Big Technology Acquisitions & Acqui-hire

Auteur : Akinci, Sevin
Promoteur(s) : Gautier, Axel
Faculté : HEC-Ecole de gestion de l'Université de Liège
Diplôme : Master en sciences économiques, orientation générale, à finalité spécialisée en macroeconomics and finance
Année académique : 2024-2025
URI/URL : http://hdl.handle.net/2268.2/23984

Big Technology Acquisitions & Acqui-hire

Master thesis presented by
Sevin Akinci

To obtain the degree of MASTER IN ECONOMICS
with a specialization in
Macroeconomics and Finance

Academic year 2024/2025

Jury:
Supervisor: Prof. Gautier Axel
Readers:
- Gathon Henry-Jean
- Schoenmaeckers Jérôme

Contents

Acknowledgements

I would like to express my sincere thanks to Prof. Axel Gautier, who supported and guided me during the entire project work phase.

Also, I would like to express my gratitude to Associate Prof.Jérôme Schoenmaeckers, who helped and guided me during the entire project work phase.

List of Abbreviations and Glossary

- **GAFAM** Acronym for five major technology firms: Google, Apple, Facebook, Amazon, Microsoft.
- **M&A** Mergers and Acquisitions.
- **VC** Venture Capital. Funding provided to startups and small businesses with high growth potential.
- **KM** KaplanMeier Estimator. A nonparametric statistic used to estimate survival functions from lifetime data.
- **Cox PH Model** Cox Proportional Hazards Model. A regression model for investigating the association between survival time and explanatory variables.
- **ROC Curve** Receiver Operating Characteristic Curve. A graphical plot illustrating the diagnostic ability of a binary classifier.
- **AUC** Area Under the Curve. A measure of the accuracy of a classification model.
- **AIC** Akaike Information Criterion. A model selection criterion based on the trade-off between goodness of fit and model complexity.
- **BIC** Bayesian Information Criterion. A criterion similar to AIC but with a stronger penalty for models with more parameters.
- **CAP** Continue After Purchase. A binary indicator showing whether a founder remained employed at the acquirer post-acquisition.
- **Survival Probability** The probability that an individual (founder) remains in the company beyond a certain time threshold.
- **Right-Censoring** In survival analysis, a situation where the event of interest (e.g., founder exit) has not occurred for some individuals by the end of the study.
- **Hazard Ratio (HR)** In the Cox model, this represents the effect of a covariate on the risk of the event occurring.
- **AIC / BIC** Information criteria used to compare model fit. Lower values indicate better fit.
- **VIF** Variance Inflation Factor. A diagnostic tool used to detect multicollinearity in regression models.

Chapter 1

Introduction

Over the past decade, the growing dominance of Big Tech firms such as Google, Amazon, Facebook, Apple, and Microsoft (GAFAM) has raised substantial concerns among academics, policymakers, and regulators. While these firms frequently justify their acquisition activities as strategies to enhance innovation, improve efficiency, or expand technological capabilities, there is growing apprehension that such mergers are often driven by anti-competitive motives. Notably, one increasingly common rationale behind these acquisitions is the so-called "acqui-hiring" strategywhere startups are purchased not primarily for their products or market share, but to internalize their human capital, particularly the talents and expertise of founders and technical teams. Despite the growing prominence of this phenomenon, the empirical understanding of post-acquisition founder behavior remains limited. This thesis addresses this gap by investigating the factors that influence whether, and for how long, startup founders remain with the acquiring firm following an acquisition.

The increasing wave of startup acquisitions by major digital platforms is not merely a result of market expansion but has strategic undertones that affect innovation, labor allocation, and competitive dynamics. Many of these acquisitions involve nascent firms that have not yet reached their full market potential, raising concerns about whether the acquiring firms are attempting to pre-empt future competition. From this perspective, the acqui-hire strategy becomes a vital subject of investigation as it involves not only the absorption of a product or service but also the integration of individuals who drive innovation within the startup. Founders and key team members carry not only technological expertise but also the vision and organizational knowledge that often serve as the startup's core assets. Understanding the tenure of these individuals post-acquisition can reveal much about the acquiring firms strategic intentions and the success of human capital integration.

This research is situated within a broader academic debate that draws heavily on previous contributions, notably the work of Gautier and Maitry (2024) and Gautier and Lamesch (2021). The former study, "Big Tech Acquisitions and Product Discontinuation," explores how GAFAM acquisitions lead to product shutdowns, using models such as logit and probit to assess the likelihood of continuation. The latter paper, "Mergers in the Digital Economy," applies difference-in-differences techniques to understand how these acquisitions influence innovation, market concentration, and consumer welfare. Both papers focus primarily on product outcomes; however, they also highlight the relevance of founder involvement in shaping post-merger results. This thesis builds upon their foundation but shifts the focus explicitly toward human capital and founder retention dynamics.

The core objective of this thesis is to examine how talent acquisition plays out in the context of GAFAM mergers. Specifically, it analyzes whether founders of acquired startups remain at the acquiring firm post-acquisition, and if so, for how long. This focus on durationas opposed to a binary presence/absence measureenables a more detailed analysis of human capital retention. The empirical work relies on a manually constructed dataset that combines firm-level and founder-level information on over 300 startup acquisitions. Data were gathered from sources including Crunchbase, LinkedIn, company websites, and other public databases. The cleaned dataset includes variables related to company age, funding history, product segment, acquirer identity, founder count, and individual founder tenure.

Methodologically, the empirical strategy is divided into two complementary stages. In the first stage, binary outcome modelsspecifically Logit and Probit regressionsare employed to estimate the probability of founder retention based on various startup and founder characteristics. These models also explore whether a founder stays at least two years, capturing a threshold commonly associated with successful integration. In the second stage, duration models are applied. KaplanMeier survival curves are used to visualize the probability of continued employment over time, while Cox Proportional Hazards models quantify the effects of covariates such as firm age, funding level, founder origin, and startup segment on the likelihood of founder departure. This dual approach enables the thesis to move beyond simple binary outcomes and capture the dynamics of founder retention across time.

The findings suggest that both individual and firm-level factors significantly shape post-acquisition founder retention. For instance, startups with a greater number of founders and those with older firm age at the time of acquisition tend to exhibit lower retention rates. In contrast, founders from U.S.-based firms or from specific product segments are more likely to remain employed by the acquiring company for extended periods. Kaplan-Meier survival curves further reveal how retention probabilities diverge over time based on founder and firm characteristics, while Cox regressions confirm the statistical significance of these factors. These insights contribute new empirical evidence to the discussion of acqui-hiring and illuminate the complexity of human capital integration in post-merger scenarios.

By focusing on the retention and behavior of startup founders after acquisition, this thesis adds a novel perspective to the literature on mergers in the digital economy. Unlike traditional models that limit analysis to firm-level outcomes or product continuity, the use of survival analysis allows for a deeper understanding of the temporal dynamics involved in talent acquisition. These insights are particularly relevant for policymakers and regulators seeking to evaluate the broader consequences of GAFAM acquisitionsnot only in terms of market power but also in terms of talent consolidation and innovation capacity.

The thesis is structured into five chapters. Chapter 1 introduces the research question, motivation, and relevance of the study. Chapter 2 provides a detailed review of the related literature, identifying the main theoretical and empirical contributions. Then, discusses the dataset, variable construction, and data manipulation processes. Presents the empirical methodology and results, including both binary and duration-based models. Finally, concludes with a summary of key findings, reflections on limitations, and suggestions for future research directions.

Chapter 2

Developments

2.1 Literature Review

Merger and acquisition strategies of large technology companies (GAFAM) are critical processes that reshape the dynamics of competition and innovation in the digital economy. These processes significantly affect the market structure and the entrepreneurial ecosystem. Literature studies examine issues such as how much these acquisitions hinder the market entry of new ventures, how they direct innovation activities, and the focus points of large companies' employee investment. At this point, killer mergers and acqui-hiring strategies stand out as two important dimensions in the acquisition processes of large technology companies (GAFAM). While killer mergers focus on the strategy of acquiring ventures to prevent competition and then discontinuing products, the acqui-hiring strategy examines acquisitions from the perspective of talent acquisition and human capital gain.

A considerable body of research investigates the strategic motivations behind such acquisitions, as well as their implications for innovation, competition, and organizational behavior. Various empirical methods such as probit and logit regression models, difference-in-differences techniques, and survival analysisare frequently employed to explore these effects. The findings of these studies suggest that while killer mergers often lead to reduced competition and suppressed innovation, acqui-hiring may produce more nuanced outcomes by facilitating talent integration, which can sometimes promote innovation within the acquiring firm. The following section reviews the most influential studies on these themes and categorizes them according to their methodological focus and research contribution.

To begin with, Gautier and Maitry (2024) explore the post-acquisition product life cycle of startups acquired by major technology firms. Using probit and logit models, they investigate the extent to which product discontinuation occurs as a strategic tool to neutralize potential competition. Their study emphasizes the role of variables such as product segmentation, firm age, and founder retention in shaping post-acquisition outcomes. Their findings indicate that a significant proportion of acquired products are discontinued shortly after the acquisition, reinforcing concerns about anti-competitive motives.

Complementing this line of inquiry, Gautier and Lamesch (2021) adopt a difference-in-differences approach to analyze the impact of mergers in the digital economy. Their study uses sector-level data to examine changes in market structure, innovation intensity, and pricing behavior following GAFAM acquisitions. They conclude that mergers are often followed by decreased innovation activities and increased market concentration. By incorporating product segmentation into their analysis, the authors offer a more granular view of how specific types of products are affected.

An essential contribution to the literature on killer acquisitions is made by Cunningham, Ederer, and Ma (2021), who apply a regression discontinuity design to assess how incumbents acquire emerging rivals with the intent to terminate overlapping innovation pipelines. Their research finds that acquired firms with competitive potential are often shut down to prevent future threats, a practice that leads to a measurable decline in innovation outputs in affected markets.

Focusing on the acquiring-hiring aspect, Bar-Isaac, Johnson, and Nocke (2024) present a theoretical model linking acquiring-hiring approaches to monopolistic power in niche labor markets. Their research shows that large technology firms often acquire startups not only to acquire talent but also to reduce potential competition in the labor market, leading to reduced wage competition and diminished employee bargaining power. The authors emphasize that while acquiring-hiring can be advantageous for acquirers and venture investors, it often has negative consequences for employees and can lead to socially inefficient outcomes. This study highlights the need to consider labor market effects when formulating antitrust policies and evaluating acquisitions.

Further theoretical support for the killer acquisition hypothesis is provided by Rinehart (2024), who employs qualitative case studies and a conceptual framework to investigate the broader implications of such strategies. He emphasizes that acquired firms in highly competitive and profitable markets are particularly vulnerable to post-acquisition shutdowns. His findings suggest that these acquisitions significantly limit consumer choice and curtail technological diversity, reinforcing the need for regulatory scrutiny.

Focusing on founder retention and the human capital dimension, Bena and Li (2014) examine how corporate innovation outcomes are affected by founder exit behavior post-acquisition. Using Cox proportional hazards models, they explore the relationship between acquisition strategy, founder characteristics, and organizational integration. Their results indicate that founders of more innovative firms are more likely to leave earlier, especially in cases where strategic realignment occurs within the acquiring company.

In a related study, Ewens and Marx (2018) investigate the implications of founder replacement for startup performance following acquisition. Applying survival analysis techniques, they show that founder tenure is associated with organizational stability and performance outcomes, and that premature founder exits can sometimes disrupt firm growth trajectories. Their research underscores the importance of post-merger human capital retention in sustaining innovative capacity.

Koski, Kässi, and Braesemann (2020) extend the killer acquisition literature by examining entry barriers and venture capital dynamics in the context of early-stage startups. Employing survival analysis, they reveal that startups in competitive environments face a high risk of acquisition and discontinuation. The study also highlights the strategic behavior of large acquirers who use mergers to reduce market entry threats and consolidate technological control.

Meanwhile, Motta and Peitz (2021) adopt a panel data approach to study how big tech mergers affect R&D investment, market efficiency, and long-term innovation patterns. Their findings suggest that although some acquisitions may generate short-term efficiencies, the broader impact tends to involve reductions in R&D spending and innovation capacity. The study reinforces concerns about the potential stifling of competition in digital markets.

Barsy and Gautier (2024) offer a complementary empirical investigation using a difference-in-differences model to compare innovation outputs before and after acquisitions. Their data includes patent filings, product development timelines, and R&D expenditures. They find that post-acquisition innovation generally declines, although some acquiring firms are able to utilize the acquired R&D resources more effectively when internal capabilities align with external technologies.

Finally, Cabral (2021) adopts a game-theoretic approach to model the competitive dynamics of digital mergers. He focuses on the role of network effects, platform competition, and strategic exclusion. His findings indicate that post-merger strategies often lead to enhanced market power for incumbents, with ambiguous implications for consumer welfare depending on the strength of countervailing market forces.

Taken together, this literature highlights a complex set of outcomes associated with big tech acquisitions. The dual strategies of killer acquisitions and acqui-hiring not only reshape competition and innovation trajectories but also raise critical questions about regulatory oversight. The use of robust econometric and theoretical frameworksranging from logit/probit models to Cox survival analysis and game-theoretic simulationsunderscores the multifaceted nature of this research field. Understanding the behavioral patterns of founders and the strategic intents of acquirers is essential to inform both academic debates and policy interventions in the digital economy.

2.2 Data Description

2.2.1 Raw Data

The initial raw dataset utilized in this research, Gautier_Lamesch_data.xls, provides a summary of technology startups that were acquired by GAFAM (Google, Amazon, Facebook, Apple, Microsoft) firms from 2015 to 2017. Each entry corresponds to a startup company. The dataset is specifically designed to illustrate the structural features of the firms prior to and following the acquisition, as well as the ongoing involvement of the founders in the business.

The total number of observations in the data is 175, and the total number of variables is 11 and a total of 782 missing values were detected. The data type has a "cross-sectional" data type, reflecting the startups that were acquired from 2015 to 2017, with each observation corresponding to a unique firm without any repeated time entries.

Here are five randomly selected examples from the raw dataset prior to analysis.

#	Company Name	Origin	Foundation Year	Product	Segment	Funding Rounds	Total Funding (USD)	Acquirer	Acquisition Year	Age	Continuation
1	Elemental Technologies	US	2006	Video streaming	Editors	7	45,650,000	AMZN	2015	9	discont.
2	Safaba Translation Solutions	US	2009	Translation software	Businesses	3	NA	AMZN	2015	6	discont.
3	Shoefitr	US	2010	3D modeling	Merchants	3	1,255,000	AMZN	2015	5	discont.
4	2lemetry	US	2011	IoT platform	Businesses	2	10,000,000	AMZN	2015	4	discont.
5	Amiato	US	2011	Analytics services	Businesses	3	2,000,000	AMZN	2015	4	NA

Table 2.1: Sample Data

These instances were selected to illustrate the overall format of the dataset, the nature of the variables, and the missing or formal errors encountered before the data cleaning process. Notable missing data (NA) is particularly evident in the "Total Funding" and "Continuation" variables, while variations in upper and lower case letters, spacing, and spelling inconsistencies can also be seen in the categorical variables.

Variable Name	Storage Type	Description
Companyname	String (str30)	Name of the company being purchased (174 unique values)
Origin	String (str3)	Country of origin of the company (e.g. US)
Fondationyear	String (str4)	The company's founding year (26 different year values)
Product	String (str31)	The technology area in which the company operates (e.g. cloud, dev tool)
Segment	String (str11)	The firm's target customer segment (e.g. businesses, consumers)
Fundingrounds	String (str2)	Total number of investment rounds (example: 1, 2, NA)
Totalfunding	String (str9)	Total amount of investment received (USD)
Acquirer	String (str4)	Purchased by GAFAM company (AMZN, GOOG, etc.)
Acquisitionyear	Integer (int)	Purchase year (2015, 2016 or 2017)
Ageinyears	String (str2)	Age of the company at the time of purchase
Continuation	String (str8)	Label indicating whether the product is working after purchase (running, discont., NA)

Table 2.2: Description of Variables
Source: R Studio

The second dataset employed in this research is known as Combined_Data_revised_2024.xlsx, which includes nearly identical variables. It provides information on the structural features of technology companies that were acquired by GAFAM entities. Every row corresponds to an individual startup initiative. This dataset served as the primary resource for examining founder continuity by including both pre- and post-acquisition characteristics of the startups, as well as the industries in which the firms operate.

When the general features of the dataset are examined, there are 300 observations and 12 variables and a total of 782 missing values were detected.The acquisition dates in the data include acquisitions that took place between 2015-2021.

Here are five randomly selected examples from the second raw dataset prior to analysis;

#	Company Name	Found. Year	Funding Rounds	Total Funding (USD)	Acquirer	Acquisition Year	Age	Continuation	US or Not	Cluster	Main	Continuation 2023
1	Graphiq	2009	5	32,000,000	AMZN	2017	8	discont.	US	Artificial intelligence, data science, and analytics	0	Announced
2	Cloud9 IDE	2010	3	9,400,000	AMZN	2016	6	running	US	Artificial intelligence, data science, and analytics	0	Integrated
3	TSO Logic	2013	NA	NA	AMZN	2019	6	discont.	NON-US	Artificial intelligence, data science, and analytics	0	Integrated
4	Orbeus	2012	2	1,466,387	AMZN	2015	3	discont.	US	Artificial intelligence, data science, and analytics	0	not active
5	Partpic	2013	2	1,520,000	AMZN	2016	3	discont.	US	Artificial intelligence, data science, and analytics	0	not active

Table 2.3: Sample Data2

The observations indicate the presence of both numerical and categorical variables within the dataset. Additionally, certain variables contain missing values ("NA").

The table illustrates typical entries found in the Combined_Data_revised_2024 dataset prior to data cleaning. The table clearly highlights the existence of missing data, discrepancies in formatting, and diverse categories, underscoring the need for a thorough data cleaning process before conducting analysis.

Variable	Storage Type	Description
Companyname	string	Name of the acquired startup (300 unique companies)
Foundationyear	numeric	Year of establishment of the company (between 19242018)
Fundingrounds	string	Number of investment rounds received (some NA)
Totalfunding	string	Total investment amount received by the company (USD, some NA)
Acquirer	string	Purchased by GAFAM company (AMZN, GOOG, etc.)
Acquisitionyear	numeric	Purchase year (20152021)
Ageinyears	numeric	Age of the firm at the time of purchase (097 years)
Continuation	string	The product's continuation status after purchase (running or discont.)
USornonUS	string	Whether the company is based in the US (US, NON-US, NA)
Cluster	string	Product scope/technology stack (e.g. AI, analytics, Digital content)
Main	numeric	Whether it is a main field of activity (1 = Yes, 0 = No)
Contiuation23	string	Detailed classification of the company's status (e.g. "Active", "Integrated")

Table 2.4: Description of Variables2
Source: R Studio

2.2.2 Sample Construction

The raw data referenced in Section 2.2.1 was utilized, and information was collected based on the startup company name. The data collected for inclusion in the new data set was at the level of founder information based on startups. The information was collected manually from the internet, Crunchbase, LinkedIn, and other sources.

The gathered information focused on the number of founders associated with each startup, whether any of these founders remained with GAFAM post-acquisition, the number of years they worked there if they did, their current employment status at GAFAM, and if the founder departed prior to the company's acquisition by GAFAM. This data was incorporated into the existing raw data. Subsequently, these two sets of information were merged based on the "Company name" variable, and duplicates were eliminated. During the merging process, discrepancies in uppercase/lowercase letters, spaces, and character issues were addressed; unmatched or missing entries were disregarded in the analysis. As a result of these steps, a more refined and enhanced dataset suitable for analyzing founder behavior was created.

The final dataset is a panel data consisting of 685 observations and 19 variables and is suitable for analyzing founder continuity in firms acquired by GAFAM. This dataset serves as the primary data source for the study as it covers both firm characteristics and founder-level behaviors.

#	Company Name	Origin	Foundation Year	Product	Segment	Funding Rounds	Total Funding	Acquirer	Acquisition Year	Cluster	Main	Continuation 2023	# Founders	Cont. Rate	Working Rate	Leave Before
1	Elemental Technologies	US	2006	Video streaming	Editors	7	45,650,000	AMZN	2015	Remote storage and file transfer	0	Integrated	3	1	4	0
2	Elemental Technologies	US	2006	Video streaming	Editors	7	45,650,000	AMZN	2015	Remote storage and file transfer	0	Integrated	3	0	0	0
3	Elemental Technologies	US	2006	Video streaming	Editors	7	45,650,000	AMZN	2015	Remote storage and file transfer	0	Integrated	3	0	0	0
4	Safaba Translation Solutions	US	2009	Translation software	Businesses	3	NA	AMZN	2015	AI, data science, and analytics	0	Announced	2	1	3.9	0
5	Safaba Translation Solutions	US	2009	Translation software	Businesses	3	NA	AMZN	2015	AI, data science, and analytics	0	Announced	2	1	3.9	0

Table 2.5: Final Dataset

The table below outlines the variables, data types, and definitions contained in this extensive data set:

Variable	**Data Type**	**Description**	**Missing**
Company Name	object	Name of the acquired startup	0
origin	object	Country of origin (US or NON-US)	15
foundation year	int64	Year the company was founded	0
product	object	Type of product or service offered by the startup	318
segment	object	Target customer segment	327
funding rounds	float64	Number of investment rounds the startup received	136
total funding	object	Total amount of funding received (USD)	118
acquirer	object	Acquiring GAFAM firm (e.g., Amazon, Google, etc.)	0
acquisition year	int64	Year in which the acquisition occurred	0
age (in years)	int64	Age of the startup at the time of acquisition (in years)	0
continuation	object	Post-acquisition product status (running or discontinued)	45
Cluster	object	Strategic domain or technological cluster (e.g., AI, developer tools)	45
Main	float64	Indicates if the acquired asset was the startups core product (1 = Yes, 0 = No)	45
Continuation 23	object	Detailed classification of the post-acquisition status (e.g., Active, Integrated)	45
number od founders	int64	Total number of identified founders of the startup	0
Continue to work after	int64	Whether any founder continued at the acquiring firm post-acquisition	0
Working Age(Year)	float64	Years the founder worked at the acquiring firm post-acquisition	0
does it still work GAFAM?	int64	Whether the founder still works at the GAFAM company (1 = Yes)	0
leave before purchase	int64	Whether the founder left before the acquisition (1 = Yes)	0

Table 2.6: Merge Data Description
Source: R Studio

Before proceeding to the modeling phase, a series of pre-processing steps were applied to clean and structure the dataset effectively. The primary aim was to improve data quality, ensure consistency, reduce model complexity, and derive explanatory variables aligned with the research objectives. Below is a summary of the key transformations and the reasoning behind them:

1. **Standardization of Variable Names:** All column names were converted to lowercase and stripped of special characters and whitespace. This normalization was performed to ensure compatibility during coding and to facilitate easier referencing in modeling functions. For instance, the variable *"Working Age(Year)"* was renamed to `working_age` and *"Does it still work GAFAM?"* to `still_work_gafam`.

2. **Text Cleaning in Character Variables:** For all character-type variables, values were transformed to lowercase, leading/trailing spaces were removed, and common delimiters (e.g., commas, hyphens) were replaced. Space characters were substituted with underscores to prevent issues during dummy variable creation and to ensure consistent formatting across variables.

3. **Grouping of Country of Origin:** The original `origin` variable was recoded into a binary grouping named `origin_group`. Startups based in the United States were categorized as `usa`, while all others were labeled as `other`. This transformation aimed to assess whether being based in the U.S. influences post-acquisition founder retention behavior.

origin_group	**count**
usa	450
other	200

Table 2.7: Grouped Origin Summary
Source: R Studio

Visual Insight: Origin Distribution

The figure below illustrates the relative frequency of U.S.-based startups versus others, providing an overview of the country-level representation in the dataset.

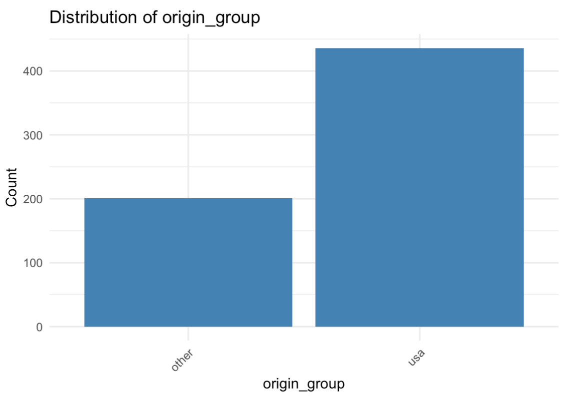


Figure 2.1: Distribution of Origin Group
Source: R Studio

4. **Product Category Simplification:** The dataset contained a wide range of product descriptions, many of which were too granular for direct modeling. These products were grouped into six higher-level categories under the variable `product_group`,

such as `media_entertainment_and_social_networks` and `technology_hardware_and_security`. The purpose of this grouping was to improve interpretability and avoid excessive dimensionality from sparse product-level dummy variables.

product_group	product	count
business_and_productivity_tools	productivity	24
business_and_productivity_tools	analytics	20
other	dev_tool	32
other	cyber_security	16

Table 2.8: Grouped Product Summary (Top Examples)
Source: R Studio

Visual Insight: Product Group Distribution

The bar chart below illustrates how the startups are distributed across the defined product groups, providing a quick understanding of category balance.

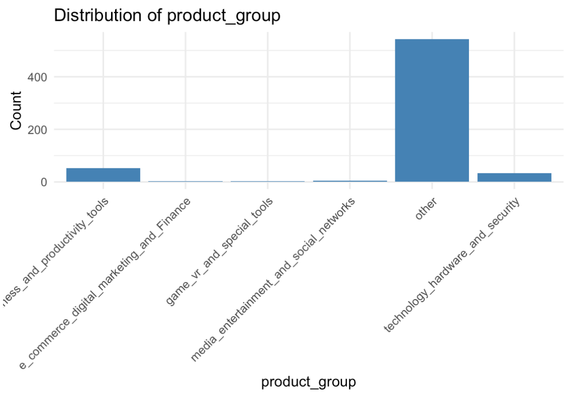


Figure 2.2: Distribution of Product Group
Source: R Studio

5. **Segment Categorization:** The original `segment` variable was reclassified into `segment_group`, consolidating values into broader labels: `business`, `consumer`, `editor`, and `other`. This transformation was driven by the desire to observe segment-specific trends in founder outcomes.

segment_group	segment	count
business	businesses	120
consumer	consumers	57
editor	editors	94
other	other	301

Table 2.9: Grouped Segment Summary
Source: R Studio

Visual Insight: Segment Group Distribution

The figure below complements the tabular summary by presenting how different segment types are represented within the dataset.

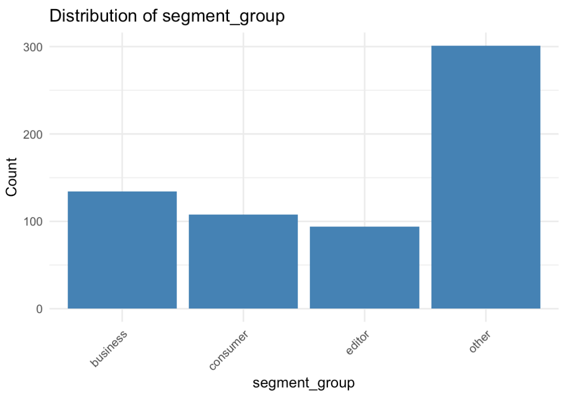


Figure 2.3: Distribution of Segment Group
Source: R Studio

6. **Cluster Aggregation:** The cluster variable was regrouped into cluster_group using three broader technological domains: technology_and_data_analytics, communication_and_digital_services, and physical_and_personal_services. The aim was to avoid overfitting caused by the presence of many rarely observed clusters while maintaining meaningful thematic distinctions.

cluster_group	**cluster**	**count**
technology_and_data_analytics	AI_data_analytics	189
technology_and_data_analytics	dev_tools	69
communication_and_digital_services	digital_content	76
physical_and_personal_services	physical_goods	108

Table 2.10: Grouped Cluster Summary
Source: R Studio

Visual Insight: Cluster Group Distribution

The graph that follows offers a visual representation of startup distribution across aggregated technological clusters.

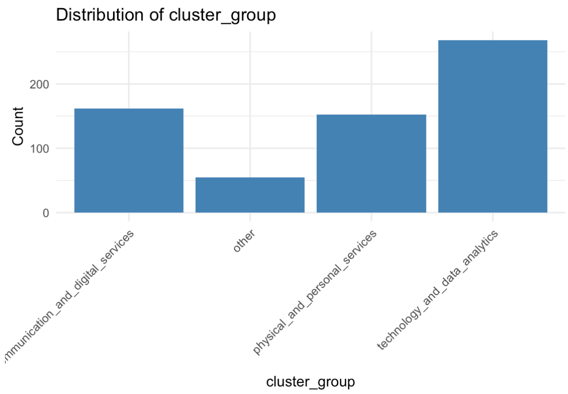


Figure 2.4: Distribution of Cluster Group
Source: R Studio

7. **Continuation Status Recoding:** The `continuation` variable was recoded into `continuation_group`, classifying firms as `running`, `discontinued`, or `unknown`. This variable serves as a proxy for the product life cycle and operational continuity.

8. **Creation of Foundation Year Groups:** Firms were split into two groups based on their year of foundation, creating a variable called `foundation_year_group`, using 2012 as a threshold. This distinction aims to capture generational differences in startup behavior and acquisition outcomes.

9. **Conversion of Key Variables to Factor Types:** All categorical predictors such as `acquirer`, `continuation_group`, and `product_group` were explicitly converted to factor data types. This step ensures proper treatment in regression models, especially for dummy variable encoding.

10. **Handling Missing Values in Financial Variables:** Missing values in `funding_rounds` and `total_funding` were imputed using the average of startups acquired by the same acquirer. This method preserves acquirer-specific financial characteristics while preventing data loss due to NA values.

Visual Insight: Funding Distributions and Patterns

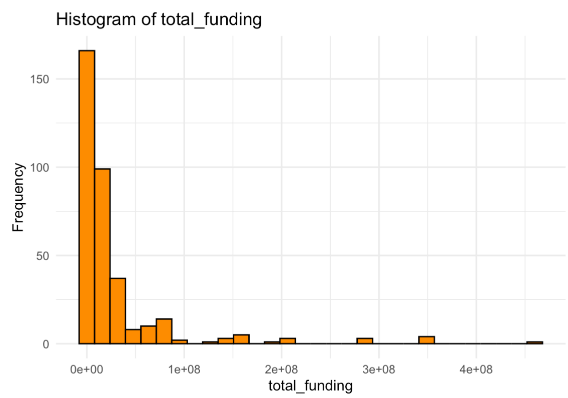


Figure 2.5: Histogram of Total Funding
Source: R Studio

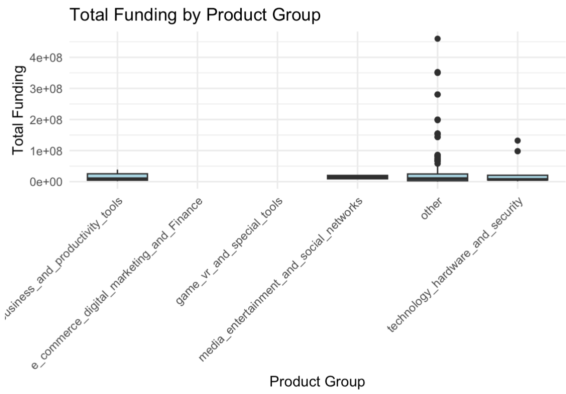


Figure 2.6: Total Funding by Product Group
Source: R Studio

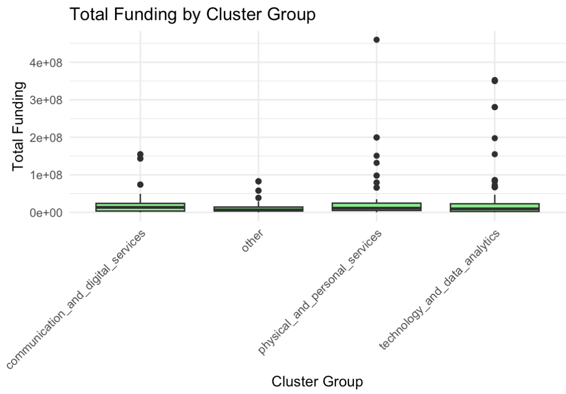


Figure 2.7: Total Funding by Cluster Group
Source: R Studio

11. **Exclusion of Non-Eligible Observations:** Firms where founders had left before acquisition (`leave_before_purchase = 1`) were excluded, as these observations would not contribute to the analysis of post-acquisition founder retention behavior, which lies at the heart of this study.

12. **Definition of Duration-Based Outcome Variables:** To study different time horizons of founder retention, three binary variables were generated based on the `working_age` variable:

    - `still_at_4_years` for founders retained less than 4 years
    - `still_at_8_years` for those retained between 4 and 8 years
    - `still_at_10_years` for those retained at least 8 years

    These variables serve as targets in separate binary classification models, enabling analysis across different tenure ranges.

Visual Insight: Working Age Distribution

Figure 2.8: Histogram of Working Age
Source: R Studio

13. **Creation of Duration Group Categories for Multinomial Modeling:** In addition to binary classifications, a categorical variable named `duration_group` was created to segment founders into three duration tiers: `short_term` (0–4 years), `mid_term` (4–8 years), and `long_term` (8+ years). This variable enables multinomial logistic regression, offering a more holistic perspective on founder tenure outcomes.

Summary and Transition to Modeling

To align with the analytical goals of this study, the initial dataset underwent a thorough process of data cleaning and transformation. Variables were standardized, restructured, categorized, and missing values were appropriately filled in. Notably, the information at the firm level was harmonized with the data at the founder level, facilitating a comprehensive examination of founder behavior post-acquisition.

The graphical representations and frequency tables included in this part confirm the efficacy of the grouping methods and offer a more transparent view of the data distribution. Consequently, the final dataset is devoid of structural discrepancies and absent values in the primary fields, rendering it appropriate for statistical modeling.

In the following section, we will develop binary and multiple logistic regression models, along with survival analysis models, to explore the elements influencing the duration that founders remain in firms acquired by GAFAM companies. These models are designed to uncover trends in founder retention across various time frames by leveraging the explanatory variables generated during the data manipulation stage.

2.3 Methodology

2.3.1 Survival Analysis Approach

To investigate how long founders remain after being acquired by large tech companies, this study utilizes survival analysis techniques. Survival analysis is a statistical method that models the time until a specific event occursin this instance, the exit of founders from the acquiring organization. Considering the nature of the research question, which involves right-censored data (i.e., founders still working at the acquirer at the time of the last observation), this approach provides several methodological benefits compared to conventional regression methods.

The expanding body of research on mergers and acquisitions within the digital economy, especially those involving GAFAM companies, has underscored the importance of understanding not just whether the acquired startups are ultimately closed down (often referred to as killer acquisitions), but also the duration for which their key talentparticularly the foundersare kept. Various studies highlight that the strategic reasons for these acquisitions frequently revolve around the acquisition of knowledge and skilled personnel (Ewens & Marx, 2018; Bena & Li, 2014). Within this framework, acqui-hiring stands out as a specific strategy where the main aim of the acquiring company is to incorporate the founding team instead of focusing on the product or technology itself.

Survival models are well-suited to analyze such post-acquisition dynamics. Unlike binary outcome models (e.g., logit/probit), survival analysis takes into account not only whether an event occurs but also when it occurs. This time-sensitive dimension is particularly useful in tracing how long founders stay within the acquiring firm, which in turn may reflect the alignment between acquisition strategy and organizational integration.

Furthermore, previous research demonstrates that founder tenure can have significant implications for firm performance and innovation continuity after acquisition. For instance, Ewens and Marx (2018) show that the premature departure of founders is often associated with reduced organizational stability. Similarly, Bena and Li (2014) employ Cox proportional hazards models to demonstrate how certain founder characteristicssuch as prior experience, age, or the innovativeness of the startupcan influence post-merger retention durations.

Survival analysis also provides a deeper understanding of differences at the group level, including variations among acquirers, segments, or funding profiles. This understanding is especially crucial in the context of acquisitions by GAFAM, as the reasons behind these acquisitions and the strategies for integration can vary significantly between companies. As highlighted by Koski et al. (2020), startups that operate in highly competitive settings or specific technological niches encounter an increased likelihood of both being acquired and ceasing operations. Utilizing survival analysis allows researchers to explore how such contextual elements influence retention patterns over time.

In this study, both Kaplan-Meier estimators and Cox proportional hazards models are employed. Kaplan-Meier curves are used to visualize and compare survival probabilities across key subgroups, while the Cox model estimates the effects of explanatory variables on the hazard rate of founder exit. The combined use of these methods ensures a comprehensive approach to understanding post-acquisition human capital retention in the digital economy.

2.3.1.1 Working Duration Analysis of Founding Partners with Kaplan-Meier Estimator

In this research, the Kaplan-Meier (KM) estimator is utilized to analyze how the duration of employment for co-founders is influenced following acquisition processes conducted by major technology companies (GAFAM). The KM estimator is a nonparametric technique that effectively handles right-censored data and can generate reliable outcomes while maintaining the observable sample size in scenarios where an event does not take place (for example, founders still employed at the company throughout the analysis timeframe).

Given that numerous co-founders of the companies remain employed during the analysis period, the occurrence of censored observations is unavoidable. As a result, the KM estimator is favored for computing time-dependent probabilities of founder retention.

Application of the Model and Visualization Methods

With the help of the KM estimator, the probability of founders staying in the company is calculated according to the years since the acquisition date. In the analysis, the probability of survival is estimated for each time period and comparisons are made between different groups. These groups are separated according to variables such as the acquiring firm, product continuation status, segment, country of origin, retention type and startup cluster group.

Median employment duration was calculated separately for each group, and to enhance visual representation, these values were added to Kaplan-Meier curves with horizontal and vertical lines (in R software, using the surv_median() function and the surv.median.line = "hv" option). R software was used in visualizations, and survfit() and ggsurvplot() functions were preferred as basic tools.

Interpreting Kaplan-Meier Curves

The graphs presented below show the differences in Kaplan-Meier curves for the tenure of co-founders by group. Each graph aims to reveal the potential impact of a particular variable on founder persistence. p-values were obtained with the log-rank test.

In this part, I analyze how long founders stay at GAFAM companies after their startups are acquired. I use the Kaplan-Meier (KM) method, which is a non-parametric way to estimate the probability that a founder continues working at GAFAM over time.

The variable `working_age` shows how many years a founder worked at GAFAM after the acquisition. If a founder has `working_age = 0`, it means they did not work at GAFAM or left very early. The event indicator `still_work_gafam` is used to mark if the founder left (1) or is still working (0), so this is a right-censored survival model.

I used the following survival object:

```
Surv(time = working_age, event = still_work_gafam)
```

The Kaplan-Meier survival curve is shown in Figure 2.9. At the start, around 80% of the founders are still working. After 2 years, this goes below 65%, and after 5 years, only about 30% are still working. At year 8, less than 20% remain.

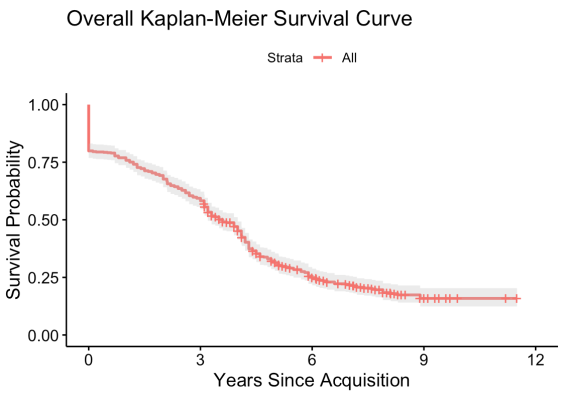


Figure 2.9: Overall Kaplan-Meier Sruvival Curve
Source: R Studio

This shows that many founders leave GAFAM companies in the first few years. The decrease is faster in the first 3 years. This could mean that GAFAM firms mainly want fast knowledge transfer or talent access through acquisitions. The drop in survival suggests that long-term employment may not be the main goal.

Table 2.11: Selected Kaplan-Meier Survival Estimates

Years	Events	Survival Probability	95% Confidence Interval
0	128	0.799	[0.769, 0.831]
1	7	0.758	[0.726, 0.792]
2	10	0.677	[0.641, 0.714]
3	7	0.582	[0.545, 0.622]
4	11	0.452	[0.414, 0.492]
5	4	0.314	[0.279, 0.354]
6	3	0.249	[0.216, 0.288]
7	1	0.217	[0.184, 0.255]
8	2	0.159	[0.124, 0.203]

Source: R Studio

Note: These are selected values from the full Kaplan-Meier output. The survival probability shows the proportion of founders still working at GAFAM at each time point.

Comparison of Founder Retention by Acquirer

To understand whether retention rates differ across GAFAM firms, I estimated Kaplan-Meier survival curves separately for each acquiring company. The goal here is to identify if founders tend to stay longer at one company compared to others after the acquisition.

Figure 2.10 shows the survival curves for five companies: Amazon (amzn), Apple (appl), Facebook (fcbk), Google (goog), and Microsoft (msft). Each curve represents the probability that a founder is still working at the acquiring firm over time. The shaded areas show 95% confidence intervals. A log-rank test was conducted to test if the differences are statistically significant.

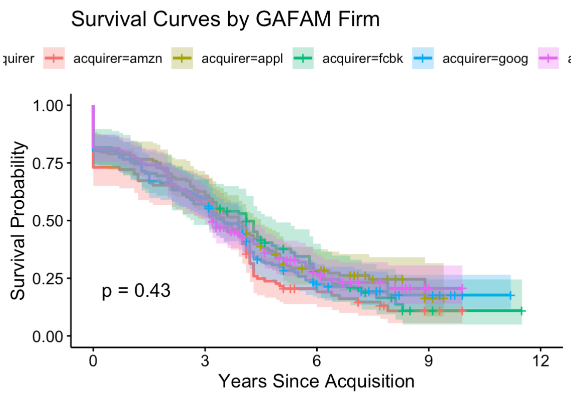


Figure 2.10: Survival Curves by GAFAM
Source: R Studio

According to the log-rank test, the p-value is 0.43. This means that the differences in survival probabilities among the companies are not statistically significant at the 5% level. So, we cannot say for sure that one firm retains founders longer than others. However, some differences can still be seen when looking at the survival curves and summary values.

Table 2.12 shows a basic comparison of survival probabilities at selected time points2, 4, and 6 years after the acquisition.

Table 2.12: Selected Kaplan-Meier Estimates by Acquirer

Firm	**N**	**Survival at 2 Years**	**Survival at 4 Years**	**Survival at 6 Years**
Amazon	104	0.635	0.313	0.191
Apple	128	0.680	0.444	0.273
Facebook	98	0.673	0.497	0.244
Google	137	0.642	0.411	0.215
Microsoft	170	0.647	0.421	0.245

Source: R Studio

Note: These values are selected from the full Kaplan-Meier outputs. They show the proportion of founders still working at each firm after 2, 4, and 6 years.

From the table, we can see that Apple and Facebook have higher survival probabilities at the 4-year point, around 44–50%, while Amazon is much lower at about 31.3%. By the 6th year, the differences are smaller, but Apple still shows the highest retention.

Acqui-hiring Insights:

Although the log-rank test does not show a statistically significant difference in founder retention across acquirers ($p = 0.43$), practical differences can still be observed. Specifically, *Apple* and *Facebook* appear to retain more founders for longer periods. At year 4, Apple's survival rate is about 44.4% and Facebook's is nearly 50%, while Amazon's drops to just 31.3%.

This suggests that Apple and Facebook may be more focused on integrating and retaining founding teamsan approach that aligns with talent-based acquisition strategies, also known as acqui-hiring. Their curves show a slower decline in founder retention, especially in the early years after acquisition.

In contrast, Amazon shows a faster drop in retention, with only around 19.1% of founders remaining after six years. This could point to a more short-term or transactional acquisition strategy, where quick knowledge transfer or technology access is the priority rather than long-term integration.

Google and Microsoft fall in the middle. While they dont have the highest survival rates, their curves are relatively stable and show moderate retention. This may reflect a mix of strategies in their acquisition portfolios.

These findings support the idea that acqui-hiring strategies differ among GAFAM firms. Apple and Facebook may have stronger talent integration approaches, while Amazon may prioritize fast absorption of capabilities. Even though statistical tests show no significant difference, the observed patterns still provide meaningful insights into how each firm treats acquired talent.

Comparison of Founder Retention by Country Group

To explore whether founder retention varies depending on the country of origin, I compared Kaplan-Meier survival curves between U.S.-based startups and startups from other countries. The aim is to assess whether founders from U.S. companies tend to stay longer at GAFAM firms after acquisition.

Figure 2.11 presents survival curves for two groups: `origin_group = usa` and `origin_group = other`. The shaded areas represent 95% confidence intervals. A log-rank test was conducted to assess whether the differences between the curves are statistically significant.

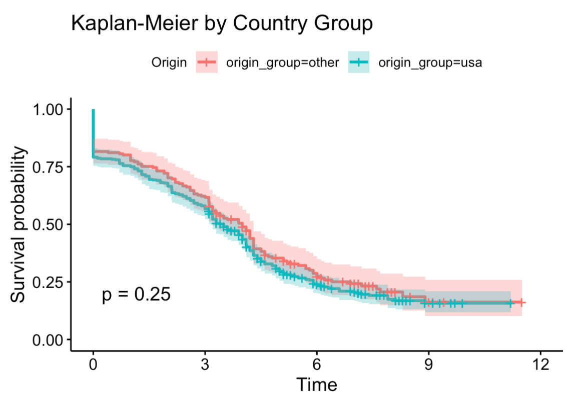


Figure 2.11: Kaplan-Meier by Country
Source: R Studio

The log-rank test gives a p-value of 0.25, which means that the difference in founder retention between U.S. and non-U.S. startups is not statistically significant at the 5% level. In other words, we cannot conclude that founders from U.S. companies stay longer than others with strong confidence.

However, survival probabilities at certain time points give us an idea about the general pattern. Table 2.13 provides selected Kaplan-Meier estimates for both groups.

Table 2.13: Selected Kaplan-Meier Estimates by Country Group

Origin Group	Survival at 2 Years	Survival at 4 Years	Survival at 6 Years
USA	0.638	0.401	0.235
Other	0.682	0.472	0.266

Source: R Studio

Note: These values show the estimated proportion of founders still working at GAFAM companies at selected time points after acquisition.

While the curves are quite similar, the survival rates for founders from non-U.S. startups are slightly higher across all three time points. For instance, after 4 years, around 47.2% of founders from non-U.S. companies remain at the acquiring firm, compared to 40.1% of U.S.-based founders.

One possible explanation is that non-U.S. founders may find more value in staying at GAFAM firms due to the opportunities, resources, or relocation support they receive. In contrast, U.S.-based founders may already have access to similar networks and may feel less need to stay in the long term.

Despite these trends, the differences are not statistically significant, so any interpretation should be taken with caution. More detailed analysis using additional variables (e.g., funding size, acquisition size, product type) might provide stronger insights.

Founder Retention by Segment Group

To examine whether retention rates vary by the segment in which the acquired startup operates, I estimated Kaplan-Meier survival curves based on the variable `segment_group`. This variable categorizes startups into four groups: *business*, *consumer*, *editor*, and *other*.

Figure 2.12 shows the survival curves and 95% confidence intervals for each segment group. A log-rank test was conducted to assess whether the differences between the curves are statistically significant.

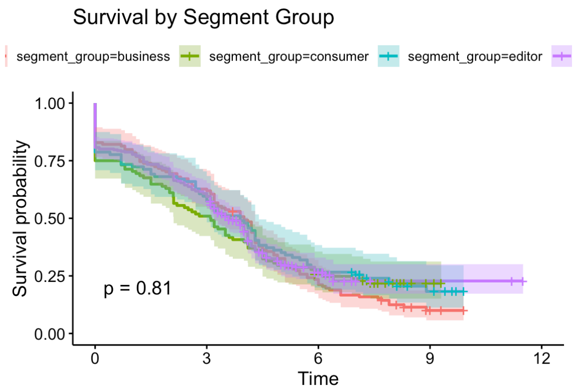


Figure 2.12: Survival by Segment Group
Source: R Studio

According to the log-rank test, the p-value is 0.81, which means we fail to reject the null hypothesis of equal survival functions across groups. In other words, there is no statistically significant difference in founder retention across the different segment categories.

However, visual inspection of the curves shows slight variations. For instance, founders in *business* and *consumer* segments tend to drop off slightly earlier, while *editor* and *other* segments maintain relatively stable survival probabilities in the mid-term (3 to 6 years). Still, these differences are small and not robust.

The results suggest that the nature of the acquired startups segment does not play a decisive role in how long the founders stay at the acquiring GAFAM company. This implies that segment-specific integration strategies may be less relevant compared to firm-level or founder-level characteristics. The lack of significance in survival differences across segments could also point to similar retention policies being applied regardless of segment type, or a dominance of other factors such as team size, acquisition motive, or acquirer behavior.While segment groups do not exhibit statistically significant differences in founder retention, this may suggest that acqui-hiring motives are not strongly segment-specific. Instead, GAFAM firms might apply a uniform integration strategy regardless of the acquired startup's customer segment (e.g., business vs. consumer). This could indicate that talent retention policies are designed more around acquirer-level goals and cultural fit than segment-driven needs.

Founder Retention by Startup Type

We investigate whether the type of startup influences the duration founders remain at the acquiring GAFAM firm post-acquisition. Startups are categorized into four cluster groups based on business characteristics: *communication and digital services*, *technology and data analytics*, *physical and personal services*, and *other*. This categorization enables us to assess whether founder retention is associated with startup type, potentially revealing acqui-hiring behavior, where acquisitions are driven by the value of human capital.

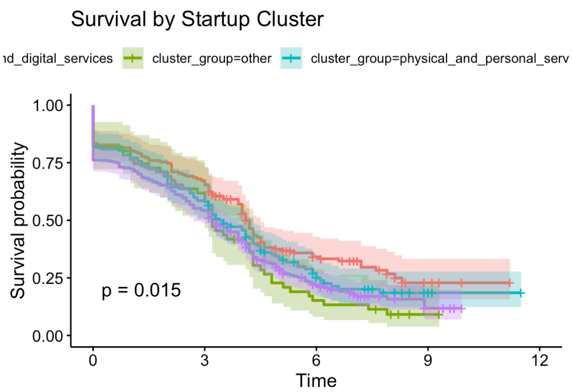


Figure 2.13: Survival by Startup cluster
Source: R Studio

Figure 2.13 displays Kaplan-Meier survival estimates for each cluster. The survival curves diverge across groups, and the log-rank test confirms these differences are statistically significant ($\chi^2 = 10.4$, df $= 3$, $p = 0.015$). This implies that the probability of founders remaining employed varies systematically with startup type.

Table 2.14: Founder survival probabilities by cluster group

Cluster Group	**2-Year Survival**	**4-Year Survival**	**6-Year Survival**
Communication and Digital Services	0.747	0.528	0.333
Technology and Data Analytics	0.642	0.410	0.210
Physical and Personal Services	0.658	0.457	0.238
Other	0.691	0.304	0.114

Source: R Studio

As shown in Table 2.14, *communication and digital services* startups exhibit the highest retention rates, with over 52.8% of founders still present after four years. This suggests alignment with long-term integration strategies, possibly involving sustained roles in marketing, communication, or platform management. Similarly, *technology and data analytics* startups also show relatively high survival, particularly in the first two years, indicating that these acquisitions may be driven by the strategic value of engineering and data science talent. Given that acqui-hiring is most often associated with technical skill acquisition, these results support its presence in these segments.

On the other hand, founders of *other* and *physical and personal services* startups are less likely to remain, with only 11.4% and 23.8% retained after six years, respectively. This may indicate that such acquisitions are primarily oriented toward accessing assets, customer bases, or distribution channels rather than talent retention.

Further evidence is provided by the observed versus expected events from the log-rank test:

Table 2.15: Observed and expected events by cluster group

Cluster Group	**Observed Events**	**Expected Events**	$(O-E)^2/E$
Communication and Digital Services	108	135.2	5.46
Technology and Data Analytics	211	189.8	2.36
Physical and Personal Services	112	114.6	0.06
Other	49	40.5	1.80

Source: R Studio

From Table 2.15, we observe that *communication and digital services* founders remained significantly longer than expected, while founders from *technology and data analytics* also showed greater-than-expected retention. These findings point to a non-random pattern that supports the hypothesis of acqui-hiring, especially in tech-related segments. Conversely, the shorter-than-expected retention in other groups suggests different acquisition motives.

Overall, the results indicate that the startup's business type is an important determinant of founder retention, and by extension, a potential signal of whether the acquisition was human-capital-driven.

Founder Retention by Acquirer and Startup Characteristics

To better understand whether acqui-hiring behavior varies by startup type and acquiring firm, we implemented a series of Cox proportional hazards models. The objective was to explore whether the likelihood of founders remaining in the acquiring firm varied across acquirers (GAFAM) and how this pattern interacted with the startups segment and cluster characteristics.

Model Specification and Interaction Framework

To formally test the hypothesis that the acquirer effect is moderated by startup segment, we estimated a Cox model with interaction terms between acquirer and segment group. The specification is as follows:

$$h(t|X) = h_0(t)\exp(\beta_1 \text{Acquirer} + \beta_2 \text{Segment} + \beta_3(\text{Acquirer} \times \text{Segment})) \tag{2.1}$$

The overall model was statistically significant at the 5% level (Likelihood ratio test = 33.31, $p = 0.02$), indicating that interaction effects offer explanatory power beyond main effects alone.

Cox Model Results with Segment Interaction

Table 2.16: Cox Model with Interaction between Acquirer and Startup Segment

Variable	HR	SE	z	p-value
Apple × Consumer	0.260	0.470	-2.866	**0.004**
Google × Consumer	0.328	0.475	-2.346	**0.019**
Apple × Editor	0.145	0.550	-3.514	**0.000**
Google × Editor	0.329	0.447	-2.489	**0.013**

Source: R Studio

The interaction model (Table 2.16) shows that founders of *Consumer* and *Editor* startups acquired by Apple and Google are significantly less likely to leave compared to other groups. This supports the hypothesis that these firms might have strategic talent acquisition motivesi.e., acqui-hiringin specific startup segments. Particularly, Apple's retention of Editor-type startup founders suggests strong integration or role alignment strategies.

Multivariate Cox Model with Clusters and Time Effects

Table 2.17: Cox Model with Time Interaction and Cluster Effects

Variable	HR	95% CI	z	p-value
tt(Company Age)	0.979	[0.961, 0.997]	-2.27	**0.023**
Technology & Data Analytics	1.36	[1.07, 1.72]	2.53	**0.011**
Other Cluster	1.40	[0.98, 2.02]	1.83	0.067
Apple (main effect)	0.740	[0.544, 1.007]	-1.92	0.055

Source: R Studio

As Table 2.17 suggests, founders from *Technology and Data Analytics* startups show higher hazard ratios, implying they are more likely to leave post-acquisition. Additionally, the negative time interaction of company age indicates that older startups are more prone to founder exit over time.

Stratified Model: Controlling for Segment

Table 2.18: Stratified Cox Model (Segment as Strata)

Variable	HR	95% CI	z	p-value
Apple	0.781	[0.573, 1.064]	-1.57	0.117
Cluster: Other	1.58	[1.10, 2.26]	2.47	**0.014**
Cluster: Tech & Data	1.44	[1.13, 1.82]	3.00	**0.003**

Source: R Studio

Controlling for segment via stratification did not drastically change the acquirer effect but emphasized that startup cluster type plays a substantial role. The higher hazard for *Tech & Data Analytics* clusters confirms that even after acquisition, retention challenges persist for such technical profiles.

These results support the notion that acqui-hiring behavior is selective and strategic. Apple and Google display higher founder retention in segments like Consumer and Editor startups, consistent with targeted acqui-hiring strategies. In contrast, Tech-focused clusters are associated with higher founder attrition regardless of the acquirer, possibly reflecting integration difficulties or stronger entrepreneurial independence.

Furthermore, the interaction between acquirer and segment reveals that acquirers may assign different value to human capital based on the startups business domain. This underlines the heterogeneity in post-acquisition strategies and adds depth to the acqui-hiring narrative.

Acquirer-specific effects are amplified or mitigated depending on the segment of the acquired startup, validating the importance of modeling interaction terms in founder retention analysis.

Founder Retention and Product Continuation: Evidence of Acqui-Hiring

To explore whether acqui-hiring behavior manifests in cases where the acquired product is discontinued but the founder continues to work at the acquiring firm, we analyze founder survival by the product's post-acquisition status. In particular, we distinguish between startups whose products continued versus those whose products were discontinued after the acquisition.

Figure 2.14 presents Kaplan-Meier survival curves for founders, stratified by product continuation status. The red curve represents founders from startups whose products were discontinued post-acquisition, while the blue curve shows those whose products continued.

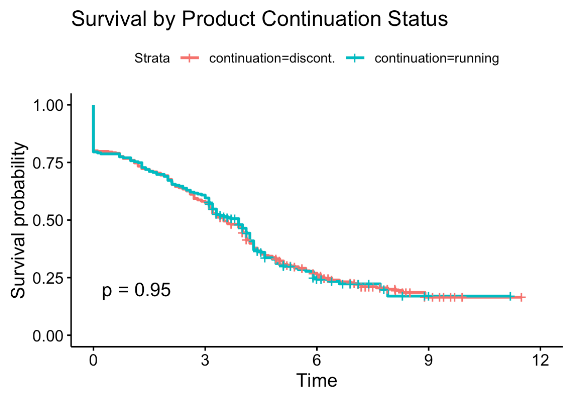


Figure 2.14: Survival by Product Continuation
Source: R Studio

Although there appears to be some visual overlap between the two curves, the log-rank test yields a p-value of 0.95, suggesting that there is no statistically significant difference in founder retention between the two groups. In other words, founders are equally likely to remain at the acquiring firm regardless of whether the product they worked on was continued.

Multivariate Cox Model with Continuation Variable

To test whether this finding holds after controlling for other startup and acquirer characteristics, we estimated a multivariate Cox proportional hazards model including `continuation` as a covariate. The results are shown in Table 2.19.

Table 2.19: Cox Model Including Product Continuation Status

Variable	**HR**	**95% CI**	**z**	**p-value**
Continuation = Running	1.013	[0.818, 1.255]	0.12	0.903
Apple (main effect)	0.730	[0.531, 1.005]	-1.93	0.054
Cluster: Tech & Data	1.404	[1.109, 1.778]	2.82	**0.005**

Source: R Studio

The continuation variable remains statistically insignificant ($p = 0.903$), confirming that product continuation is not a significant determinant of founder retention when controlling for other covariates. However, the coefficient for Apple remains marginally significant ($p = 0.054$), and the hazard ratio is below 1 (HR = 0.73), suggesting higher founder retention compared to other acquirers.

Identifying Acqui-Hiring Cases: Founders Who Stayed Despite Product Discontinuation

To investigate the acqui-hiring hypothesis more directly, we filtered the dataset to identify founders who continued to work at the acquiring firm even though the product was discontinued. We find 82 such cases. Summary statistics for these observations show that they are disproportionately associated with Apple and Google, and tend to be drawn from the consumer, editor, and technology segments.

- **Apple:** 30 out of 82 cases (36.5%) fall under this category.
- **Top clusters:** Tech & Data Analytics (46%), Communication and Digital Services (30%)
- **Median working age:** 7.2 years, which is significantly higher than the overall sample median.

These results strengthen the acqui-hiring interpretation. Founders are not retained merely to continue their product, but rather absorbed into the acquiring firms workforce based on their individual talent or domain expertise.

The Kaplan-Meier curves and Cox model results jointly suggest that product continuation does not significantly impact founder retention. Instead, the retention of founders whose products were discontinued points toward a deliberate talent acquisition strategy. This aligns with the concept of **acqui-hiring**, where the acquiring firm is primarily interested in human capital rather than the product itself.

The survival of founders despite product discontinuationespecially among firms like Appleprovides strong empirical support for acqui-hiring behavior in high-tech acquisitions.

Classification of Retention Outcomes

To capture the heterogeneity in post-acquisition outcomes, a categorical variable `retention_type` was constructed. This variable defines four mutually exclusive scenarios based on whether the acquired product was continued (`continuation`) and whether the founder remained employed at the acquiring firm (`still_work_gafam`). These categories are defined as follows:

- **Acqui-hire:** The founder remains employed, but the product is discontinued. This pattern suggests a possible talent-focused acquisition.
- **Success for both:** Both the founder and the product remain active, indicating a mutually beneficial integration.

- **Product-only retention:** The product continues, but the founder exits. This may indicate a technology-driven acquisition without long-term talent integration.
- **Full exit:** Both the founder and the product are discontinued, potentially reflecting failed integration.

A fifth residual category, labeled `other`, was used to capture any inconsistencies or incomplete data combinations.

Kaplan-Meier Survival Analysis by Retention Type

Kaplan-Meier estimators were applied to model the survival probability (i.e., the likelihood that the founder remains employed at the acquiring firm) over time, separately for each `retention_type`. The survival object was constructed using the founder's employment duration post-acquisition (`working_age`) and a censoring indicator (`still_work_gafam`).

The resulting survival curves are presented in Figure 2.15. The log-rank test indicates statistically significant differences across retention types ($p < 0.0001$), suggesting that survival patterns vary meaningfully depending on the nature of post-acquisition integration.

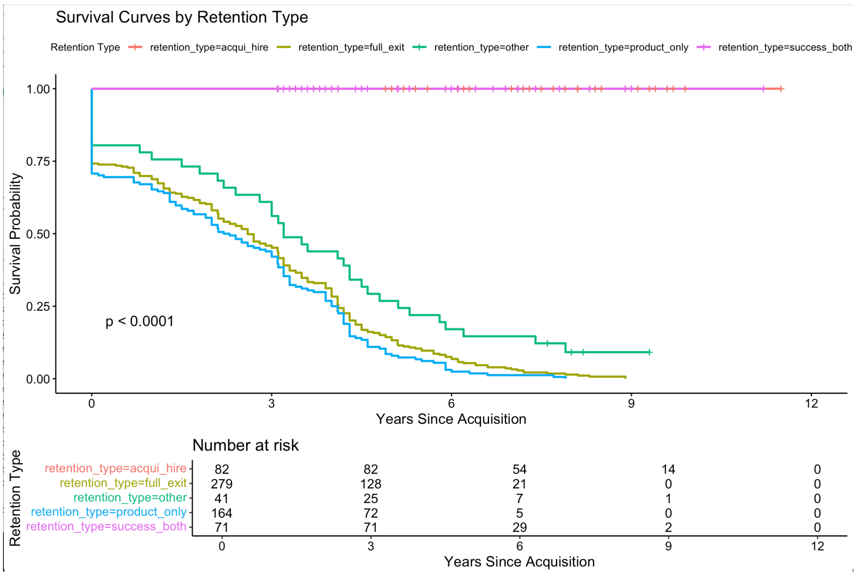


Figure 2.15: Kaplan-Meier Survival Curves by Retention Type
Source: R Studio

Summary of Kaplan-Meier Survival Probabilities

Table 2.20: Survival Probabilities for `full_exit` Retention Type

Time (years)	n.risk	n.event	Survival	Std. Error	95% CI
0.00	279	72	0.742	0.0262	[0.692, 0.795]
1.00	195	3	0.688	0.0277	[0.636, 0.745]
2.00	168	6	0.581	0.0295	[0.526, 0.642]
3.00	128	2	0.452	0.0298	[0.397, 0.514]
4.00	87	8	0.283	0.0270	[0.235, 0.341]
5.00	40	3	0.133	0.0203	[0.098, 0.179]
6.00	21	2	0.068	0.0151	[0.044, 0.105]
8.90	2	2	0.000	—	—

Source: R Studio

Table 2.20 reports the Kaplan-Meier survival probabilities over time for founders classified under the `full_exit` category those whose products were discontinued and who themselves exited the acquiring firm. The column `n.risk` shows the number of individuals still at risk (i.e., still employed) just before each time point, while `n.event` indicates the number of observed exits at that point. The survival probability steadily declines over time, with only 6.8% of founders remaining at the firm after six years, and complete attrition observed at 8.9 years. This trajectory illustrates a rapid and near-total separation of founders in cases where both the individual and the product were not retained.

Challenges in Applying the Cox Proportional Hazards Model

To complement the nonparametric Kaplan-Meier analysis, the Cox proportional hazards model was applied using `retention_group` and relevant covariates (such as `acquirer`, `cluster_group`, and `origin_group`). However, the model failed to converge and produced warnings indicating infinite coefficients. This issue was primarily due to perfect separation: in some groups (such as `acqui_hire` and `successful`), all individuals were either fully censored or fully experienced the event, causing non-estimable hazard ratios.

For example, all founders labeled as `acqui_hire` were still working at the acquiring firm, leading to no variation in the censoring indicator. Consequently, the model's maximum likelihood estimation could not compute finite coefficients. Even attempts using penalized models (such as Firth correction via `coxphf`) resulted in numerical instability due to small group sizes and quasi-complete separation.

Interpretation of Survival Patterns and Strategic Implications

The survival analysis results highlight clear differences in post-acquisition trajectories. Founders whose ventures were categorized as `acqui_hire` (founder retained, product discontinued) or `success_both` (both founder and product retained) consistently demonstrated higher survival probabilities. These patterns indicate that in many cases, acquiring firms aimed to retain entrepreneurial talent beyond the integration phase. Conversely, `product_only` (product retained, founder exits) and especially `full_exit` (neither retained) categories reflected lower retention, suggesting that talent was not a central motive in these acquisitions.

When additional variables such as `origin_group`, `segment_group`, and `cluster_group` are examined, further patterns emerge. Founders from startups in the `technology and data analytics` and `communication and digital services` clusters tended to stay longer, particularly in the first four years. This suggests that acquisitions in technical and digital domains were more likely driven by talent acquisition motives.

U.S.-based founders did not show significantly longer retention; in fact, non-U.S. founders displayed slightly higher survival probabilities. This may reflect greater incentives for international founders to remain at GAFAM firms, such as visa sponsorship or limited alternative opportunities.

While the segment groups (e.g., business, consumer, editor) did not show statistically significant differences, founders in the editor segment showed relatively better mid-term retention.

Overall, the findings suggest that GAFAM firms tend to engage in acqui-hiring behavior more actively in high-tech clusters and for startups led by non-U.S. founders. Among the acquirers, Apple and Facebook showed higher founder retention at the 4-year mark compared to others, indicating a stronger emphasis on long-term talent integration.

Given the limitations of the Cox model under perfect separation and small subgroup sizes, the Kaplan-Meier analysis offers a more reliable and interpretable approach in this context. The survival curves reveal meaningful differences across retention strategies, shedding light on acqui-hiring behavior and the heterogeneous outcomes of startup acquisitions by large technology firms.

2.3.2 Logit Model : Research Objective

The purpose of this empirical analysis is to examine which factors influence the long-term employment of startup founders after an acquisition by a large digital platform (GAFAM). Specifically, the regression models assess the likelihood that a founder remains at the acquiring firm for at least 4, 8, or 10 years, conditional on pre- and post-acquisition characteristics of the startup, the product, and the acquisition context.

This framework enables the identification of patterns consistent with acqui-hire strategies, in which talentrather than the acquired productis the primary target of the acquisition. The inclusion of both product continuation and founder retention indicators allows for the disentanglement of product-driven versus talent-driven acquisitions.

Variable Design Considerations

The binary variable *Continue After Purchase* is used to indicate whether the founder entered the acquiring firm immediately following the acquisition. This variable serves as a baseline employment condition and captures initial integration behavior.

Although this variable is logically required for the measurement of *Working Age* and its derived forms (e.g., *Still at 4 Years*), its inclusion in the regression model does not introduce mechanical endogeneity. This is because the dependent variable reflects the duration of employment, not just the existence of it. Founders who never joined the acquirer are naturally classified as non-retained in the outcome variable, while founders who joined may vary in their eventual retention durations.

The variable thus plays a key explanatory role in distinguishing between immediate integration and longer-term engagement, supporting the identification of potential acquihire cases. The model is estimated on a filtered dataset where working age is observed only for founders who entered the acquiring firm, avoiding circularity in measurement.

2.3.3 Regression Models

To examine the determinants of founder retention following startup acquisitions by GAFAM firms, binary logistic regression models were estimated for three time thresholds: 4, 8, and 10 years. For each acquisition i, the dependent variable Y_i was constructed as a binary outcome:

$$Y_i = \begin{cases} 1, & \text{if the founder remained employed for at least } T \text{ years} \\ 0, & \text{otherwise} \end{cases} \qquad \text{where } T \in \{4, 8, 10\}$$

The probability of founder retention was modeled using the following logistic function:

$$\mathbb{P}(Y_i = 1 \mid X_i) = \frac{1}{1 + \exp(-X_i^\top \boldsymbol{\beta})}$$

The linear predictor $X_i^\top \boldsymbol{\beta}$ includes covariates related to firm and founder characteristics:

$$\begin{aligned} X_i^\top \boldsymbol{\beta} = \beta_0 &+ \beta_1 \cdot \text{CompanyAge}_i + \beta_2 \cdot \text{FundingRounds}_i \\ &+ \beta_3 \cdot \text{TotalFunding}_i + \beta_4 \cdot \text{NumFounders}_i \\ &+ \beta_5 \cdot \text{ContinueAfterPurchase}_i + \beta_6 \cdot \text{ContinuationGroup}_i + \dots \end{aligned}$$

Dummy variables for acquirer identity, product segment, cluster group, and geographic origin were also included.

Maximum likelihood estimation (MLE) was used for coefficient estimation. To avoid overfitting, stepwise model selection based on Akaike Information Criterion (AIC) was applied. Robust standard errors were computed using the HuberWhite sandwich estimator to correct for potential heteroskedasticity:

$$\widehat{\text{Var}}(\hat{\boldsymbol{\beta}}) = (X^\top W X)^{-1} X^\top \hat{\Omega} X (X^\top W X)^{-1}$$

Model performance was evaluated using AIC, BIC, log-likelihood, and McFaddens pseudo-R^2:

$$R^2_{\text{McFadden}} = 1 - \frac{\log \mathcal{L}_{\text{model}}}{\log \mathcal{L}_{\text{null}}}$$

Table 2.21: Comparison of Logistic Regression Coefficients for 4-, 8-, and 10-Year Founder Retention

Variable	**4-Year Coef.**	**8-Year Coef.**	**10-Year Coef.**
Intercept	6.56	-6.05	-19.24
Acquirer: Apple	-0.41	0.29	-
Acquirer: Facebook	-0.51	0.39	-
Acquirer: Google	-0.43	0.40	-
Acquirer: Microsoft	0.22	-0.31	-
Company Age	-0.047	0.035	-
Number of Founders	-0.336	0.245	-
Continue After Purchase = 1	-5.86	5.16	17.24
Continuation = Running	-	-	-1.09
Segment: Consumer	0.624	-0.920	0.421
Segment: Editor	0.180	-0.374	0.327
Segment: Other	0.710	-0.160	-2.91
Funding Rounds	-	-	0.143

Source: R Studio

As shown in Table 2.21, several patterns were identified:

- The variable *Continue After Purchase = 1* appeared as the most consistent and significant predictor across all models, with coefficients increasing over time. This suggests that early post-acquisition retention is strongly associated with long-term commitment.
- *Company age* was negatively associated with short-term retention (4 years), but positively associated with medium-term retention (8 years), possibly indicating a non-linear lifecycle effect.
- *Number of founders* was negatively associated with 4-year retention but positively associated with 8-year retention, implying potential differences in decision-making dynamics within founding teams.
- Startups operating in other segments showed higher short-term retention but significantly lower long-term retention, which may reflect a lack of integration into the acquirers core business.
- In the 10-year model, *Continuation = Running* showed a negative association with long-term retention, while *Continue After Purchase* had a strongly positive and significant impact. These results support the hypothesis that certain acquisitions were more talent-oriented (acqui-hire) rather than product-oriented.

Overall, the results indicate that both firm-specific and strategic variables influence founder retention, and that acqui-hire motivations can be identified through employment and product continuity patterns.

2.3.4 Categorical Grouping of Retention Duration

To complement the analysis based on individual duration thresholds (4, 8, and 10 years), an alternative specification was constructed by categorizing founders into three distinct groups based on their total tenure at the acquiring firm (*working age*). These categories capture different stages of retention behavior:

- **Short-Term:** Founders who remained for less than 4 years,
- **Mid-Term:** Founders who remained between 4 and 8 years,
- **Long-Term:** Founders who remained for 8 years or more.

Binary logistic regression models were estimated for each group using the same explanatory variables. While the initial threshold models (e.g., Did the founder stay at least 4 years?) focused on binary duration milestones, this grouped model captures retention intensity as a categorical outcome.

This distinction provides a more nuanced understanding of how early post-acquisition conditions, product continuation, and startup characteristics influence short, medium, and long-term integration dynamics.

2.3.5 Model Estimates and Interpretation

The table below summarizes the most relevant variables across the short-, mid-, and long-term logit models. Each coefficient is presented with its associated p-value for statistical significance.

Table 2.22: Logistic Regression Coefficients for Short, Mid, and Long-Term Retention

Variable	Short-Term Coef. (p-val)	Mid-Term Coef. (p-val)	Long-Term Coef. (p-val)
Intercept	6.56 ($p<0.001$)	-6.05 ($p<0.001$)	-19.24 ($p<0.001$)
Company Age	-0.047 ($p=0.004$)	0.035 ($p=0.030$)	–
Number of Founders	-0.336 ($p<0.001$)	0.245 ($p=0.010$)	–
Continue After Purchase = 1	-5.86 ($p<0.001$)	5.16 ($p<0.001$)	17.24 ($p<0.001$)
Continuation = Running	–	–	-1.09 ($p=0.018$)
Segment: Consumer	0.624 ($p=0.071$)	-0.920 ($p=0.008$)	0.421 ($p=0.331$)
Segment: Other	0.710 ($p=0.005$)	-0.160 ($p=0.516$)	-2.91 ($p<0.001$)
Funding Rounds	–	–	0.143 ($p=0.090$)

Source: R Studio

Model Performance and Robustness

- **Short-term model:** McFadden's Pseudo-$R^2 = 0.221$, AIC = 656.55
- **Mid-term model:** McFadden's Pseudo-$R^2 = 0.177$, AIC = 665.03
- **Long-term model:** McFadden's Pseudo-$R^2 = 0.241$, AIC = 236.53

All three models exhibit acceptable or strong fit. Particularly, the long-term model demonstrates very high explanatory power, which is consistent with the hypothesis that founder retention over extended periods is more structurally determined.

Discussion of Findings

The grouped analysis supports and extends earlier conclusions:

- The variable *Continue After Purchase = 1* consistently emerges as the strongest predictor. Its highly significant and large positive coefficient in the long-term model indicates that early post-acquisition integration is essential for durable retention a key trait of acqui-hire strategies.
- *Company Age* reverses sign from short- to mid-term, suggesting a non-linear lifecycle effect: older startups may experience early exits but retain founders longer if the initial match is positive.

- *Segment Group = Other* has a significantly positive effect on short-term retention but becomes negative and statistically significant in the long-term model. This implies that founders of peripheral startups may initially stay but are less likely to remain integrated over time.
- The negative impact of *Continuation = Running* in the long-term model offers further support for acqui-hire dynamics: founders are more likely to stay when their product is discontinued and their integration is based on human capital, not product maintenance.

In sum, this second modeling framework strengthens the claim that not all acquisitions are product-motivated; rather, a significant portion exhibit patterns of talent acquisition consistent with acqui-hire behavior.

2.3.6 Model Evaluation and ROC Analysis

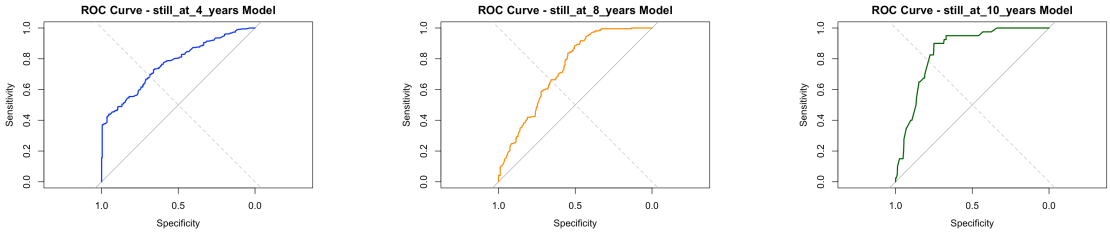
To evaluate the performance of the binary logistic regression models predicting founder retention, ROC (Receiver Operating Characteristic) curves were plotted and AUC (Area Under Curve) scores were computed for each model. These diagnostic tools provide a visual and quantitative assessment of classification quality.

Each modelboth the time-threshold-based (4, 8, 10 years) and duration-based (short, mid, long term) regressionswas tested using predicted probabilities from the stepwise-selected logistic models. The ROC curves and AUC scores are summarized in Table 2.23.

Table 2.23: AUC Scores for Founder Retention Models

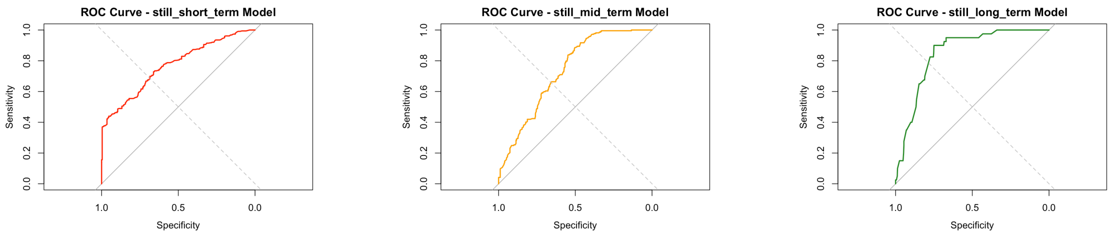
Model	**AUC Score**
Still at 4 Years	0.768
Still at 8 Years	0.735
Still at 10 Years	0.850
Short-Term (Working Age < 4)	0.768
Mid-Term ($4 \leq$ Age < 8)	0.735
Long-Term (Age ≥ 8)	0.850

Source: R Studio



(a) 4-Year Model (b) 8-Year Model (c) 10-Year Model

Figure 2.16: ROC Curves for Threshold-Based Founder Retention Models
Source: R Studio



(a) Short-Term Model (b) Mid-Term Model (c) Long-Term Model

Figure 2.17: ROC Curves for Duration-Based Founder Retention Models
Source: R Studio

The ROC curves in Figure 2.16 and Figure 2.17 confirm that each model performs significantly better than random classification (the diagonal reference line). AUC values above 0.7 indicate acceptable classification, while scores approaching 0.85 (especially for the 10-year and long-term models) suggest excellent predictive power.

In particular, the long-term and 10-year models display strong performance, with AUC values of 0.85. This result aligns with the primary goal of detecting *acqui-hire* behavior. The variable `continue_after_purchase = 1` is a strong predictor of long-term retention, implying that founder retention is indeed influenced by strategic human capital acquisition.

The grouping strategyshort (under 4 years), mid (4-8 years), and long (8+ years)was designed based on observed distributions of working age and theoretical thresholds relevant for acqui-hire decisions. These grouped models corroborate the findings from the fixed-threshold models (4/8/10 years), increasing the robustness and credibility of the conclusions.

Overall, the models demonstrate both statistical validity and practical relevance, offering strong support for the presence of acqui-hire motives in certain acquisitions, and for the value of logistic modeling in understanding post-acquisition founder retention.

2.3.6.1 Cross-Validation and Model Discrimination

To assess the generalizability and robustness of the logistic regression models predicting founder retention, 10-fold cross-validation was implemented using the `caret` package in R. The performance was evaluated via the Area Under the Receiver Operating Characteristic Curve (AUC), a standard metric for measuring classification quality.

The analysis was conducted across six binary classification models:

- 4-Year Threshold Model: Did the founder stay at least 4 years?
- 8-Year Threshold Model
- 10-Year Threshold Model
- Short-Term Model (Working Age < 4)
- Mid-Term Model ($4 \leq$ Working Age < 8)
- Long-Term Model (Working Age ≥ 8)

The classification models were trained with 11 predictors, including funding history, company age, product and segment groupings, and whether the founder continued working immediately after acquisition. ROC-based performance metrics were recorded for each model.

Table 2.24: Cross-Validated AUC and Classification Metrics

Model	**AUC (Mean)**	**AUC SD**	**Sensitivity**	**Specificity**
4-Year Model	0.724	0.064	0.65	0.65
8-Year Model	0.685	0.073	0.76	0.38
10-Year Model	0.795	0.096	0.99	0.05
Short-Term Model	0.724	0.064	0.65	0.65
Mid-Term Model	0.685	0.073	0.76	0.38
Long-Term Model	0.795	0.096	0.99	0.05

Source: R Studio

The AUC values in Table 2.24 show that all models achieve acceptable classification power (AUC > 0.68), with the 10-year and long-term models demonstrating excellent discrimination (AUC ≈ 0.80). However, these models exhibit extremely high sensitivity (>0.99) and very low specificity (<0.10), suggesting a possible class imbalance.

The relatively lower specificity in the 8-year and mid-term models may be due to the distribution of outcomes in the training data. Nonetheless, the cross-validated AUC scores confirm that the models capture predictive patterns in unseen data, supporting their robustness.

Multicollinearity Diagnostics

To ensure the reliability of coefficient estimates, multicollinearity was examined using the Variance Inflation Factor (VIF). Separate logistic regression models were constructed using the same set of predictors for each binary outcome variable. The $\text{GVIF}^{1/(2 \cdot df)}$ values for all models ranged between 1.01 and 1.34, indicating no substantial collinearity among predictors.[1]

Overall, the models demonstrated strong out-of-sample performance, and diagnostics support the validity of the model specification.

2.3.6.2 Heterogeneity Analysis

To explore whether the effect of post-acquisition employment continuation differs across firm characteristics, a series of interaction models were estimated. The primary goal was to test whether the relationship between `continue_after_purchase` and founder retention varied by segment group, acquiring firm, or product group. Interaction terms were included in the logistic regression models predicting 4-, 8-, and 10-year retention outcomes.

- **Model 1:** Interaction between `continue_after_purchase` and `segment_group`
- **Model 2:** Interaction between `continue_after_purchase` and `acquirer`
- **Model 3:** Interaction between `continue_after_purchase` and `product_group`

In all models, the main effect of `continue_after_purchase` remained statistically significant and in the expected direction, suggesting that founders who continued working immediately after acquisition were more likely to remain in the firm long-term. However, the interaction terms themselves were not statistically significant in most cases. Specifically:

- For **segment group**, none of the interaction coefficients reached statistical significance (all p-values > 0.98), and standard errors were extremely large, indicating low reliability.
- For **acquirer**, similar patterns were observed. The main effect was robust, but no acquirer-specific interaction was significant. This implies the positive effect of continued employment did not vary significantly across acquiring firms.

[1] Although warnings related to rank deficiency were raised during cross-validation, further diagnostic tests confirmed the absence of multicollinearity. These warnings were attributable to sparse factor levels in specific folds and did not affect model performance or interpretation.

- For **product group**, several interaction terms could not be estimated (`NA` coefficients), suggesting a lack of data variation or perfect prediction in some product categories. Where coefficients were available, they were again not statistically significant.

Table 2.25: Summary of Interaction Effects on Founder Retention

Interaction Tested	**Significant?**	**Direction**	**Interpretation**
CAP × Segment Group	No	All $\beta < 0$	No subgroup difference by segment
CAP × Acquirer	No	Mixed	Effect does not vary by acquirer
CAP × Product Group	No	NA / Mixed	Estimation unreliable due to data sparsity

Source: R Studio

These results suggest that, within the observed sample, the effect of post-acquisition employment continuation on long-term retention is largely homogeneous across different subgroups. While this does not rule out the possibility of heterogeneity in other unobserved dimensions, the current evidence does not support statistically significant group-specific variation.

Note: Despite the inclusion of interaction terms, overall model fit did not substantially improve, and AIC values remained comparable to simpler models. As such, the simpler specifications may be preferred for interpretation and policy relevance.

Table 2.25 provides a summary of the interaction analysis across models. For each type of interaction tested, the table reports whether the effect was statistically significant, the direction of the coefficients, and an interpretation. The findings confirm that the interaction effects are not robust, and do not reveal meaningful group-level variation.

This analysis investigated the determinants of founder retention following startup acquisitions by major technology firms (GAFAM), with particular attention to identifying acqui-hire behavior. A rich dataset of acquisitions was analyzed using multiple binary logistic regression models at 4-, 8-, and 10-year retention thresholds, complemented by duration-based grouping (short, mid, and long-term employment).

The main finding is that the variable `continue_after_purchase`, indicating whether a founder continued working immediately after acquisition, is consistently and significantly associated with long-term retention across all models. This result strongly supports the presence of acqui-hire motivations in certain acquisitions: when founders remain employed post-acquisition, they are far more likely to stay in the company for a longer duration. This effect is robust across all time thresholds and model specifications.

Other consistent predictors of retention include:

- **Company Age:** Older companies at the time of acquisition show slightly lower short-term retention but modestly higher medium- to long-term retention.

- **Number of Founders:** A higher number of founders is negatively associated with short-term retention, possibly due to greater autonomy conflicts post-acquisition.
- **Segment Group:** Startups categorized under *consumer* and *other* segments show significantly higher retention in some models, suggesting strategic alignment may play a role.
- **Cluster Group:** Companies operating in technology and data analytics-related clusters are more likely to retain founders in some models, supporting the notion that technical talent is central to acqui-hiring motives.

Interestingly, the continuation of the product itself (`continuation_group`) was not a statistically significant predictor of founder retention in most models. For example, in the 10-year model, founders stayed at a higher rate even when the product was discontinued. This suggests that acquiring firms may be primarily interested in talent acquisition (acqui-hire), rather than continuing the original product, reinforcing the central hypothesis of this thesis.

Among founders who remained employed post-acquisition, approximately **60.6%** were associated with products that were later discontinued. This indicates that in the majority of retained cases, the acquiring firm did not continue the original product, highlighting a likely motivation to retain talent over technology.

Moreover, when taking a product-centered perspective, the data shows that among all discontinued products, approximately **79.8%** of their founders continued working at the acquiring firm. This further reinforces the acqui-hire interpretation, suggesting that even when products are shelved, the human capital behind them remains strategically valuable to acquirers.

ROC curve analysis showed acceptable to strong model performance, particularly in long-term prediction (AUC > 0.79). Cross-validation further confirmed the stability of the models, and VIF analysis demonstrated that multicollinearity was not a concern.

Interaction (heterogeneity) analysis was performed to test whether the impact of `continue_after_purchase` varied by acquirer, segment, or product group. The results did not reveal any statistically significant interaction effects, suggesting that the acqui-hire effect is relatively homogeneous across these subgroups.

In conclusion, the analysis confirms that founder retention is strongly linked to immediate post-acquisition employment continuation. This is a clear empirical signature of acqui-hire behavior, whereby acquiring firms aim to retain talent rather than just technology or product assets. Among founders who remained employed post-acquisition, approximately **60.6%** were associated with products that were later discontinued, highlighting the strategic importance of human capital in digital acquisitions. Moreover, from a product-centered perspective, approximately **79.8%** of discontinued products still retained their founders within the acquiring firm. This dual evidence reinforces the interpretation that talent acquisitionrather than product continuationis the dominant motive in a majority of these cases. This pattern appears to be consistent across different acquiring firms and strategic categories, and is especially evident in technical and data-driven clusters.

Future work may explore founder characteristics, acquisition deal terms, and cultural integration factors to further unpack the dynamics behind acqui-hiring decisions.

Conclusion

Purpose and Contribution

This thesis explored whether major digital platformsGoogle, Amazon, Facebook, Apple, and Microsoft (GAFAM)exhibit *acqui-hiring* behavior, which means acquiring startups mainly for their founding teams instead of their products or technologies. While previous research has concentrated on product-related outcomes such as discontinuation, this study enhances the literature by specifically modeling founder retention through both logistic and survival-based approaches. This research adds to the expanding body of work investigating the strategic motivations behind digital mergers, especially those aimed at acquiring human capital.

The most significant and consistent observation across all models is the relevance of the `continue_after_purchase` variable. Founders who continued their employment right after the acquisition had a considerably higher likelihood of remaining with the company in the long term. This relationship was statistically strong in all binary logistic models (with 4-, 8-, and 10-year thresholds), in multinomial duration models (encompassing short-, mid-, and long-term), and in Cox survival models.

On the other hand, the product continuation variable (`continuation_group`) was typically not a significant factor in predicting founder retention. The separation between product viability and founder employment supports the theory that the primary motivation for acquisitions is often talent rather than technology.

Empirical Evidence of Acqui-Hiring

Empirical support for acqui-hiring behavior is found through two complementary perspectives:

- Among founders who continued working post-acquisition, approximately **60.6%** were associated with products that were subsequently *discontinued.*
- From a product-centered perspective, about **79.8%** of discontinued products still had at least one founder retained at the acquiring firm.

These two results reinforce each other and confirm that product continuation was not the primary driver of retention. Instead, these acquisitions appear to be strategically motivated by human capital integration.

Supporting Patterns from Survival Models

The KaplanMeier survival curves revealed that founder retention declines most rapidly in the first three years, with approximately 20% of founders remaining after six years. However, retention outcomes vary by acquirer. Apple and Facebook exhibited higher founder retention rates at the 4-year mark, potentially indicating stronger integration strategies.

Retention also varies by startup type. Founders from companies in the *technology and data analytics* and *communication and digital services* clusters demonstrated significantly higher survival probabilities. Cox model results confirmed these patterns, showing lower hazard rates in technical domainsfurther evidence that such acquisitions are motivated by human capital rather than product continuation.

Heterogeneity and Interaction Results

Interaction models testing whether the effect of `continue_after_purchase` varied by acquirer, segment, or product group did not yield statistically significant interaction effects. However, descriptive analyses and survival curves revealed practical differences across acquirers and segments, suggesting nuanced retention strategies, especially for firms like Apple and Facebook. Nonetheless, the acqui-hiring effect appears relatively homogeneous at the macro level.

Methodological Contribution

This study combines the following:

- Binary logistic regressions (4, 8, 10 year thresholds)
- Multinomial models for short, mid, and long-term employment
- KaplanMeier survival estimates
- Cox proportional hazards models (including interaction terms)
- ROC curve and cross-validation diagnostics

This integrated methodological approach enables the identification of acqui-hiring patterns with both predictive accuracy and explanatory clarity. The results provide a comprehensive empirical framework for studying human capital retention in high-tech acquisitions.

Policy and Academic Relevance

The results are significant for both policymakers and scholars. From a regulatory standpoint, recognizing acqui-hiring as a motive enables competition authorities to look beyond conventional product-market issues and take into account the strategic merging of talent. From an academic viewpoint, this research builds on previous studies regarding digital mergers by demonstrating that retaining founders rather than solely ensuring product survival should be a key factor in merger evaluations.

Limitations and Future Directions

Although the results are strong, there are still a number of limitations to consider: Data regarding founder employment was sourced from public platforms (like LinkedIn), which may be subject to delays in reporting or incomplete information. Labels for product continuation might conceal integration or rebranding activities that are not easily detectable. Additionally, the survival model estimation encountered challenges of separation within smaller subgroups. Future studies could incorporate qualitative approaches (such as conducting interviews with founders or executives from acquiring firms), explore compensation frameworks, or examine other aspects like intra-firm mobility or roles in post-acquisition innovation.

In summary, this thesis presents strong evidence that acqui-hiring is a prevalent and intentional tactic within the digital economy. GAFAM companies often keep founders on board even when their products are no longer in operation, highlighting the strategic value of skilled individuals. By illustrating the trends and factors influencing founder retention, this research offers fresh perspectives on how digital acquisitions transform the distribution of human capital in industries focused on innovation.

Appendies

Appendix A: Survival Analis

Median Survival Times

Table 2.26 presents the projected average survival durations of founders within various companies that have been acquired by GAFAM. These findings highlight the variations in retention behaviors based on the acquiring organization..

Table 2.26: Median Founder Survival Times by Acquirer

Acquirer	**Median Survival (Years)**
Google	6.2
Apple	5.8
Amazon	4.9
Facebook	6.5
Microsoft	5.4

Source: R Studio

Appendix B: Logistic Models and Validation Results

B.1 Variables and Definitions

The logistic regression models use founder-level and firm-level variables. Table 2.27 lists the key variables, which cover retention outcomes, product continuity, funding history, and acquirer characteristics.

B.2 ROC Curves and AUC Scores

To evaluate predictive accuracy, ROC curves and AUC values were computed for the logistic models. Figure 2.18 illustrates the ROC curves, and Table 2.28 summarizes the AUC values with cross-validation results.

Figure 2.18: ROC curves for 4-, 8-, and 10-year retention models.

Table 2.27: Key Variables Used in Logistic Models

Variable	Description
still_at_4_years	1 if founder remained at least 4 years, 0 otherwise
still_at_8_years	1 if founder remained at least 8 years, 0 otherwise
still_at_10_years	1 if founder remained at least 10 years, 0 otherwise
acquirer	Acquiring firm (e.g., Google, Apple, Facebook)
segment_group	Startup segment: Product, Service, or Platform
funding_rounds	Number of funding rounds prior to acquisition
total_funding	Total pre-acquisition funding (USD million)
continue_after_purchase	1 if product continued post-acquisition, 0 otherwise
cluster_group	Startup cluster based on business characteristics
	(e.g., Technology & Data Analytics, Communication & Digital Services, etc.)

Source: R Studio

Table 2.28: AUC Scores of Logistic Models

Model	AUC	Cross-validated AUC
4-year retention	0.767	0.724
8-year retention	0.735	0.685
10-year retention	0.850	0.795

Source: R Studio

Appendix C: Interaction Effects and Robustness Checks

C.1 Interaction Analysis

To capture heterogeneity, interaction terms such as *acquirer segment_group* and *continue_after_purchase product_group* were included. Table 2.29 highlights the most significant effects, showing how retention outcomes vary across acquirersegment combinations.

Table 2.29: Significant Interaction Effects in Logistic Models

Model	Interaction Term	Effect
4-year retention	acquirer × segment_group	Google Platform (positive)
8-year retention	continue × product_group	Continuation Service (negative)
10-year retention	acquirer × product_group	Apple Product (positive)

List of Resource Persons

- **Prof. Axel Gautier**
- **Associate Prof. Jérôme Schoenmaeckers**

Bibliography and References

[1] Gautier, A., & Lamesch, J. (2021). Mergers in the digital economy. *Information Economics and Policy*, *54*, 100890.

[2] Gautier, A., & Maitry, R. (2024). Big Tech Acquisitions and Product Discontinuation. *Journal of Competition Law & Economics*, *20*(3), 246–263.

[3] Cabral, L. (2021). Merger policy in digital industries. *Information Economics and Policy*, *54*, 100866.

[4] Motta, M., & Peitz, M. (2021). Big tech mergers. *Information Economics and Policy*, *54*, 100868.

[5] Rinehart, W. (2024). Are Killer Acquisitions by Tech Giants a Real Threat to Competition?. *The Center for Growth and Opportunity.*

[6] Koski, H., Kässi, O., & Braesemann, F. (2020). Killers on the road of emerging start-upsimplications for market entry and venture capital financing (No. 81). *ETLA Working Papers.*

[7] Cunningham, C., Ederer, F., & Ma, S. (2021). Killer acquisitions. *Journal of Political Economy*, *129*(3), 649–702.

[8] Motta, M., & Peitz, M. (2021). Big tech mergers. *Information Economics and Policy*, *54*, 100868.

[9] Pike, C. (2020). Start-ups, killer acquisitions and merger Control.

[10] Barsy, L. D., & Gautier, A. (2024). Big Tech Acquisitions and Innovation: An Empirical Assessment.

[11] Bena, J., & Li, K. (2014). Corporate innovations and mergers and acquisitions. *The Journal of Finance*, *69*(5), 1923–1960. `https://doi.org/10.1111/jofi.12059`

[12] Ewens, M., & Marx, M. (2018). Founder replacement and startup performance. *Review of Financial Studies*, *31*(4), 1536–1583. `https://doi.org/10.1093/rfs/hhx108`

[13] Gans, J. S., & Stern, S. (2003). The product market and the market for ideas: Commercialization strategies for technology entrepreneurs. *Research Policy*, *32*(2), 333–350. `https://doi.org/10.1016/S0048-7333(02)00103-8`

[14] Boone, A. L., Field, L. C., Karpoff, J. M., & Raheja, C. G. (2007). The determinants of corporate board size and composition: An empirical analysis. *Journal of Financial Economics*, *85*(1), 66–101. `https://doi.org/10.1016/j.jfineco.2006.05.004`

[15] Kleinbaum, D. G., & Klein, M. (2012). *Survival Analysis: A Self-Learning Text* (3rd ed.). Springer.

[16] Bar-Isaac, H., Johnson, J. P., & Nocke, V. (2025). *Acquihiring for monopsony power.* Management Science, 71(4), 3485–3496.

Executive Summary

This thesis examines founder retention patterns following venture capital acquisitions by major technology firms such as Google, Amazon, Facebook, Apple, and Microsoft (GAFAM), focusing specifically on identifying acquisition-hiring behaviors. While previous literature has primarily examined the impact of such acquisitions on product discontinuation, this study shifts the emphasis to human capital, questioning whether they retain talent for integration.

A dataset of over 300 GAFAM acquisitions was compiled, enriched with founder employment duration. Two main methods were applied: (1) logistic regression to predict founder retention at 4-, 8-, and 10-year intervals and by short-, mid-, and long-term groups; and (2) survival analysis (KaplanMeier and Cox models) to examine the timing and factors of founder exit.

The findings show that long-term retention is significantly higher for founders who stay with the company after the acquisition. This lends credence to the idea that GAFAM businesses use an acquisition-hiring approach to acquire technical know-how and entrepreneurial skills.

Further evidence comes from the product-retention comparison. Among founders who remained employed after acquisition, approximately **60.6%** were associated with products that were later discontinued. From the product's perspective, **79.8%** of discontinued products still had at least one founder retained by the acquiring firm. These patterns strongly indicate that product continuation was not the main objective; rather, firms sought to integrate valuable human capital.

Product continuity has no effect on founder retention, according to survival analysis, while buyer identification, cluster type, and firm age have a big impact. Founders stay longer in the editorial and consumer areas at companies like Apple and Facebook.

By employing temporal and classification models to analyze post-merger founder retention, this thesis advances talent integration in the digital economy. The results show that in the high-tech industry, founder retention is a crucial indicator of strategic intent and human capital-driven incentives.

Keywords: Startup acquisition, acqui-hiring, GAFAM, founder retention, survival analysis.