



THE UNIVERSITY OF
SYDNEY

The Impact of Job Satisfaction on Employee Retention and Firm Performance in the IT Industry

MATTHIAS GIANNESINI
SID: 490390106

Supervisor: Dr. Mostafa Shahin

This thesis is submitted in partial fulfilment
of the requirements for the degree of
Bachelor of Advanced Computing

School of Computer Science
The University of Sydney
Australia

1 November 2024

Student Plagiarism: Compliance Statement

I certify that:

I have read and understood the University of Sydney Student Plagiarism: Coursework Policy and Procedure;

I understand that failure to comply with the Student Plagiarism: Academic Board Policy: Academic Dishonesty and Plagiarism can lead to the University commencing proceedings against me for potential student misconduct under the [2012 Academic Dishonesty and Plagiarism in Coursework Policy](#).

This Work is substantially my own, and to the extent that any part of this Work is not my own I have indicated that it is not my own by Acknowledging the Source of that part or those parts of the Work.

Name: Matthias Giannesini

Signature:



Date: 01/11/2024

Abstract

Job satisfaction has become a staple of the modern IT industry. Although tech companies show strong financial performance, they continue to experience high employee turnover despite extensive investments in employee benefits. This study seeks to address this contradiction by quantitatively analysing 7.4 million job reviews from Glassdoor, alongside financial data from Compustat. Feature ratings of job satisfaction are extracted from the reviews, while market capitalisation and return on assets (ROA) are chosen as measures of firm performance. Correlation and regression analyses are then performed using the individual features of job satisfaction, retention, market cap and ROA. The results indicate that job satisfaction is positively correlated with retention, while having minimal to non-existent predictive power in forecasting firm performance. These findings suggest that tech firms should continue to spend on employees with the intent of improving retention, but should take care with the expectation that these investments will translate to quantifiable firm performance.

Acknowledgements

I would like to thank my thesis supervisor Dr. Mostafa Shahin for guiding me through this year-long journey. Your ability to understand complex ideas in discussion and provide immediate, applicable advice has been invaluable.

Contents

Student Plagiarism: Compliance Statement	ii
Abstract	iii
Acknowledgements	iv
List of Figures	vi
List of Tables	vii
1. Introduction	1
2. Literature Review	3
3. Methodology	6
3.1 Glassdoor Data	6
3.2 Firm Performance Data	8
3.3 Merging Datasets	9
3.4 Summary Statistics	9
3.5 Correlation and Regression Analysis	11
4. Results	12
4.1 Correlation Analysis	12
4.2 Regression Models	15
5. Limitations	16
6. Conclusion	17
7. Future Work	18
Reference List	19
Appendix A	23

List of Figures

Figure 1: Methodology framework	6
Figure 2: Statistical summary of ratings features	9
Figure 3: Quarterly aggregated rating and retention across all firms	10
Figure 4: Quarterly aggregated rating, market cap and return on assets across all firms	10

List of Tables

Table 1: Summary of the Glassdoor review attributes	7
Table 2: Categorical feature encodings	7
Table 3: Independent variables used for correlation and regression analysis	12
Table 4: Correlation between features of job satisfaction and retention	12
Table 5: Correlation between features of job satisfaction and market capitalisation	13
Table 6: Correlation between features of job satisfaction and return on assets	14
Table 7: Performance of regression models with each dependent variable	15

1. Introduction

Tech companies are fundamentally transforming the modern workplace. With employee benefits such as gyms, cafes, and massages, the boundaries between personal and professional life are disappearing. For example, Apple's main office is complete with a state-of-the-art fitness centre, wellness programs, and nature trails for relaxation all in the name of boosting job satisfaction. Meanwhile Google offers perks like nap pods, gourmet cafeterias and even laundry services. These investments reflect a big shift in tech culture, raising an important question: are all these incentives truly effective? To investigate this issue, this thesis focuses on three questions related to job satisfaction in the tech industry. Firstly, does employee satisfaction actually lead to increased retention? What about firm financial performance? And if so, which components of job satisfaction have the biggest impact on retention and performance?

With the introduction of remote work arrangements after the COVID-19 pandemic, work-life balance has become an increasing priority for tech employees. This points to a greater focus on job satisfaction overall, considered to be one of the most desired attributes of any job. Yet, despite significant efforts by companies to increase job satisfaction through competitive salaries, flexible work environments, and positive corporate cultures, the information technology (IT) industry continues to experience high turnover. In particular, a 2023 survey found that 32% of tech employees were likely to change jobs in the next 12 months, notably higher than the 26% average across other industries, despite experiencing an 8% higher job satisfaction than the average employee (PWC, 2023). This situation presents a unique problem: despite high job satisfaction ratings, tech companies do not show a corresponding increase in retention.

The stakes of this problem are high, affecting not only businesses but human resources (HR) professionals and employees themselves. HR must spend greater amounts of time on recruitment and training, while employees experience more job instability and friction between job changes. While this problem remains unaddressed, tech companies face extensive costs by having to spend millions both on recruiting more talent as well as enhancing existing job benefits, leading to reduced capital to spend on other ventures.

To add to this, the relationship between job satisfaction and firm performance in the tech industry is just as ambiguous. Although studies generally indicate that job satisfaction is associated with firm performance (Melián-González et al., 2015), there is a notable gap in the literature when examining this connection specific to the IT sector. As such, the problem becomes further extended – within the IT industry, the impact of job satisfaction on both employee retention and firm performance is not very well understood. Failure to understand this relationship may lead to a misallocation of resources by tech companies spending excessively on employee benefits, in expectation of accordingly improved firm performance.

To better understand the nature of job satisfaction, online company review sites such as Glassdoor and Indeed provide a new source of untapped data: job reviews. In recent years, the number of job reviews has increased considerably with Glassdoor alone publishing 150 million reviews. Studies also show that 86% of employees and job seekers check company reviews before deciding on where to apply for a job, including

current employees considering a job change (Glassdoor, 2021). Job seekers look for work-life balance, salary, benefits and a positive company culture. In other words, facets of job satisfaction.

As such, review sites have become a recent and practical source of employee satisfaction data. Other sources include surveys, which provide more targeted research questions but can often be subject to social desirability bias, given that the responses are initiated by the surveyors. On the other hand, job reviews have the advantage of self-initiation, increasing the chance of honest and authentic opinion. Furthermore, sites like Glassdoor only allow anonymous reviews, further reducing the impact of biases.

Hence, this study analyses over 9.9 million Glassdoor reviews on 35,000 companies to examine job satisfaction and retention in the tech industry in detail. Firm performance data is retrieved from Wharton's CRSP/Compustat merged database and analysed in combination with the review data. Regression models are trained on the combined dataset alongside correlation analysis to understand the strength and nature of this relationship. By analysing the patterns in these datasets, this study aims to create a quantifiable relationship to better tackle the problem of how job satisfaction, retention and firm performance interact in the IT industry.

2. Literature Review

Existing literature on job satisfaction, retention and firm financial performance is abundant. Individually, the topics are fairly well researched but studies on the relationships between any two of the three are far less common. In regards to these relationships, much of the literature avoids focusing on any specific industry regardless of the IT industry itself, which often leads to discussions that are too generalised and unactionable.

Job Satisfaction

On its own, job satisfaction is a well-researched topic in literature. Clark (2015) delves into what constitutes a “good job”, finding that job satisfaction extends beyond financial remuneration to include factors like job security, future prospects, intrinsic interest of the work, interpersonal relationships, and the quality of working conditions. Judge & Klinger (2008) find a strong link between job and life satisfaction, concluding that satisfied employees are more likely to want to stay at their jobs. These findings are obtained largely from survey sources, which may be subject to some survey bias. Meanwhile, Moro et al. (2021) examine the drivers of job satisfaction in IT companies, identifying several key factors including salary, work-life balance, career advancement opportunities, and organizational culture. These factors are extracted from Glassdoor reviews, using natural language processing to separate meaning from the textual data.

Increasingly, job satisfaction is becoming a core part of the modern office worker’s priorities. Firms are correspondingly spending more on employee benefits, compensation and perks. A literature review found that firms now see employee satisfaction as central to productivity and have increased benefits to retain talent (Dulebohn et al., 2009). Hammermann & Mohnen (2014) analyzed tangible incentives and concluded that employee perks are highly effective in enhancing job satisfaction, particularly in large firms with the resources for extensive benefits. Meanwhile, studies on employee remuneration have shown that firms focusing on fair and generous compensation packages increase employee engagement and satisfaction (Magnan & Martin, 2019). In regards to the tech industry, research on firms from Silicon Valley discusses how tech companies lead the arms race of offering employee perks, with the goal of enhancing job satisfaction (Nguyen, 2015). Schneider (2016) confirms this, finding that the tech industry invests significantly more into benefits to boost employee morale compared to other industries. Undoubtedly, job satisfaction has become a staple of the modern IT industry.

Retention

Good research also exists on employee retention across all industries as a whole, but specific insights into the IT industry are lacking. It is generally accepted that job satisfaction is linked to retention (Harter et al., 2002; Hom et al., 2017; Allen et al., 2010). In surveys, employees consistently rate job satisfaction as one of the most significant reasons for staying at a company (Hausknecht et al., 2009). Across all industries, the relationship is clearly well documented. Industry-specific literature is more difficult to find, but also shows similar findings. Ghapanchi & Aurum (2011) reviewed studies on turnover in software development and found that job satisfaction was a good predictor of retention, given the competitive talent market in the tech industry. Rizwan et al. (2014) examined job satisfaction across multiple industries,

finding that satisfaction is a significant factor in employee retention, particularly in sectors with high turnover like IT. Harris (2018) examines retention strategies within IT companies, highlighting how facets of job satisfaction are leveraged to improve retention. Harris's research, centred around a case study in the Dallas-Fort Worth area, reveals that while many strategies are common across sectors, their effectiveness can vary significantly based on industry-specific factors. However, the study's geographic and demographic limitations point to a need for further research into these ideas on a broader dataset. Stein et al. (2022) extend these ideas, examining the relationship between job satisfaction with the retention of IT professionals in public and private IT institutions. They found that the most important dimension of job satisfaction for retention was satisfaction with salary, rather than salary itself. However, their research still suffers from demographic limitations, being based on survey data on IT employees solely in Santa Catarina, Brazil.

It should be evident that the research on the link between job satisfaction and retention is plentiful. This thesis aims to extend this research in a single foundational way: using a different dataset. The data for this study comes from Glassdoor, a website where company employees can post reviews of their jobs and companies. Glassdoor reviews provide many benefits as a data source. They are inherently anonymous, which improves their authenticity and quality since employees can give true feedback without fear of reproach from their employers (Milliken & Morrison, 2003). Furthermore, they avoid the social desirability bias found in surveys given that they are self-initiated rather than a response to directed questions. Studies have shown that this leads to more accurate insights in relation to job satisfaction (Edwards & Cable, 2009; Podsakoff et al, 2003). On top of this, their reviews are made from all over the world and over many years, leading to a large and diversified sample size. Compared to surveys, they are unlikely to be biased by time-specific or location-specific events.

Firm Performance

With contemporary trends in business emphasizing the improvement of employee workplaces, it has become a highly researched topic. The consensus is divided on whether it translates to verifiable firm performance. Proponents claim that increasing job satisfaction will raise productivity levels and ultimately firm financial performance. Meanwhile, critics argue that the connection between employee satisfaction and overall firm performance is weak, suggesting that the investment in perks, compensation and benefits is only impactful on an individual level. Park and Rahmani's (2020) findings suggest that satisfied employees are more likely to contribute to innovative processes and outcomes, measured by the number of patents generated by the company. Another study finds that higher levels of employee satisfaction contribute positively to firm financial outcomes, as satisfied employees tend to be more engaged, productive, and committed to the organization (Melián-González et al., 2015). In particular, this study uses Glassdoor data which provides evidence on both public and private firms from a number of different industries. The keystone study comes from Edmans (2012), who demonstrates that firms with high employee satisfaction achieve superior stock performance, suggesting that satisfaction drives financial success. Edmans uses ideal research practices, measuring firm value while controlling for risk, firm characteristics, industry performance, and outliers. Van Rooy et al. (2011) link employee engagement with organizational performance, finding that engagement and satisfaction play a critical role in improving financial outcomes.

Although these studies are exhaustive, many of them use subjective measures of firm performance when investigating the impact of job satisfaction. This is because correlation between a qualifiable and quantifiable metric is difficult to measure. Realistically, job satisfaction represents emotions and opinions while firm performance is based on data-driven calculations. It is for this very reason that a majority of studies investigating the link between job satisfaction and some other dependent variable choose a subjective metric. For example, studies will often use surveys to determine job satisfaction, asking employees for their opinions on work-life balance, culture, growth opportunities and so on. Curiously, the surveys will often also ask those same employees what they think about their company's performance, and use their responses as the variable being measured for correlation (Pang & Lu, 2018; Deb et al., 2023; Shiu & Yu, 2010; Phi et al., 2021; Kanyurhi & Akonkwa, 2016). These studies unsurprisingly find positive correlation between job satisfaction and organisational performance. This seems like a remarkable design flaw, since employees who are satisfied with their job are likely to be biased in their opinion towards their company's financial performance, skewing the results (Galant & Cadez, 2017). Understandably, another reason behind this is that retrieving consistent financial data from firms is usually a difficult and costly venture.

Besides this, some studies that do use some form of objective measure of retention or firm performance as the dependent variable tend to use absolute values as opposed to relative values, which can be a design flaw if not accounted for (Wruck & Yu, 2021). For example, Apple has a significantly larger market cap than a small tech startup, but the average employee satisfaction between the two companies may not be so different. Simply put, larger companies are far more likely to retain employees better and have much stronger metrics of financial performance due to their incumbency in the market, as opposed to smaller startups which undergo a lot of change (Labro & Omartian, 2024). Controlling for size can mitigate some of this, such as classifying companies into different size classes, however inter-class variation may still skew results (Edmans et al., 2023).

As such, the general connection between job satisfaction, employee retention and firm performance is fairly well researched. However, a gap in the literature is highlighted when researching these topics specifically in relation to the IT industry. Furthermore, there is very little research that uses new and relevant data sources such as job review sites. Therefore, this thesis aims to bridge the gap in the literature surrounding the correlation between these three variables. With much of the existing literature being published in the past 10 years, the research is fairly recent due to the increasing availability of job review data. Job review data from Glassdoor and financial performance metrics from Wharton's CRSP/Compustat merged database will be combined and analysed, focusing exclusively on the IT industry.

3. Methodology

Job satisfaction and employee retention data was collected from Glassdoor reviews, while firm financials were collected from Wharton Research Data Services (Glassdoor, 2024; Wharton, 2024). Glassdoor data was scraped from their website and cleaned, followed by feature extraction. To fetch the financial data, the names of the firms and their tickers were required first. A list containing the top 500 IT firms by market capitalisation was retrieved from Yahoo Finance to this end (Yahoo, 2024). The list served a twofold purpose. Firstly, it restricted the scope of the firms to only the IT industry, which was needed for the study. Second, only publicly traded companies were analysed for this study, since financial performance data is significantly more costly and challenging to obtain for private companies. Furthermore, the bulk of reviews on Glassdoor are made on public rather than private companies.

After retrieval from Wharton, the financial data was cleaned, converted, and merged with the review data by firm name. Correlation and regression analysis was performed on the final dataset to create models of the relationship between job satisfaction, employee retention and firm performance.

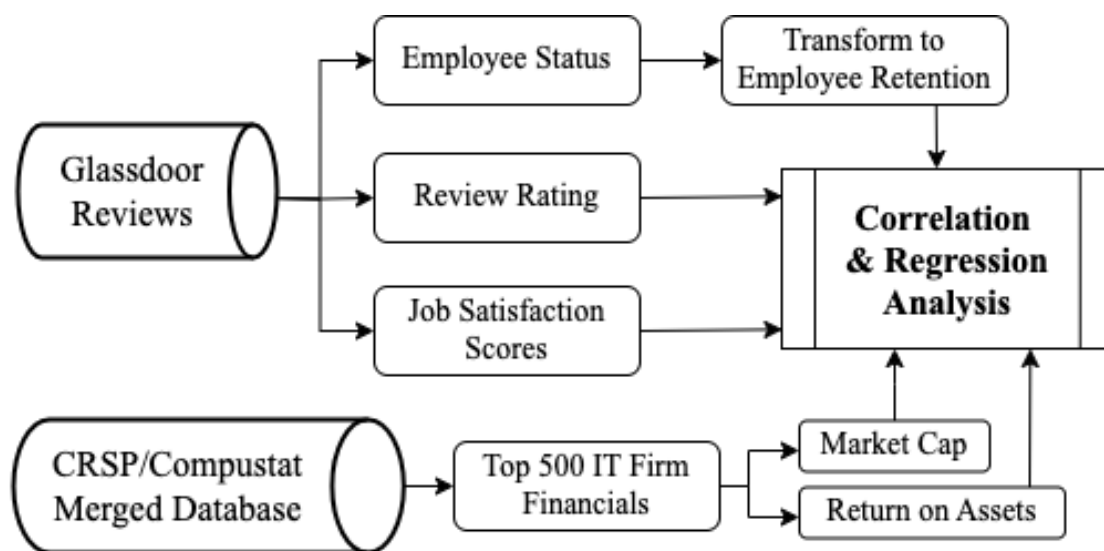


Figure 1: Methodology framework

3.1 Glassdoor Data

The review data for this study was collected solely from Glassdoor, primarily due to the sheer volume of reviews on their site compared to other review sites such as Indeed. Glassdoor was also chosen because their reviews are uniquely anonymous, largely reducing the impact of positivity bias from named employees. Furthermore, sourcing the data from an online review site ensured that reviews were not restricted by location, reducing geographical biases common in other studies of job satisfaction.

Data Collection

To retrieve the review data, a web scraping program was developed in Python and run over multiple weeks on Glassdoor.com. The resulting dataset contained over 9.9 million job reviews on 35,000 companies with dates ranging from 2012 through to 2023. The features of the dataset are outlined in Table 1 below.

Feature	Description
rating	Overall rating of the job on a scale of 1-5
title	Review title
status	Employee status at time of review
pros	Pros of working at the firm
cons	Cons of working at the firm
firm_link	URL of firm's reviews
date	Date of review published to Glassdoor
job	Job being reviewed
recommend	Whether the employee would recommend this job to a friend
ceo_approv	Whether the employee approves of the CEO
bus_outlook	What the employee thinks of the business outlook
career_opps	Rating of the career opportunities on a scale of 1-5
comp_benefits	Rating of the compensation and benefits on a scale of 1-5
wl_balance	Rating of the work-life balance on a scale of 1-5
cult_values	Rating of the culture and values on a scale of 1-5
senior_mgmt	Rating of senior management on a scale of 1-5

Table 1: Summary of the Glassdoor review attributes

Data cleaning and transformation was performed in Python with the use of the *pandas* and *NumPy* libraries. Of the almost 10 million reviews, 2.2 million of them contained null data and were dropped, leaving a final complete dataset of 7.4 million reviews. This decision was justified by the immense size of the dataset, meaning that integrity and feature patterns would not be lost by heavy reductions in size while the completeness of the resulting dataset would be far easier to work with.

Data Processing

Glassdoor's site structure creates difficulties when associating the firm name to the review, so the page URL was instead retrieved with each review. The firm name could then be extracted from the URL using regular expressions and parsing. Care had to be taken with certain accents and symbols being represented with special characters (e.g. The "é" in "Université" is represented as "%C3%A9", referring to UTF-8 encoding "c3 a9").

The *recommend*, *ceo_approv* and *bus_outlook* features had four possible string values, which were converted to match the other numerical feature scales from 1-5:

String	"Positive"	"Mild"	"Negative"
Conversion	5	3	1

Table 2: Categorical feature encodings

The employee *status* detailed whether the review was coming from a current or former employee, as well as how long they had been a current employee. This was converted to a Boolean, where True represented the current employees (“retention”) and False represented the former employees (“turnover”).

Thus, the resulting dataset had 9 features of job satisfaction to be used as independent variables, and *retention* to be used as a dependent variable in later regression models.

3.2 Firm Performance Data

Initially, Yahoo Finance was chosen as the source for firm performance data. However, attempts at retrieval revealed that their database is limited to financial metrics from the last 4 years only (i.e. from 2020 onwards). This was not sufficient for the study given that the earliest reviews were dated from 2012, meaning that a significant portion of the review data could not be used for analysis. As such, firm performance data was instead collected from Wharton’s CRSP/Compustat merged database. Quarterly metrics were retrieved for each firm in the list of top public IT firms from the period 2012-2023.

Measures of Performance

Firm performance is an extensively researched topic in the fields of finance and accounting, and the ideal method of measuring it is heavily debated (Cheboi & Keter, 2023; Herciu, 2017; Siepel & Dejardin, 2020). For this study, market capitalisation (market cap) and return on assets (ROA) were chosen as metrics.

Market cap measures the total value and size of a company. Although it is a good general indicator of firm performance, it is strongly influenced by market fluctuations (Muhammad & Scrimgeour, 2014). Meanwhile, ROA measures the effectiveness of an organisation to generate profits based on the dollar value of its assets, meaning both cash-based assets as well as equipment, property and inventory. It is far less exposed to market changes while providing a good assessment of both the financial and non-financial performance of a company (Al-Matari et al., 2014). This makes sense for the study, given that the output of productive and happy employees is not always tangible.

Data Collection

Market cap and ROA were not immediately available from the CRSP/Compustat database, and were instead calculated from their formulas:

$$\text{Market Cap} = \text{Share Price} \times \text{Number of Shares Outstanding}$$

$$\text{Return on Assets} = \frac{\text{Net Income}}{\text{Total Assets}}$$

Following retrieval, the market cap and ROA were converted to percentage change values. The conversion from an absolute to a relative measure of firm performance is important, since larger firms with bigger market caps tend to allocate more to employee benefits (Oyer, 2004). Both measures were also 1% winsorized, meaning that outlier values falling outside the 1-99% range of distribution were capped.

3.3 Merging Datasets

Given that the reviews were linked to a date while the financials were linked to a quarter, care had to be taken when merging. One option was to upscale the firm performance data, merging on a review-by-review basis by repeating the same market cap and ROA value for all reviews made for a company during a given quarter. However, this would have created too much repeated data, obfuscating any patterns that might be revealed by the regression analysis.

Instead, the review data was downsampled by averaging all the review data by both firm and quarter, in order to match up with the quarterly financials. For the numerical ratings (e.g. *career_opportunities*) this was a simple task, while the *retention* feature was handled by calculating the ratio of the retention group to the turnover group:

$$\text{Quarterly Retention} = \frac{\text{Retention}}{\text{Retention} + \text{Turnover}}$$

With the time periods matching, the datasets were merged on firm name and quarter. The firm names did not always match exactly, with 211 initial matches on equality. Through the use of fuzzy matching, which identifies similar but not identical strings (e.g. “Apple” and “Apple, Inc.”), another 21 firms were merged .

Following the merge, the final dataset contained 7,203 rows of quarterly data on 232 public IT companies.

3.4 Summary Statistics

Figure 2 shows a summary of the numerical rating features of job satisfaction across all jobs in the IT industry. Senior management is rated the lowest on average but has the highest standard deviation, indicating more polarised employee opinions.

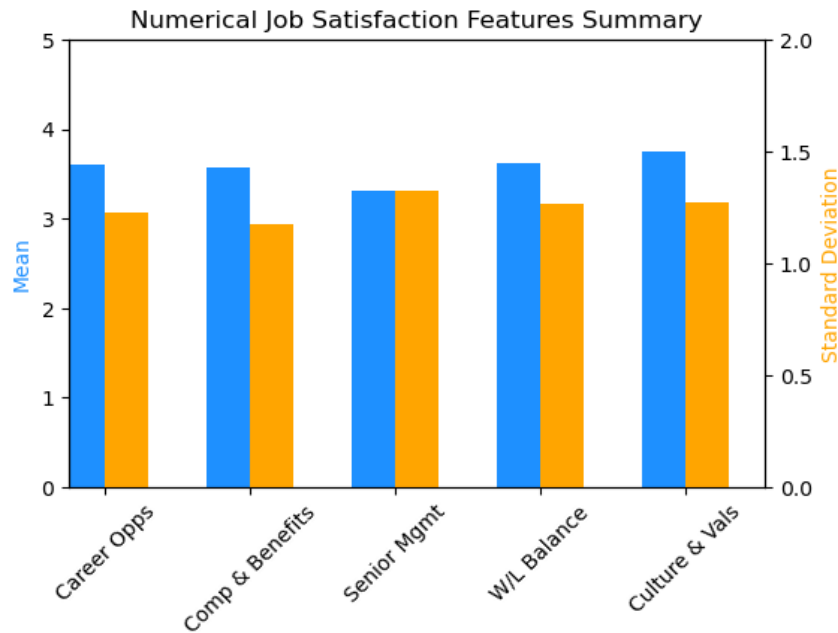


Figure 2: Statistical summary of rating features

Figures 3 and 4 display the overall review rating, retention, market cap and return on assets averaged across all companies per quarter.

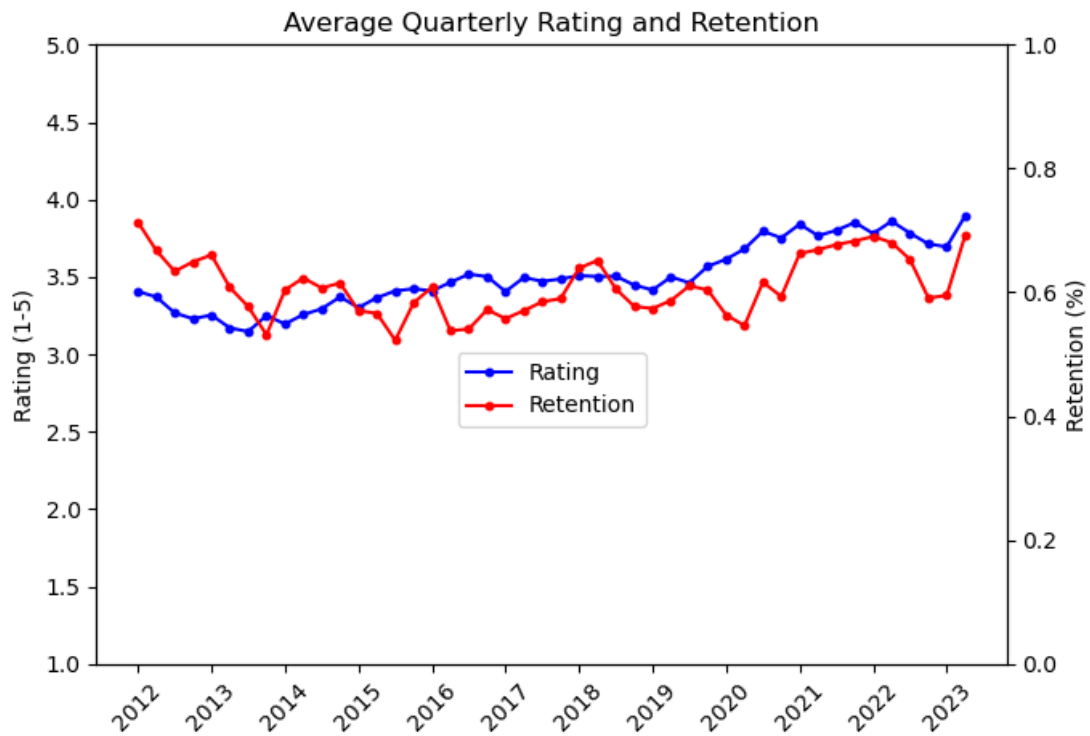


Figure 3: Quarterly aggregated rating and retention across all firms

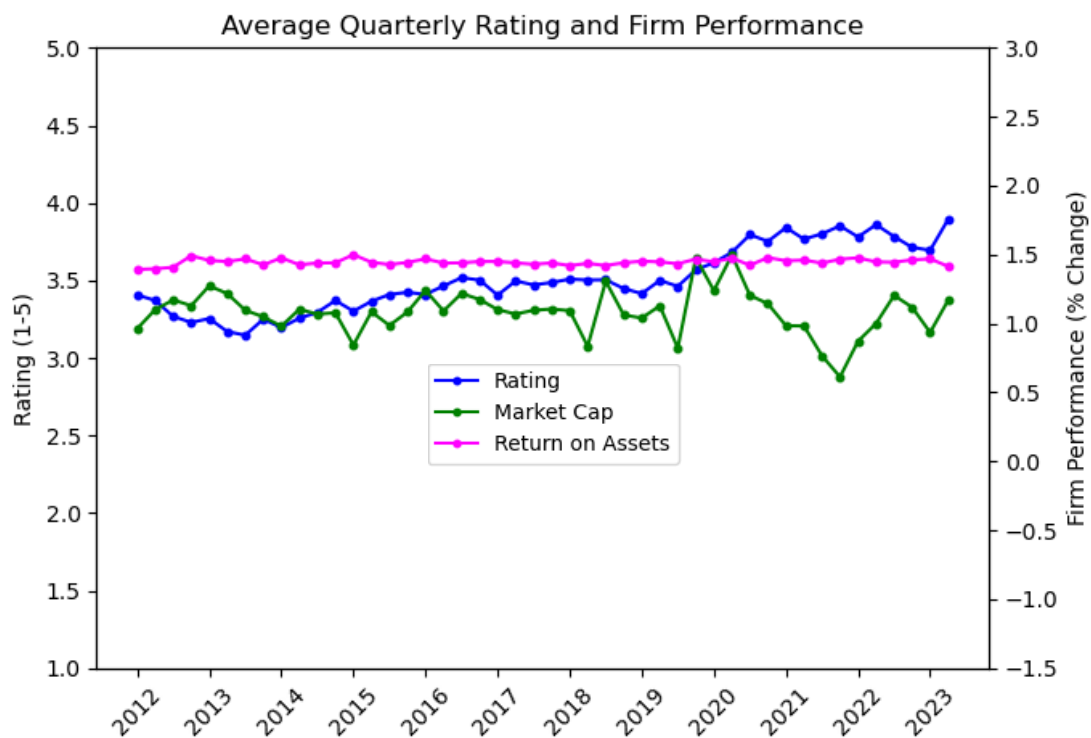


Figure 4: Quarterly aggregated rating, market cap and return on assets across all firms

3.5 Correlation and Regression Analysis

To measure the relationship between job satisfaction, retention and firm performance, this study focused on two methods of analysis: correlation and regression models. The Pearson correlation coefficient measures the strength and direction of a linear relationship between two variables. In this study, it was used to measure the relationship between each individual feature of job satisfaction with retention and firm performance. Its simplicity is easy to use and interpret, however, more complex analysis was needed to capture non-linear patterns in the data.

As such, three regression models were chosen to improve the robustness of the study: Linear Regression, Random Forest and Support Vector Machine (SVM). Regression models seek to model a predictive relationship between dependent and independent variables by finding patterns in the data. Each of them fall under the umbrella of supervised learning, meaning that they are trained on labelled data to make predictions about unseen data.

Machine Learning Models

Linear Regression is an algorithm that fits a linear equation to a dependent variable and one or more independent variables. It is the most basic regression of the three and may lead to underfitting when the patterns in the data are too complex.

Meanwhile, Random Forest is a machine learning method that constructs a number of decision trees during training. Given its ability to capture both linear and non-linear relationships, it tends to be robust to both under- and overfitting.

SVM in this study is built as a Support Vector Regressor, aiming to predict a continuous output by finding a hyperplane that best fits the data. It typically finds the most complex patterns of the three by using a high-dimensional space, but may lead to overfitting (Salam et al., 2021).

Hyperparameter Tuning

Extensive hyperparameter tuning was conducted for all three models. For Random Forest, this included tuning the maximum depth, minimum samples per split, and maximum number of features. For SVM, this included testing various kernels such as linear, polynomial, and radial basis function to search for hidden trends in the job satisfaction data.

Independent Variables

These regression models were used to determine the predictive power of all the features of job satisfaction in estimating the retention, market cap and ROA of companies in the tech industry. Table 3 displays a summary of the independent variables used for the models, noting that each variable is in quarterly average form.

Variable	Description
rating	Overall rating of the job on a scale of 1-5
recommend	Rating of their job recommendation on a scale of 1-5
ceo_approv	Rating of their approval of the CEO on a scale of 1-5
bus_outlook	Rating of the business outlook on a scale of 1-5
career_opps	Rating of the career opportunities on a scale of 1-5
comp_benefits	Rating of the compensation and benefits on a scale of 1-5
wl_balance	Rating of the work-life balance on a scale of 1-5
cult_values	Rating of the culture and values on a scale of 1-5
senior_mgmt	Rating of senior management on a scale of 1-5

Table 3: Independent variables used for correlation and regression analysis

4. Results

4.1 Correlation Analysis

Correlation with Retention

To assess the influence of job satisfaction on employee retention, the Pearson correlation coefficient is calculated between each factor of job satisfaction and retention. Table 4 summarizes the coefficients along with their associated p-values. Given the relatively large size of the dataset at 7,203 samples, a significance level of 0.01 was chosen to account for the smaller p-values.

Feature	Correlation	P-value
rating	0.332	0.000
recommend	0.327	0.000
ceo_approv	0.329	0.000
bus_outlook	0.321	0.000
career_opps	0.315	0.000
comp_benefits	0.191	0.000
wl_balance	0.244	0.000
cult_values	0.312	0.000
senior_mgmt	0.330	0.000

Table 4: Correlation between features of job satisfaction and retention

The results generally show that there is a positive correlation between factors of job satisfaction and employee retention in the IT industry. While none of the correlations are particularly strong (i.e. close to 1), they still suggest moderately strong relationships given the large sample size of the data. The uniformly low p-values support this, with all values passing the level of statistical significance.

The highest correlations are observed for rating, senior management and CEO approval. These findings propose that how employees rate their overall job experience, alongside the effectiveness of senior management and their approval of the CEO are all closely related to whether they stay at the company. In practicality, this reveals the significant role of leadership in retention for tech companies. Employees who feel positively about their leadership are more likely to stay with their company, reducing turnover.

On the other hand, the lowest correlation is observed for compensation and benefits, a surprising result. Compensation is often assumed to be a key driver of retention, yet this result suggests that other factors may be more important for employees in the IT industry. While pay is certainly a factor, things like management, culture, and career opportunities might have a greater impact on an employee's decision to stay. To add to this, work-life balance scores the second-lowest, despite often being seen as a vital part of retaining employees. Although important, it appears that leadership is the most critical part of maintaining retention in the tech industry.

Correlation with Market Capitalisation (Percentage Change)

The correlation coefficients between each factor of job satisfaction and market cap are summarised in Table 5, alongside their p-values. Note again that these statistics are calculated for the percentage change in market cap for the next quarter, as opposed to the firm's current market cap value.

Feature	Correlation	P-value
rating	-0.011	0.359
recommend	0.000	0.977
ceo_approv	0.011	0.371
bus_outlook	0.017	0.151
career_opps	-0.014	0.249
comp_benefits	0.001	0.911
wl_balance	-0.023	0.046
cult_values	-0.016	0.176
senior_mgmt	-0.005	0.685

Table 5: Correlation between features of job satisfaction and market capitalisation

The calculations reveal surprising results, with all correlations hovering around zero. This suggests that the measurable relationship between job satisfaction and market cap change is almost non-existent. The p-values further support this, with all features failing to reach the statistical significance threshold of 0.01. The individual feature correlations do not vary much between themselves either, preventing specific inferences from being drawn.

The only value close to significance is the coefficient between career opportunities and market cap change, yet it is negative. Besides that, business outlook has the strongest positive correlation with market cap despite being negligible. This at least suggests that the data does align in some way, although far less cohesively than expected.

Unlike the retention results, the correlations here suggest that employee satisfaction factors have almost no influence on the market cap changes of IT firms. While job satisfaction is important for internal factors like retention, it may not necessarily be tied to performance indicators such as market cap. For companies, this highlights the need to balance internal employee well-being with external market strategies, as the two may not always be strongly linked.

Correlation with Return on Assets (Percentage Change)

Table 6 displays the correlations between variables of job satisfaction and change in ROA. As previously, the value for ROA is the percentage change to the next quarter.

Feature	Correlation	P-value
rating	0.025	0.033
recommend	0.017	0.139
ceo_approv	0.009	0.445
bus_outlook	0.020	0.093
career_opps	0.031	0.008
comp_benefits	0.012	0.313
wl_balance	0.018	0.129
cult_values	0.028	0.017
senior_mgmt	0.041	0.001

Table 6: Correlation between features of job satisfaction and return on assets

The results show that most of the correlations between job satisfaction factors are small, indicating very weak positive relationships overall. However unlike retention and market cap, the p-values vary considerably with some correlations achieving statistical significance and others falling short.

Of all the features only career opportunities and senior management exceed the significance level of $p < 0.01$, but the correlation values themselves are not practically large. This suggests that employee opinions of their managers and careers makes a difference in driving firm performance in the IT industry. Interestingly though, leadership in general appears to be divided in influencing firm success, with the CEO approval correlation having the smallest value and significance.

In contrast, other features such as compensation and benefits and work-life balance show no meaningful correlation with ROA, implying that while these factors are important for retention they do not have a direct impact on firm performance. This is a notable result for tech firms, which may assume that the investments put into employees will be returned in measurable output. Given the importance of employee opinions of upper management and career potential in both retention and financial measures like ROA, tech companies should thus focus not only on keeping employees satisfied but also on improving middle management and maintaining internal growth opportunities.

4.2 Regression Models

The regression models in this section are represented by the equation below:

$$y = \beta_0 Rat + \beta_1 Rec + \beta_2 CEO + \beta_3 Bus + \beta_4 Opp + \beta_5 CnB + \beta_6 WLB + \beta_7 Cul + \beta_8 Mgm + \epsilon$$

Where y = average quarterly firm retention or market cap or return on assets,
 Rat = average quarterly employee job review rating,
 Rec = average quarterly employee job recommendation rating,
 CEO = average quarterly employee approval rating of CEO,
 Bus = average quarterly employee rating of business outlook,
 Opp = average quarterly employee rating of career opportunities,
 CnB = average quarterly employee rating of compensation & benefits,
 WLB = average quarterly employee rating of work/life balance,
 Cul = average quarterly employee rating of culture & values,
 Mgm = average quarterly employee rating of senior management.

Each model estimates the equation differently in order to predict company retention, market cap or return on assets. The main results are present below in Table 7. Accuracy is measured with r-squared and mean squared error (MSE), allowing for performance comparison between models.

y	Model	R^2	MSE
Retention	Linear Regression	0.106	0.047
Retention	Random Forest	0.127	0.046
Retention	Support Vector Machine	0.099	0.047
Market Cap	Linear Regression	0.002	0.201
Market Cap	Random Forest	0.004	0.201
Market Cap	Support Vector Machine	0.000	0.202
Return on Assets	Linear Regression	0.000	0.075
Return on Assets	Random Forest	0.001	0.075
Return on Assets	Support Vector Machine	0.000	0.075

Table 7: Performance of regression models with each dependent variable

Overall the regression models report similar results to the correlation statistics. The incredibly low R-squared values across all models indicate that the trained models can only predict a minimal portion of the variance in the dependent variables. Although surprising given the hyperparameter tuning undertaken, it is worth noting that social data is often much less precise than that of other fields such as computing and sciences.

The MSE shows very little variation between models, suggesting that they perform similarly regardless of the dependent variable. Similarly to the correlation results, the MSE when predicting ROA was slightly better than predicting market cap. This is not entirely supported by the R-squared values, but they are so small as to be inconsequential in analysis.

Comparing the models individually, Random Forest reliably outperformed Linear Regression and SVM for all variables, albeit with only slight improvements in performance. This can be attributed to Random Forest's ability to capture both linear and non-linear relationships, which could exist in the data but remain weak. Linear regression performed comparably, while SVM made consistently suboptimal predictions. The most probable cause is overfitting in combination with low correlation, given SVM's tendency to search for the most complex patterns of the three models.

Thus, the regression models confirm the results found in the earlier correlation analysis. Weak or non-existent relationships are observed between job satisfaction and market cap, as well as ROA. The low predictive power in these models suggests that employee satisfaction variables, while important for internal metrics like retention, are not strong indicators of financial performance in the IT industry. This reinforces the idea that factors outside of employee satisfaction such as innovation, resources and skill level may play a more dominant role in determining financial outcomes.

For employees and firms in the IT industry, these findings imply that while job satisfaction is crucial for retention, its direct influence on financial performance is minimal. Firms may need to look beyond employee satisfaction when aiming to drive market valuation or profitability. This suggests that retention strategies focused on job satisfaction could help reduce turnover but are unlikely to immediately improve firm financial metrics like market cap or ROA.

5. Limitations

Recognising the limitations of the dataset is necessary in light of the modest results. As a subjective topic of interest, job satisfaction is inherently difficult to measure. Glassdoor reviews provide many advantages as a data source, such as the anonymity of feedback, the diversity of samples internationally, the continuity of samples over many years, and self-initiated participation. Each of these contribute towards a reduction in bias, improving the dataset's value for analysis.

However, one criticism of the data comes from the polarized nature of reviews, since employees are more likely to post a review if they feel strongly about their job. Fortunately, according to a study of different types of online review sites, Glassdoor was one of the few with a normal distribution in its ratings (Schoenmueller et al., 2019). This is likely due to Glassdoor's access policies, where all users are required to contribute a review after accessing a certain number of pages on their website. After submission, reviews are processed and screened for accuracy by Glassdoor before being published. This broadens the range of reviewer demographics, ensuring a more normal distribution of reviews.

Another limitation comes in the fact that review data is very unspecific to the purpose of research when compared to data acquired from surveys, making precise real-world implications difficult. Survey data may also provide other benefits, capturing all degrees of polarity equally and allowing for much more targeted questions. However, surveys have difficulties dealing with biases of their own.

A host of difficulties arise from measuring the practical relationship between job satisfaction, retention and firm performance. Many come from the fact that job satisfaction is a qualifiable measure while the other two variables are quantifiable. Realistically, job satisfaction is a reflection of subjective emotions, opinions and cultural trends while firm performance is a reflection of objective numbers, ratios and calculations. Additionally, job satisfaction is an entirely individual measure while firm performance is determined by the output of the group. These divides make any sort of correlation difficult to identify, much less quantify. This is likely why the results failed to show much correlation, besides the presence of modest data patterns.

Regardless of the context of the IT industry, another limitation of this study is encountered in the time-lagged nature of job satisfaction's impact on retention and firm performance. The output of highly satisfied employees may only manifest itself in the overall firm performance metrics many months later (Kessler et al., 2020). Although a one-quarter time lag was briefly examined in this study (see Appendix A), the lack of correlation irrespective of time period discouraged further investigation.

Regarding the measures of firm performance, market cap and ROA are useful in considering a firm's equity value and asset value. However, they do not reflect company debt. For example, a highly leveraged firm may appear to perform well based on market cap and ROA, but the risk posed by heavy debt is not reflected in these measures.

6. Conclusion

In this study, the problem of high job satisfaction combined with low retention yet strong firm performance in tech companies was explored. The uniqueness of this conundrum to the IT industry provided good motivation to find out what was occurring in the data. A review of the literature indicated that the topics of job satisfaction, retention and firm performance were well researched individually, but few studies analysed the link between any of the three variables. Fewer still specialized to any one industry when exploring these topics regardless of the IT industry itself. Additionally, the literature that did cover these relationships often used puzzling data sources, such as using surveys as a measure of financial performance or failing to control for company size. To fill this gap, job satisfaction and retention statistics were pulled from Glassdoor while financial metrics for tech firms were retrieved from Compustat to determine the impact of job satisfaction on retention and firm performance in the IT industry.

First, correlation analysis was performed on the individual features of job satisfaction with retention. This was done by converting the retention data to Boolean values, then to a numerical format. The findings showed that job satisfaction has a modest, positive correlation with employee retention. In particular, variables associated with leadership had the greatest coefficients, while compensation, benefits and work-life balance scored lower. Although salary, perks and lifestyle are oft considered vital to retaining employees, these results suggest that leadership has a more compelling role in tech companies.

Next, correlations were calculated between job satisfaction and firm performance, using market cap and ROA as metrics. This was done by aggregating ratings by quarter in order to match up with the quarterly financial data, and converting the financials from absolute to relative values. The calculations for market cap returned surprising results, finding no statistically significant correlation with any of the features of job satisfaction. Although the limitations of time lag and differing data types should be taken into account, this implies that employee satisfaction has a minimal effect on market cap in the IT industry. Similarly, ROA generated weak correlations overall. Statistical significance was only reached with career opportunities and senior management, suggesting that employee perspectives on career prospects and managers have some influence on firm performance in tech companies, while other facets of job satisfaction have no effect.

Finally, regression analysis was performed on the factors of job satisfaction using retention, market cap and ROA each as a dependent variable. The machine learning models confirmed the results found in the correlations. Job satisfaction weakly predicted firm retention rates, but produced inconsequential or non-existent predictions for both market cap and ROA. This suggests that factors beyond employee satisfaction such as innovation, assets and expertise may be more critical to financial success in tech firms.

Overall, the significance of this study lies in the precise nature of the data and analysis methods used. The results indicate that while job satisfaction is important for retention, its direct impact on financial performance is minimal in the tech industry. Firms should take care with the expectation that spending on employees will translate to quantifiable firm performance, and instead focus not only on keeping employees satisfied but also improving other factors tied to financial success.

7. Future Work

Given that Glassdoor was the only source of job reviews, reproducing the study with data from other sites such as Indeed could help determine if the results are consistent and can be generalised to other datasets. Furthermore, adding financial measures such as the debt-to-equity ratio and beta could be used to cover the debt limitation. Besides these, with the modern business environment being shaped by a growing focus on corporate social responsibility, it would be intriguing to test the correlation of job satisfaction with non-financial metrics such as environmental impact, social factors, governance and productivity. Each of these additions could also be added to the regression models for further results.

Future studies could focus on longitudinal measures of firm performance which account for time lag, such as latent growth modelling. Correlation and regression analysis on the relationship between job satisfaction, retention and firm performance could also be performed between different industries to compare how the relationship evolves between sectors. Perhaps other industries would show a more prominent connection between the three variables.

Reference List

- Al-Matari, E., Al-Swidi, A., & Fadzil, F. (2014). The measurements of firm performance's dimensions. *Asian Journal of Finance and Accounting*, 6(1), 1-24. <https://doi.org/10.5296/ajfa.v6i1.4761>
- Allen, D. G., Bryant, P. C., & Vardaman, J. M. (2010). Retaining talent: Replacing misconceptions with evidence-based strategies. *Academy of Management Perspectives*, 24(2), 48–64. <https://doi.org/10.5465/AMP.2010.51827775>
- Cheboi, J., & Keter, C. (2023). Financial Performance and Firm Value of Listed Companies: Financial Performance Measure ROA versus ROE. *Journal of Business and Economic Research*.
- Clark, A. E. (2015). What makes a good job? Job quality and job satisfaction. *IZA World of Labor*. <https://doi.org/10.1787/670570634774>
- Cosa, M., Torelli, R. (2024). Digital Transformation and Flexible Performance Management: A Systematic Literature Review of the Evolution of Performance Measurement Systems. *Global Journal of Flexible Systems Management*, 25, 445–466. <https://doi.org/10.1007/s40171-024-00409-9>
- Deb, S., Nafi, S., Mallik, N., & Valeri, M. (2023). Mediating effect of emotional intelligence on the relationship between employee job satisfaction and firm performance of small business. *European Business Review*, 35(5), 624-651. <https://doi.org/10.1108/EBR-12-2022-0249>
- Dulebohn, J. H., Molloy, J. C., Pichler, S. M., & Murray, B. (2009). Employee benefits: Literature review and emerging issues. *Human Resource Management Review*, 19(2), 86-103. <https://doi.org/10.1016/j.hrmr.2008.10.001>
- Edmans, A. (2012). The Link Between Job Satisfaction and Firm Value, with Implications for Corporate Social Responsibility. *Academy of Management Perspectives* 26(4), 1-19. <https://doi.org/10.2139/ssrn.2054066>
- Edmans, A., Pu, D., Zhang, C., & Li, L. (2023). Employee Satisfaction, Labor Market Flexibility, and Stock Returns Around the World. *Management Science*, Finance Working Paper No. 433/2014. <https://doi.org/10.2139/ssrn.2461003>
- Edwards, J. R., & Cable, D. M. (2009). The value of value congruence. *Journal of Applied Psychology*, 94(3), 654–677. <https://doi.org/10.1037/a0014891>
- Feng, S. (2023). Job Satisfaction, Management Sentiment, and Financial Performance: Text Analysis with Job Reviews from Indeed.com. *International Journal of Information Management Data Insights*, 3(1), 100084. <https://doi.org/10.1016/j.jjime.2023.100084>
- Galant, A., & Cadez, S. (2017). Corporate social responsibility and financial performance relationship: a review of measurement approaches. *Economic Research-Ekonomska Istraživanja*, 30(1), 676–693. <https://doi.org/10.1080/1331677X.2017.1313122>
- Ghapanchi, A.H. and Aurum, A. (2011). Antecedents to IT Personnel's Intentions to Leave: A Systematic Literature Review. *Journal of Systems and Software*, 84, 238-249. <https://doi.org/10.1016/j.jss.2010.09.022>
- Glassdoor LLC. (2024). *Glassdoor Reviews*. <https://www.glassdoor.com.au/Reviews/index.htm>
- Glassdoor. (2021). Statistical Reference Guide for Recruiters. Retrieved 15 Sep 2024. <https://www.glassdoor.com/employers/resources/hr-and-recruiting-stats/#employer-branding-and-recruiting-hiring-and-retention>

- Hammermann, A., & Mohnen, A. (2014). Who benefits from benefits? Empirical research on tangible incentives. *Review of Managerial Science*, 8(3), 327–350. <https://doi.org/10.1007/s11846-013-0107-3>
- Harris, C. N. (2018). Employee Retention Strategies in the Information Technology Industry. *Walden Dissertations and Doctoral Studies*. 5559. <https://scholarworks.waldenu.edu/dissertations/5559>
- Harter, J. K., Schmidt, F. L., & Hayes, T. L. (2002). Business-unit-level relationship between employee satisfaction, employee engagement, and business outcomes: A meta-analysis. *Journal of Applied Psychology*, 87(2), 268–279. <https://doi.org/10.1037/0021-9010.87.2.268>
- Hausknecht, J.P., Rodda, J. and Howard, M.J. (2009). Targeted employee retention: Performance-based and job-related differences in reported reasons for staying. *Human Resource Management*, 48(2), 69-288. <https://doi.org/10.1002/hrm.20279>
- Herciu, M. (2017). Drivers of Firm Performance: Exploring Quantitative and Qualitative Approaches. *Studies in Business and Economics* 12(1), 79-84. <https://doi.org/10.1515/sbe-2017-0006>
- Hom, P. W., Lee, T. W., Shaw, J. D., & Hausknecht, J. P. (2017). One hundred years of employee turnover theory and research. *Journal of Applied Psychology*, 102(3), 530–545. <https://doi.org/10.1037/apl0000103>
- Judge, T. A., & Klinger, R. (2008). Job satisfaction: Subjective well-being at work. In M. Eid & R. J. Larsen (Eds.), *The science of subjective well-being*, 393–413. The Guilford Press.
- Kanyurhi, E., & Akonkwa, D. (2016). Internal marketing, employee job satisfaction, and perceived organizational performance in microfinance institutions. *International Journal of Bank Marketing*, 34(5), 773-796, <https://doi.org/10.1108/ijbm-06-2015-0083>
- Kessler, S., Lucianetti, L., Pindek, S., Zhu, Z., & Spector, P. (2020). Job satisfaction and firm performance: Can employees' job satisfaction change the trajectory of a firm's performance?. *Journal of Applied Social Psychology*, 50, 563–572. <https://doi.org/10.1111/jasp.12695>
- Labro, E., & Omartian, J. (2024). Managing Employee Retention Concerns: Evidence from U.S. Census Data. *The Accounting Review*, 1-27. <https://doi.org/10.2308/TAR-2021-0363>
- Lee, J., & Kang, J. (2017). A Study on Job Satisfaction Factors in Retention and Turnover Groups Using Dominance Analysis and LDA Topic Modeling with Employee Reviews on Glassdoor.com. *ICIS 2017 Proceedings*. 26. <https://aisel.aisnet.org/icis2017/DataScience/Presentations/26>
- Magnan, M. & Martin, D. (2019). Executive Compensation and Employee Remuneration: The Flexible Principles of Justice in Pay. *Journal of Business Ethics*, 160(1), 89-105. <https://doi.org/10.1007/s10551-018-3786-5>
- Melián-González, S., Bulchand-Gidumal, J. & López-Valcárcel, B. (2015). New Evidence of the Relationship Between Employee Satisfaction and Firm Economic Performance. *Personnel Review*, 44(6), 906-929. <https://doi.org/10.1108/PR-01-2014-0023>
- Milliken, F. J., Morrison, E. W., & Hewlin, P. F. (2003). An exploratory study of employee silence: Issues that employees don't communicate upward and why. *Journal of Management Studies*, 40(6), 1453-1476. <https://doi.org/10.1111/1467-6486.00387>

- Moro, S., Ramos, R.F. & Rita, P. (2021). What drives job satisfaction in IT companies?. *International Journal of Productivity and Performance Management*, 70(2), 391-407. <https://doi.org/10.1108/IJPPM-03-2019-0124>
- Muhammad, N., & Scrimgeour, F. (2014). Stock returns and fundamentals in the Australian market. *Asian Journal of Finance and Accounting*, 6(1). <https://doi.org/10.5296/ajfa.v6i1.5486>
- Nguyen, T. (2015). Employee Perks in Silicon Valley: Technology Companies Lead the “Arms Race” as Corporate Law Trails in Representing Shareholder Interests. *Hastings Business Law Journal*, 12(1). https://repository.uclawsf.edu/hastings_business_law_journal/vol12/iss1/3
- Oyer, P. (2004). Salary or benefits?. *Work, Earnings and Other Aspects of the Employment Relation*, 28, 429-467. <https://doi.org/10.2139/ssrn.632862>
- Pang, K., & Lu, C. (2018). Organizational motivation, employee job satisfaction and organizational performance: An empirical study of container shipping companies in Taiwan. *Maritime Business Review*, 3(1), 36-52. <https://doi.org/10.1108/MABR-03-2018-0007>
- Park, H., & Rahmani, M. (2020). Employee Satisfaction and Firm Innovation Performance. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.3860303>
- Phi, H., Huu, L., Quang, M., Minh, T., & Thi, O. (2021). Impact of Job Satisfaction on Organizational Performance, the Intermediate Role of Loyalty in the Health Industry: An Empirical Study from the University Medical Center of Ho Chi Minh City -Branch 2, Vietnam. *International Journal of Business Marketing and Management*, 6(9), 63-76.
- Podsakoff, P. M., MacKenzie, S. B., Lee, J.-Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology*, 88(5), 879–903. <https://doi.org/10.1037/0021-9010.88.5.879>
- PwC. (2023). Technology job market and employee trends. Retrieved 17 Apr 2024. <https://www.pwc.com/gx/en/issues/c-suite-insights/the-leadership-agenda/technology-job-market-employee-trends.html>
- Rizwan, M., Iqra, S., Waseem, M., & Sikander, S. (2014). The relationship of turnover intention with job satisfaction, job performance, leader member exchange, emotional intelligence and organizational commitment. *International Journal of Learning and Development*, 4(2), 242-256. <https://doi.org/10.5296/ijld.v4i2.6100>
- Salam, M., Azar, A., & Elgendy, M. (2021). The effect of different dimensionality reduction techniques on machine learning overfitting problem. *International Journal of Advanced Computer Science*, 12(4), 311–328. <https://doi.org/10.14569/IJACSA.2021.0120480>
- Schneider, C. (2016). The purposes of workplace benefits: A qualitative study of workers’ perceptions of work perks in tech and non-tech industries. *University of Texas at Austin*. <https://repositories.lib.utexas.edu/server/api/core/bitstreams/8838bb72-c60b-4f0a-bd2e-f154f5f30617/content>
- Schoenmueller, V., Netzer, O. & Stahl, F. 2019. The Extreme Distribution of Online Reviews: Prevalence, Drivers and Implications. *Columbia Business School Research Paper*, 18(10). <https://doi.org/10.2139/ssrn.3100217>
- Shiu, Y., & Yu, T. (2010). Internal marketing, organisational culture, job satisfaction, and organisational performance in non-life insurance. *The Service Industries Journal*. 30. 793-809. 10.1080/02642060701849840.

- Siepel, J., & Dejardin, M. (2020). How do we measure firm performance? A review of issues facing entrepreneurship researchers. *Handbook of Quantitative Research Methods in Entrepreneurship*, 2, 4-20.
<https://doi.org/10.4337/9781786430960.00006>
- Steil, A. V., Bello, J. S. A., Cuffa, D., & Freitas, A. F. (2022). Job Satisfaction and Employee Retention by Public and Private IT Organizations. *Revista de Administração da UFSM*, 15(3), 610-625.
<https://doi.org/10.5902/1983465963945>
- The Wharton School, University of Pennsylvania. (2024). *Wharton Research Data Services*. <https://wrds-www.wharton.upenn.edu/>
- Van Rooy, D. L., Whitman, D. S., Hart, D., & Caleo, S. (2011). Measuring Employee Engagement During a Financial Downturn: Business Imperative or Nuisance? *Journal of Business and Psychology*, 26(2), 147–152.
<http://www.jstor.org/stable/41474862>
- Wruck, K. & Wu, Y. (2021). Relative versus Absolute Performance Evaluation and CEO Decision-Making. *Journal of Financial and Quantitative Analysis*, Working Paper No. 2017-20. <https://doi.org/10.2139/ssrn.3048401>
- Yahoo. (2024). *Largest Companies In The Technology Sector*.
https://finance.yahoo.com/screener/predefined/sec-ind_sec-largest-equities_technology/

Appendix A

