sense of accomplishment. Moreover, our modern lives are driven by materialistic components, such as pay and compensation packages, which permit individuals to execute the desired life-style and to achieve social status. Moreover, the post internet society is driven by growing amounts of mega information on the one hand, and a shift in various information search related processes, on the other. As a result, the online career market has become larger and more vibrant than ever before. It has been noted that less than half a decade ago, the weak ties in a social network, were found to be the major source for job hunting related information [1].

The digital world we live in results in weaker social connections between individuals. Furthermore, close relationships, such as family and intimate friends, are found to be less important for fetching information that might result in finding a new job. Web search engines seem to replace information search methods such as word-of-mouth [2], which was once the main source of information for job hunting. This trend drives the development of new HR models, which are based on social networks and artificial intelligence. These new models are converted at times to successful match making between job applicants on the one hand, and employers seeking for talent, on the other hand.

In recent years, labor markets went through fundamental changes which were led by the centrality of the internet in general, and social networks platforms, more specifically. As a result, employers adjusted their recruitment and placement processes accordingly. Nowadays, the term e-HRM is used [3] to describe electronic human
resource management systems and processes at the intersection between human resource management (HRM) and information technology (IT). Notwithstanding, for many individual job applicants, internet-based job search has become a fundamental method towards effective career development. While the web-based job search has become the common method for job hunt, a 2007 meta-analysis of over 30 e-HRM related studies [4] claims that e-HRM is still in its early stages, is quickly developing and is going through rapid fluctuations.

The purpose of our study is to examine the e-HRM arena through the lens of career analytics. This study examines a uniquely large data set and explores vocational related data. The study reveals hidden patterns in employee job satisfaction, turnover and job seeking patterns.

This study utilizes two data sources. First, a data set which includes more than 970k CV’s. Second, benchmark market data retrieved from annual surveys which measures employee job satisfaction and compensation levels in several organizations. These surveys, such as for example, “Fortune 100 Best Companies to Work For at 2018” [5], [6], [7], [8] measures employee job satisfaction in a diverse sample of organizations, and includes salaries, as well as total compensation data [5].

The combined data sources enabled a robust examination of the relationships between three factors. (a) Compensation / Compensation level, (ii) reported employee job satisfaction and (iii) employment period. We assume that the average employment period per employee represents, up to some degree, a realistic and deep aspect of job satisfaction, while we neglect a formal causality analysis and focus on correlation measures, following the conventional practice in this field. Our assumption is supported both practically and theoretically. Practically, it has been accepted that employees stay on a job and/or in the organization as long as they experience job satisfaction.
Theoretically, a large body of literature supports this assumption [9], [10], [11], [12], [13], [14], [15], [16].

This study reveals hidden patterns in turnover which are not necessarily aligned with existing theories from the literature. Moreover, we uncover some disruptive aspects of the turnover phenomenon which are sometimes counterintuitive. Perhaps the most striking finding is an inverse relation between job satisfaction level, average employment period and job seeking patterns in specific high technology companies. We also found that while higher compensation packages often lead to higher job satisfaction, they do not translate into longer employment periods for some employees. This counterintuitive effect was found to be strongest in global high technology companies, such as Google and Facebook.

II. Theoretical Background

A. Early Studies

Employee job satisfaction is a research field as ancient as modern human labor itself. The early published articles, approximately a century ago [14], defined no less than 16 dimensions of work satisfaction. Thirty years later, the common view of work motivation assumed that the motivation for staying in a workplace is rooted in hedonism. Scholars found that employees are engaged in workplace partially because of enjoyment. Additionally, individual’s hedonistic needs, result in an effort to enjoy life in the workplace. Maintaining longer periods of employment can thus be associated with the human effort to prolong personal state of satisfaction [10].

Moreover, differences between genders were found to have an important effect on work motivation [17]. Notwithstanding, for male and female employees alike, turnover (i.e. leaving a job) is associated with low job satisfaction [18]. Employee
commitment and job satisfaction are considered to have complex sets of interactions. It is unclear if commitment creates job satisfaction, or if job satisfaction creates commitment in the workplace. Unfortunately, the relationship between these two important antecedents is still rather unclear [19].

Employee job satisfaction is also rooted in the exploration of motivation theory, thus, a multi-dimensional phenomenon. Literature differentiates between intrinsic motivation factors (i.e. recognition, responsibility) and extrinsic motivation factors (i.e. job security, compensation, employment conditions) [12]. As a result, the discussion is broadened to include variations in explaining motivation, satisfaction and term of employment.

B. Contemporary Works

Employee turnover has long drawn scholarly attention. A recent study summarizes the vast literature of decades of theoretical advancement. This study presents the complex relationship between job satisfaction, motivation and turnover through an in-depth review [13]. It emphasizes the importance of job expectations and their influence on turnover [18]. It also highlights recent approaches, which distinguishes between factors influencing the decision to leave a job and those responsible for staying in a job. Apparently, while employee’s decision to leave a job, or to stay in a job are excluding choices, the factors for leaving a job and the factors for staying in a job are different [16], [20].

We utilize the Job Embeddedness Theory as a central theoretical approach to explain our findings. The Embeddedness Theory completes an earlier theory which describes the reasons for staying in a job as “a net or a web in which an individual can become stuck” [21]. Thus, the decision to stay in a job, includes many emotional
dependent and independent variables which influence the decisions of employees. The Embeddedness Theory enables a more robust and holistic view of vocational turnover processes, through the separation of factors responsible for leaving a job, and factors responsible for staying in it, thus serves us in explaining our study results.

When exploring the advantages and disadvantages of turnover, it is estimated [15] that in a retail setting, for example, one standard deviation increase in turnover, shrinks the annual profit by 8.9%. Nonetheless, while we could not retrieve similar studies with regards to high technology companies, we believe the knowledge required to perform within high technology companies is broader, therefore the learning curve is more complex resulting in a slower learning curve, thus the costs of turnover might be even higher.

Nowadays, turnover is examined in the literature as a more complex multi-dimensional phenomenon. Employee turnover varies between industries and between job types and employee types. Some researchers have distinguished between employees who leave organizations voluntarily and those who are dismissed or who leave as part of a reduction in force. Voluntary turnover captures the departure of employees from organizations on their own free will. Because people who are more skilled or perform better in their jobs are believed to have more external employment opportunities available to them than average or poor employees, they are more likely to leave. This distinction broadens our understanding of how different types of turnover impact organizations and employees alike. However, many organizations fail to differentiate between voluntary and involuntary turnover [13], [15].

Nevertheless, high turnover rates are generally perceived as a negative workforce sign, especially if the employee leaves within a period shorter than one year. Some knowledge intensive companies attempt to reduce turnover rates and thus retain their top talents by
implementing a wide range of legal arrangements, for example patent registrations and legal protections procedures. Within these policies, employees are restricted to legal constraints regarding limitations on knowledge transfer, etc. However, the implementation of legal enforcement as means to reduce turnover rates were found rather unsuccessful, and talented employees leave organizations despite such limitations [11].

III. Methods

In this study we implemented the following methodology. We calculated the average employment period based on a data set of 973,134 CV`s. The data set includes 44 features or variables, including control variables such as: gender, country of employment seniority ranking (low, medium and high management level), number of endorsed skills, employment archival data per last job only (company name, job title, employment duration in previous firms), industry type etc. This data set was merged with two additional data sources, which contributed the employee job satisfaction and the compensation dimensions. This benchmark datasets were collected from surveys [5], [6], [7], [8], where an aggregative score for each company is published. In regards the “Top Best Companies” lists, although much difference exists between the different lists, Google and Facebook were found to be among the top 10 employers in all the lists. The list used in the study was based on the average scores of the lists above. Finally, the following are the dimensions measured in this study: (a) average employment period, (ii) reported job satisfaction and (iii) average salaries (compensation level). We analyzed the data using R software version 3.4.3. Also, we used the data.table package in R, for the purpose of fast table filtering, and the stringer package for the purpose of string manipulation. Finally, we utilized the ggplot2 and ggrepel packages for plotting.
IV. Results

The analysis presented in Figure 1 below demonstrates the relationship between employment durations, satisfaction and compensation levels. As demonstrated in Figure 1, the companies Google and Facebook are located in the bottom left corner. These companies offer higher compensation packages (the red color indicates a compensation package of 130-140k USD per year), and demonstrate higher job satisfaction levels (represented in the $x$-axis, where lower scores reflect an employee’s high satisfaction workplace). Surprisingly, these companies demonstrate low scores of average employment periods (represented in the $y$-axis). Results indicate that Facebook has an average employment duration of 16.9 months.

[INSERT FIGURE 1 ABOUT HERE]

Figure 1 above illustrates the average job satisfaction (represented in the $x$-axis), the average duration in a job (represented in the $y$-axis) and the average compensation package (represented in colors) for 39 organization, for which a valid satisfaction score could be extracted.

As illustrated in Figure 1, organizations such as Facebook or Google, which offer high compensation packages, which seem to influence higher job satisfaction levels, demonstrate a surprisingly low employment periods for their employees. This finding is surprising and counterintuitive also compared to other organizations. For example, comparing this finding to the situation in Intel, which demonstrates an average satisfaction score. However, demonstrates a very long employment duration. Additionally, both Symantec and Bristol-Myers Squibb have lower compensation packages, however demonstrate higher job satisfaction levels, as well as longer
employment duration. Finally, Apple or EMC, have low job satisfaction scores, high compensation packages and demonstrate average employment durations.

We further explored the employment periods utilizing another methodology. Our findings are demonstrated in Figure 2 below. In Figure 2 we present a comparison of eight company’s employment (work) duration histograms: Facebook, Google, eBay, IBM, Apple, 3M, Intel and Motorola (ordered by descending average employment periods). The sample size and mean are presented in the label and are marked by a horizontal line. Note that there are peaks in the histograms in 1, 2, 3,…, k full years, indicating a tendency of employees to round up employment durations to the next full year period (i.e. 12 months).

[INSERT FIGURE 2 ABOUT HERE]

Examining the findings presented in Figure 2 above, Facebook scored the lowest reported average employment duration of 16.9 months. Facebook is followed by eBay, Apple, 3M, Intel and Motorola, with average employment durations of 26, 28.7, 38.2, 52.6 and 55.9 months, respectively.

The number of CV’s used to generate and analyze the histograms presented in Figure 2 above ranges from ~1000 CV’s to ~7000 CV’s. This reflects a large and representative sample. The study results also indicate a peak in a 12/24/36 months periods, which reflects a tendency to round up a short or partial employment duration to the near full year period, both in practice by the job applicant, or for CV writing purposes only.

Our study results also indicate, as presented in Figure 2 above, that the longer the average employment period, the less the data represented in the histograms
resembles an exponential distribution, hence forming a more long-tailed Gaussian shape. For example, while the upper two figures, for Facebook, eBay and Google fully resemble an exponential decaying function, the lower two histograms, of Intel and Motorola, have a peak at an employment period of ~24 month, and smaller bins in lower and in higher values. Assuming that an entry-level job in such companies usually requires a minimum of six months up to one year in order to reach mastery level of performance, this unexpected pattern seem more natural than those of Facebook or Google, where the average employment period is 16.9 and 23.3 months, respectively.

As illustrated above in Figures 1 and 2, study results indicate that despite higher job satisfaction scores and higher compensation packages, both Facebook and Google demonstrate shorter employment periods. This counterintuitive finding may be explained by employee’s perceptions of these companies. Some employees may indicate the companies “Google” and “Facebook” on their curriculum vitae profiles as a demonstration for past career excellence. This explanation is also supported by the demonstrated high peaks within reported year cycle, as demonstrated in Figure 2. Examining Figure 2 findings, in the upper two figures (i.e. Facebook and eBay) the employment period histograms reveal a sharp peak at end of each year. This indicates that employees who worked for 1 to 11 months in Facebook, may indicate a 12 month (one year) period on their curriculum vitae profiles. Knowledge intensive firms operating in the global media market – e.g. Google, Facebook and eBay - are known to be technological pioneers, therefore might be perceived by possible job applicants as “jumping platforms” towards their next desired job.

In order to deepen our understanding of the counterintuitive findings illustrated above, we analyzed career path directions and patterns of employees. Figure 3a below illustrates “career network clusters”, i.e., employees’ career path movements across
industries. Examining Figure 3a below, three main career network clusters can be identified: the financial cluster (represented in red color), the consulting cluster (represented in green color), and the high technology cluster (represented in blue color). This cluster analysis methodology enables us to examine employment and career moves within and across the career network cluster, and thus reveal surprising patterns of employee’s career path.

[INSERT FIGURE 3a ABOUT HERE]

As illustrated in Figure 3a above, employees tend to make more frequent career moves and changes within their own career network cluster. For example, employee working in the financial cluster (represented in red color), will tend to make career moves to other financial companies, but less frequently to other career network clusters. Similar career path movement patterns exist in the consulting cluster (represented in green color), and the high technology cluster (represented in blue color). Moreover, within the consulting cluster, we identified IBM company as an “industry hub”. A “industry hub” is a company that serves as a central crossroad junction for employees from which they can easily transfer to a different career network cluster. This finding is surprising when compared to the Facebook and Google results. Both Facebook and Google’s positioning represent a less central point as potential employers, and thus do not serve as industry hubs.

Analyzing the results generated from Figure 3a, one can detect that working at IBM may serve as a strategic career bridge to other employment industries, compared to working in other companies. In these companies, employees are more likely to stay within the borders of the career network cluster, especially when performing career
choices. Further analysis of career paths within and across career network clusters, revealed additional finding and patterns, represented in Figure 3b below.

[INSERT FIGURE 3b ABOUT HERE]

Figure 3b above indicates that both Facebook and Google are companies that dominate two distinct employment clusters. Analyzing the employment clusters within a network that consists of almost 50 thousand career moves between large organizations, only one single career move was detected between these two media giants. This finding is surprising and counterintuitive due to some similarities in employment patterns of the two knowledge intensive firms.

Earlier we indicated that according to the research findings, the average employment period in IBM is 42.6 months. It is much longer compared to the average employment period in Google and Facebook (16 and 25 months, respectively). As demonstrated in Figure 3b above, IBM is centrally located within its employment cluster, based on its physical central location, as well as on its evident large size. This finding represents IBM is a central company within its employment cluster. Moreover, it is an industry hub, from which employees can easily make career moves to other industries. This is surprising, also considering the demonstrated short employment duration identified previously in Google and Facebook. According to our results, IBM is an example of a unique company, which facilitates ongoing strategic career moves for its employees, together with a prolonged employment duration.
v. Discussion

This study presents some counterintuitive results related to employee turnover and career moves in knowledge intensive firms. In this section we discuss some possible explanations for the findings and results presented above.

Employees of knowledge intensive firms in general, and more specifically in the high technology industry, are a valuable organizational asset and are in high demand. Therefore, these employees might be approached more actively by competitive companies. Specifically, with regards to Google and Facebook, these are both relatively young companies compared to the other companies that are included in our analysis, e.g., Intel and Motorola. In the latter, longer term employment periods affect both the average employment duration, as well as the employee average age.

Nonetheless, from the companies’ perspective, the employment duration anomaly, may be explained by the internship system that these firms are executing. Both Facebook and Google are well known for having a vibrant and effective system of summer internships geared at relatively fresh employees. Thus, employees included in such internship programs, document this on their online CV profiles, thus shifting employment duration as presented in Figure 2.

From the employee perspective, engaged in employment relations in highly competitive knowledge intensive and global firms, may sometimes result in rather high employee burnout level. This implies that there exists a difference between the reported level of employees’ satisfaction, and the actual ability of companies to retain their employees. Finally, from a macro workforce perspective, the relatively short period of time, (1-2 years as presented in Figure 2), reflects the potential future of employment patterns. Decades ago, life-long careers in a single organization was considered a common average career. However, nowadays, a 3-4 years position (most specifically
in high technology work settings) is considered acceptable and in line with existing and predicted workforce movements.

Considering modern online digital career tools, shorter term employment cycles are sometimes perceived as a positive career management strategy, by organizations and by peer employees alike. Nonetheless, these global technology firms, which compete on talent on wide and global scales, can afford having shorter employment duration, since their worldwide global presence, as well as their reputation, transform the drawbacks employment distances into positive resources. As a result, these companies are not limited to recruiting their talents within a limited geographical area, but rather may leverage and better acquire their human resources by performing a truly global talent search.

One proposed explanation we offer that might explain the short average employment durations for employees in top high technology global media firms such as Google or Facebook, is that these companies serve as “jumping platforms” for employees to enter other competitive knowledge intensive firms. Thus, employees might prefer working in Facebook or Google, simply in order to move in the career path to their next job, which can be reached, only by demonstrating previous employment experience within these specific companies. Accepting this explanation, the shorter demonstrated employment period is initiated by the employee (i.e. voluntary turnover), and not by the company itself.

Finally, in order to validate this explanation, we constructed the industry network clusters, from the curriculum vita profiles data set. In the industry network cluster companies are represented by nodes, while employees moving across companies represent the edges between nodes (Figure 3a). Surprisingly, our findings revealed counterintuitive results. We found that, while Google and Facebook are central in the
employment network cluster, and thus many employees pass through them during their career journey. However, other companies (IBM) serve as industry hubs representing a much more central node.

Additionally we found that both Facebook and Google, two large knowledge intensive global media firms, seem to dominate two distinct employment clusters (Figure 3b). Thus, a distinct employment separation exists, even more so two employment “fire walls”. This is due to the fact that a surprisingly small number of employees make career moves between these two companies. Our study reveals that out of over 130,000 career moves in the data set, of which 50,000 were between large organizations, only one single employee made a career move between those two giants. Moreover, analyzing career moves represents IBM in a central position within the employment cluster. IBM was found to be located in the center of the employment network cluster. This finding reflects IBM’s flexibility as an employer. Moreover, employees who moved from IBM, continued on to diverse industry clusters, as opposed to employees from Google or Facebook, who tend to remain within in the same industry cluster.

Our findings may be explained by the Job Embeddedness Theory. The Embeddedness Theory completes an earlier theory which describes the reasons for staying in a job as “a net or a web in which an individual can become stuck” [21]. Thus, the decision to stay in a job, includes many emotional dependent and independent variables which influence the decisions of employees. The Embeddedness Theory enables a more robust and holistic view of vocational turnover processes, through the separation of factors responsible for leaving a job, and factors responsible for staying in it, thus serves us in explaining our study results.
An important aspect that requires careful consideration is the reliability of companies ranking. There are numerus company ranking methods available on the web, each having its own system of “Best 100 companies to work in”. While the reliability of the rankers is rather high, since the ranks are published by known and respectable media sources (e.g. Fortune 100, Glassdoor, and Business Insider), they often rely on the outcome of a commercial firm, which might gain from adding a specific company into the list. Furthermore, branding a company as an extraordinary employer, has a direct influence on the firm’s ability to recruit top employees, thus companies have a clear motivation for improving their ranks on these surveys.

Furthermore, pay and compensation levels, as collected from PayScale, are based upon a rather more stable study, and is thus more reliable to the best of our judgment. While the different job satisfaction scores vary between the different sources, Facebook and Google are always within the top 10 best places to work in lists and therefore their top location is unquestionable. Interestingly, these two companies are the ones with the shortest average employment duration, which is a surprising and somewhat counterintuitive result.

In light of all of the analysis illustrated above, our study emphasizes the importance of a nuanced perspective in the emergent field of career analytics. We demonstrate that by utilizing data-driven tools, one can gain a strategic granular view on employees’ career path initiatives.

VI. Conclusion and Contribution

As human resources and management challenges in many organizations become increasingly resource consuming and complex, the need arises for a robust tool to increase organizational effectiveness through the investigation of career analytics.
This study is an example of how an innovative and robust methodological approach based on data-driven and career analytics tools deepens our understanding of workforce movement and organizational phenomena, thus increasing ROI. Perhaps the most striking finding of our research is that in contrast to what was previously assumed, career planning is a non-linear and disruptive phenomenon in certain cases.

In this study we collected data extracted from approximately 970K curriculum vita profiles. Our purpose was to investigate an important HR related issue - career paths - in knowledge intensive companies - all indexed in the “Fortune 500 Best Companies to Work for in 2018”.

We found counterintuitive patterns in turnover and career moves which are not necessarily aligned with existing theories from the literature. Moreover, we uncover some disruptive aspects of the turnover phenomenon which are sometimes unintuitive. Another striking finding is an inverse relation between job satisfaction level, average employment period and job seeking patterns in specific high technology companies. We also found that while higher compensation packages often lead to higher job satisfaction, they do not translate into longer employment periods for some employees. This counterintuitive effect was found to be strongest in global high technology companies, such as Google and Facebook.

Specific antecedents may explain these counterintuitive results, e.g., workforce competitiveness, employee burnout, organizational recruitment policies and procedures, workforce mobility or employee’s perceptions regarding the ideal career path, these notable results require further investigation and research.

Our study presents three notable contributions. First, from a theoretical perspective, our findings advance the motivation literature by highlighting the counterintuitive effect of intrinsic and extrinsic motivation. Second, from a practical perspective...
perspective, our findings contribute the ongoing extant examination of career analytics which are at the forefront of the investigation of the organizations of the future. Moreover, our findings lay the groundwork for the adoption of practical career planning management tools, in the form of career analytics, which are at the forefront of the investigation of the organizations of the future. Finally, from a methodological perspective, this study contributes to the increasingly growing use of data-driven tools for organizational and human resources investigation [22].
References


Figure 1

Job Satisfaction, Average Work Duration and Salary

Based on 973K LinkedIn profiles from 39 Companies
Figure 2

Histograms showing work durations for different companies:
- Facebook: n=5661, Mean = 16.691
- eBay: n=1350, Mean = 26.044
- Google: n=21882, Mean = 25.597
- IBM: n=33602, Mean = 42.016
- Apple: n=7603, Mean = 26.7
- 3M: n=1836, Mean = 38.233
- Intel: n=7162, Mean = 52.686
- Motorola Solutions: n=2383, Mean = 55.926