Acquiring Expertise and Societal Productivity in a World of Artificial Intelligence

by

Diptasri Gupta

B.S Computer Science and Engineering, Massachusetts Institute of Technology, 2023

Submitted to the Department of Electrical Engineering and Computer Science
in partial fulfillment of the requirements for the degree of

MASTER OF ENGINEERING IN ELECTRICAL ENGINEERING AND COMPUTER SCIENCE

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

May 2024

Authored by: Diptasri Gupta
Department of Electrical Engineering and Computer Science
May 17, 2024

Certified by: David Atkin
Professor of Economics, Thesis Supervisor

Accepted by: Katrina LaCurts
Chair, Master of Engineering Thesis Committee

Acquiring Expertise and Societal Productivity in a World of Artificial Intelligence

by

Diptasri Gupta

Submitted to the Department of Electrical Engineering and Computer Science on May 17, 2024 in partial fulfillment of the requirements for the degree of

MASTER OF ENGINEERING IN ELECTRICAL ENGINEERING AND COMPUTER SCIENCE

ABSTRACT

This thesis investigates the impact of automation and advanced technologies, specifically focusing on Large Language Models (LLMs), on traditional employment structures in the modern workplace. Historically, the master-apprentice model has been integral to vocational training across various industries, facilitating the transfer of knowledge, skills, and professional ethics from one generation to the next. However, the rise of AI and machine learning challenges the viability of this model, raising critical questions about the nature and quality of mentorship and skill acquisition in work environments.

Part of a broader research initiative led by Professors Atkin, Li, and Beraja, this study explores the hypothesis that apprentices promoted without foundational mentorship may struggle in their advanced roles, potentially reducing long-term productivity gains from AI. Utilizing a comprehensive dataset from Brazilian Social Security records (RAIS) spanning 2003-2015, the research focuses on industries with a clear apprentice-master dynamic, such as finance, legal, and insurance sectors. By analyzing job code changes and pay adjustments, the study aims to correlate technological influx within companies with the productivity of workers promoted to master roles, using pay as a proxy for productivity.

Findings indicate that while technological influx does not significantly affect immediate post-promotion wages, it negatively impacts wages one and two years after promotion, suggesting potential wage stagnation or reduction. Additionally, technological influx initially increases promotion likelihood and stabilizes employee retention, though longer-term effects are less clear. These results imply that apprentices are more likely to be promoted and retained in the short term but face reduced wage growth and potentially diminished performance.

The study concludes that technological advancements can alter the traditional apprenticeship model, affecting skill acquisition and long-term productivity. Recommendations are provided for educators, industry leaders, and policymakers on optimizing apprenticeship models in an increasingly automated world. Further research will involve AI-focused evaluations to observe the real-world impact of AI integration on team dynamics, productivity, and skill development, aiming to refine our understanding of its effects on employment structures.

Thesis supervisor: David Atkin

Title: Professor of Economics

Acknowledgments

I would like to extend my gratitude to Professors Atkin, Li, and Beraja for their guidance and support throughout the course of this research. Their insights and expertise have been instrumental in shaping the work, and I am grateful for their dedication and commitment. Meeting with them weekly greatly enriched my understanding and development in the field. Thank you for your support and mentorship.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Motivation

The integration of automation and technologies such as Large Language Models (LLMs) is reshaping traditional employment structures in the modern workplace. Historically, the master-apprentice model has been a cornerstone of vocational training across various industries. This model, which involves apprentices engaging in lower-level tasks under the tutelage of experienced masters, has facilitated the transfer of knowledge, skills, and professional ethics from one generation to the next.

However, the advent of technologies such as AI and machine learning presents a paradigm shift that challenges the viability of this model. The automation of low-level tasks (previously the domain of apprentices) raises critical questions about the nature and quality of mentorship and skill acquisition in work environments.

My thesis is part of a research undertaking by Professors Atkin, Li, and Beraja, which aims to explore the implications of this shift. The research explores the hypothesis that apprentices, who rise to the rank of masters without the foundational mentorship once inherent in their training, might face difficulties in effectively performing in their advanced roles. Immediate productivity gains from AI may be reduced, even reversed, in the long term if novice workers fail to acquire the skills needed to perform complex tasks as seniors.

This research may provide insights and recommendations for educators, industry leaders, and policymakers on how to navigate and optimize the apprenticeship model in a world increasingly dominated by automation and advanced technologies.

1.2 Problem Statement

The study utilizes a comprehensive dataset from Brazilian Social Security records (RAIS) from 2003-2015, focusing on industries with a clear apprentice-master dynamic, such as finance, legal, and insurance sectors. By analyzing changes between job codes signifying a worker being promoted and pay changes, the research aims to establish a correlation between technological influx within a company and the productivity of workers that are subsequently promoted to master roles, with pay serving as a proxy for productivity. Our key research question is: When professional service firms adopt new information and communication technologies, what happens to the skills of subsequent generations of "experts"?

1.3 Related Work

Numerous studies have explored the impact of automation on skill development in the workforce. For instance, Autor, Levy, and Murnane discuss how computerization alters job tasks, emphasizing the displacement of routine, manual tasks [1]. More recent works expand on this, examining how the automation of such tasks affects the learning curve for apprentices who traditionally relied on these tasks as stepping stones in their career development.

Additionally, Higgins Kram provide foundational insights into the functions and benefits of mentorship [2]. However, studies focusing on mentorship in the context of high automation and reduced low-level task engagement are still emerging. These works are crucial in understanding the gap in experiential learning when apprentices bypass traditional mentorial stages.

The paper most relevant to the research currently being undertaken is 'Artificial Intelligence and the Modern Productivity Paradox: A Clash of Expectations and Statistics' [3]. The paper explores the decline in productivity among workers since the turn of the century, despite vast promises made by the advent of AI. It offers potential explanations of this paradox. This research explores the same domain but on a somewhat microeconomic level, focusing on the master-apprentice structure specifically, and in a more quantitative fashion.

While this provides valuable insights into the dynamics of automation, mentorship, and skill development, this research carves out a unique niche in this domain. Unlike previous studies that primarily focus on the immediate effects of automation on task allocation and skill acquisition, this work delves into the long-term implications of reduced mentorship on the professional efficacy of apprentices as they transition to master roles.

1.4 Contributions

I worked as a Research Assistant from September 2023 to May 2024. My primary role was to analyze the Brazilian Social Security dataset, with the goal of identifying patterns that were pertinent to our research questions. I frequently wrote and improved code to execute various regression models, adjusting our strategies based on the insights gained from the data and the evolving requirements of our research. Throughout the year, we developed and refined numerous regression models, progressively increasing their complexity. The iterations are detailed in this thesis.

1.5 Thesis Outline

This thesis is structured into four main chapters, excluding the introduction. In Chapter 2, I present a comprehensive overview of the RAIS dataset and the definition of derived variables that are essential throughout the thesis. I also present a more detailed conceptual framework, our estimation strategy, and our model and hypotheses within this chapter. Chapter 3 documents initial exploratory analysis that was foundational to creating the formal definitions used in Chapter 2 and throughout the paper. In Chapter 4, I detail the final version of the regression, present the conclusive results, and interpret these results. Finally, in Chapter 5, I discuss the implications of the results and examine the limitations of our findings.

Chapter 2

Background and Data Description

2.1 Conceptual Framework and Hypothesis

Without AI: Producing output involves two complementary types of tasks. An "expert" performs Task 1, and an "apprentice" performs Task 2. For instance, a senior consultant might formulate strategy or interface with clients (Task 1), while junior consultants perform supporting analysis (Task 2). Apprentices learn to perform Task 1 by collaborating with and observing the expert. Consequently, this generation's apprentices become the next generation's experts.

With AI - hypothesis: Task 2 can now be performed by AI. As a result, the expert spends less time working with apprentices, reducing the apprentices' opportunities to learn Task 1. Over time, this diminishes the skill level of the next generation of experts. In extreme cases, the expert may stop mentoring apprentices entirely, leaving the next generation of experts without any training on how to perform Task 1.

Implications: To address this potential skill gap, regulations or policies may be necessary to incentivize mentorship and ensure that apprentices continue to receive the guidance needed to become proficient experts.

2.2 Dataset

The RAIS (Relação Anual de Informações Sociais) dataset is a database compiled by the Brazilian Ministry of Labor (Ministério do Trabalho). It captures information on formal employment relationships in Brazil and is a resource for analyzing labor market dynamics, employment trends, and socioeconomic indicators in the country. It includes key variables such as employee ID (anonymized), wages, occupation codes, the firm a worker is employed in, their 'establishment' (similar to a branch), and the year of the record. This allows us to track the trajectory of a worker over time and the impact of technological changes on their careers. The scope of this research was limited to industries identified as having a master-apprentice structure: firms in the insurance, legal, real estate, and financial industry.

Year	**Worker ID**	**Firm ID**	**Establishment ID**	**Occupation Code**	**Earnings**
2007	18700712	2134 Itaú Unibanco	2134794 Araguari	413210 Bank Teller	406.86
2008	18700712	2134 Itaú Unibanco	2134794 Araguari	413210 Bank Teller	432.84
2009	18700712	2134 Itaú Unibanco	2134794 Araguari	253215 Bank Manager	504.95
2010	18700712	2134 Itaú Unibanco	2134794 Araguari	253215 Bank Manager	522.93
2011	18700712	4292 Banco Bradesco	4292020 Araguari	253215 Bank Manager	510.75
2012	18700712	4292 Banco Bradesco	4292020 Araguari	253215 Bank Manager	515.44

Table 2.1: Employment and Earnings Data

Table 2.1 displays the trajectory of a fictitious worker over 2007-2012. It is evident that this worker underwent a change in occupation from 2008 to 2009, along with an increase in earnings, which will hereafter be referred to as a promotion. The change in Firm ID between 2010-2011 signifies that the worker changed firms but remained in the same occupation.

I manually flagged certain occupation codes as ones belonging to tech professions. Individuals that occupied a tech profession were referred to as ’tech workers’.

Year	**Firm ID**	**Estab ID**	**Tech Workers**	**Non-Tech Workers**	**% Tech Workers**	**Tech Influx**
2007	2134 Itaú Unibanco	2134001 Head Office	301	6,885	4.4%	False
2008	2134 Itaú Unibanco	2134001 Head Office	335	8,001	4.1%	False
2009	2134 Itaú Unibanco	2134001 Head Office	933	10,015	9.3%	True
2010	2134 Itaú Unibanco	2134001 Head Office	1100	11,671	9.4%	True

Table 2.2: Tech Worker Data - Head Office

Year	**Firm ID**	**Estab ID**	**Tech Workers**	**Non-Tech Workers**	**% Tech Workers**	**Tech Influx**
2007	2134 Itaú Unibanco	2134794 Araguari	1	23	4.3%	False
2008	2134 Itaú Unibanco	2134794 Araguari	1	41	2.4%	False
2009	2134 Itaú Unibanco	2134794 Araguari	1	33	3.0%	True
2010	2134 Itaú Unibanco	2134794 Araguari	3	62	4.8%	True

Table 2.3: Tech Worker Data - IU Branch Araguari

Tables 2.2 and 2.3 present summary statistics for the head office of Itaú Unibanco and its

Araguari branch. We define a 'tech-influx' as a change of at least $x\%$ in the proportion of tech workers at a company's head office, coupled with a net increase of at least y workers. For the purpose of this research, we conducted regressions using various values for x and y. In the tables mentioned, x is set to 5% and y to 3, meaning that a 'tech influx' is characterized by a minimum 5% increase in the proportion of tech workers and the hiring of at least 3 additional tech workers. Since there was a 5.2% increase in the number of tech workers between 2008 and 2009, as well as 2014 additional workers, means that 'Tech Influx' was set to true in 2009. Within some regressions (as with the example above), 'Tech Influx' is set to true in **all subsequent years**. It is important to note that if a tech influx occurs within the head office in a year, the tech-influx variable will be set to true within branches of the main office as well even when the branches themselves don't meet the definition.

2.3 Key Definitions

Tech influx A change of at least $x\%$ in the proportion of tech workers at a company, coupled with a net increase of at least y workers. In different iterations of the project, 'tech influx' was set to TRUE if there was an influx meeting this definition into the head office only . In other iterations . Wherever possible, I will specify which definition of tech influx we are using.

Promotion A change in occupation code of an individual worker, accompanied by a wage increase

Tenure The number of years a worker has worked at a specific firm

2.4 Summary Statistics

I initially produced some summary statistics to ensure the definitions provided above were reasonable, seen below:

Table 2.4: Summary Statistics

Statistic	**Value**
No. of unique workers in Legal + FIRE professions	7,233,576
No. of observations per worker (mean)	2.6
No. of promotions	1,123,187
Years between promotions (mean)	2.0 years
No. of unique firms	228,680
Branches per firm (mean)	1.49
No. of firms with a "tech influx"	2,260
No. of workers employed in firms with a tech influx	83,966

I concluded that the definitions were reasonable, as the ratio of promotions to unique workers was as expected. Additionally, the ratio of firms with a 'tech influx' to those without allowed us to produce regressions with statistically significant results while maintaining a narrow enough definition.

I also produced a summary in Fig. 2.1. The x-axis groups firms by the number of establishments within that firm, and the text on each bar indicates the number of firms that fell into that grouping. The y-axis indicates how many employees work in firms in that grouping.

The majority of firms have only 1-2 establishments, and these firms collectively employ a large portion of the workforce. There is a sharp decline in both the number of firms and the number of employees as the number of establishments per firm increases. Firms with over 1000 establishments, while very rare, are key employers. Overall, the data suggests a highly skewed distribution of firms where small and mid-sized firms are critical for employment, and very large firms, though rare, also play a significant role in the job market.

2.5 Estimation Strategy

In our estimation strategy, we focus on our key prediction: that technology investments today reduce the productivity of experts (those performing Task 1) in the future. We measure this by examining a worker's productivity upon promotion (i.e., once they become the expert) and determining whether they were exposed to a significant influx of ICT investments

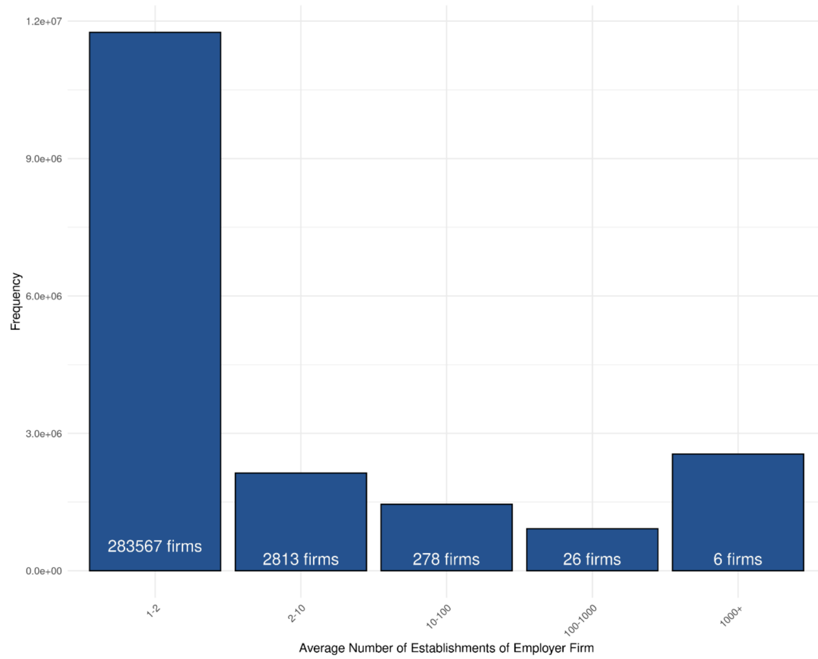

Figure 2.1: Employees by Firm-Size Bin

during their pre-promotion years (i.e., when they were apprentices).

We use wages upon promotion as a proxy for Task 1 productivity. Large changes in ICT staff serve as a proxy for the firm's ICT investments. Our comparison involves similar workers: two individuals who worked in the same office (establishment), had the same pre-promotion occupation, and were promoted to the same new position. We also include additional controls for pre-promotion wage, worker tenure, years since the last promotion, and the year of promotion.

2.6 Model and Hypothesis

The two main models on which we will run regressions, with their descriptions, results, and evaluations detailed in Chapters 4 and 5, are as follows. The full explanations of the models are detailed in chapter 4.

2.6.1 Model A

$$y_{ijkt} = \alpha + \beta 1(\text{tech\_influx}_{j,\tau} = 1 \text{ for any } \tau < t) + \gamma_t + \gamma_{o_{t-1},o_t,k} + \gamma_p + \gamma_q + \gamma_r + \epsilon$$

The model aims to analyze the impact of a tech influx in a company on wages of employees as a proxy for productivity. We aim to measure the value of the coefficient β, our hypothesis being that this value will be negative. This would signify that the appearance of tech influx within a company would dampen wages of apprentices upon promotion.

- **Dependent Variable** (y_{ijkt}):
  - y_{ijkt} is the measure of wages post-promotion. This is in a few different forms:
    - Log of real wage immediately post-promotion.
    - Log of real wage one year post-promotion.
    - Log of real wage two years post-promotion.
  - Using the log of real wages helps to stabilize variance and interpret coefficients in terms of percentage changes.
- **Independent Variable** ($1(\text{tech\_influx}_{j,\tau} = 1 \text{ for any } \tau < t)$):
  - This is our variable that measures tech influx in company j at any time before the year of promotion t.
  - We are measuring the coefficient β, which captures the impact of this tech influx on post-promotion wages.
- **Fixed Effects** (γ Terms):
  - γ_t: Year
  - $\gamma_{o_{t-1},o_t,k}$: Branch, pre-promotion, and post-promotion occupations
  - γ_p: Number of years between promotions

– γ_q: Tenure at the firm

– γ_r: Pre-promotion wage

2.6.2 Model B

$$y_{ijkt} = \alpha + \sum_{\tau=-2}^{0} \beta_\tau \, \text{tech\_influx}_{j(t+\tau)} + \sum_{\tau=-2}^{-1} \phi_\tau \, \text{replaced}_{j(t+\tau)} + \lambda_p + \gamma_w + \gamma_p + \epsilon$$

The model aims to analyze the effects of a tech influx on a worker's likelihood of being promoted or leaving the firm. This analysis helps us understand whether technological advancements lead to increased employee attrition, potentially due to a reduced need for apprentices, and whether certain employees face barriers that prevent them from advancing beyond the apprentice level.

- **Dependent Variable** (y_{ijkt}):
  - y_{ijkt}: worker variable that measures the probability of a certain event occurring upon tech influx. The events include:
    - Leaving the firm
    - Receiving a promotion
- **Independent Variables**
  - $\sum_{\tau=-2}^{2} \beta_\tau \, \textbf{tech\_influx}_{j(t+\tau)}$: The variables that measures the impact of tech influx in company j at time $t + \tau$.
  - $\sum_{\tau=-2}^{-1} \phi_\tau \, \textbf{replaced}_{j(t+\tau)}$: Variables indicating whether a previous tech influx value was replaced by FALSE, to mitigate losing large portions of the data
  - λ_p: Whether a promotion with wage increase occurred. This variable is omitted when the
- **Fixed Effects** (γ Terms):

- γ_w: Establishment
- γ_p: Number of years between promotions

Chapter 3

Initial Data Exploration

The definitions and some of the methods outlined in the previous chapter were the result of data exploration conducted early on in the research, which will be detailed below.

3.1 Fixed Effect Investigation

The first investigation (displayed in Fig. 1) 3.1 shows the percentage wage change of a worker year-to-year plotted for every year on the primary axis relative to that of 2005. This was performed as I knew some variation of wage change would be our independent variable in the final regression. GDP Growth for each year was plotted on a secondary axis to determine if macroeconomic effects played a role in influencing wage growth year-on-year. The data shows that there is some positive correlation between wage growth and GDP growth (except for in 2009 and 2011).[1]

The graph also displays dummies for each year for the establishment a worker worked in, the firm they worked in, and an interaction variable that describes the establishment a worker worked in, their initial occupation, and their final occupation year-on-year. All these

[1] It must be noted that 2010 was omitted from this initial exploration as the occupation codes were missing for that year. I later retrieved 2010 data by subsetting the 2010 dataset to include all the workers that were present in other years of the dataset.

variables were identified as potentially being used as fixed effects. The fixed effect "estab x occ x occ" (green) appears to be the most consistently aligned with both wage jumps and GDP growth fluctuations, making it a potentially valuable fixed effect.

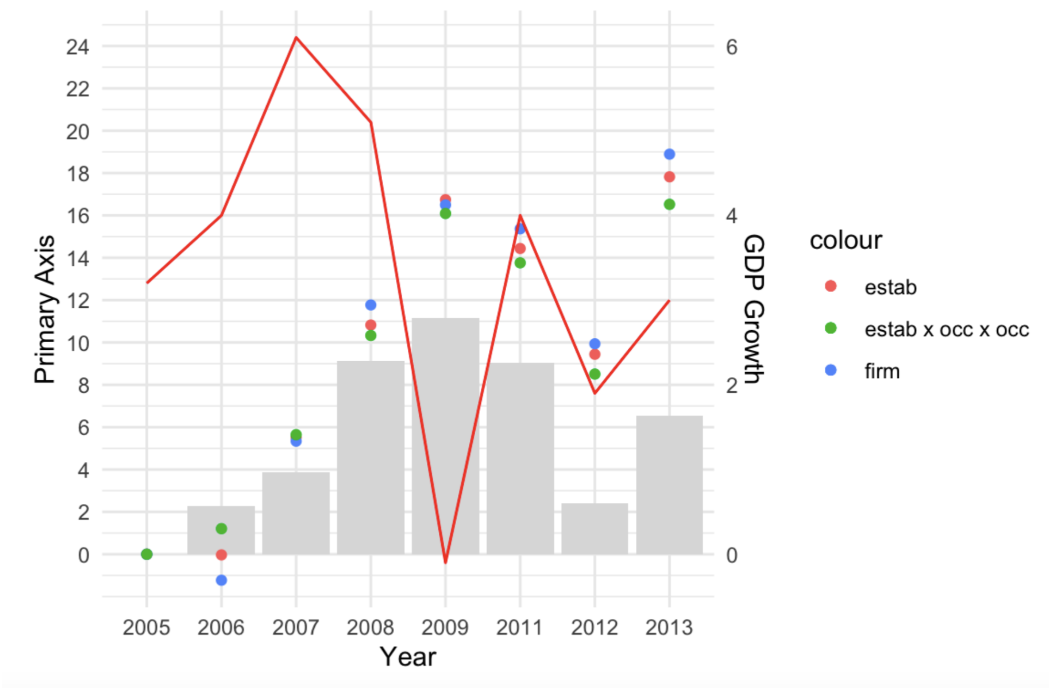

Figure 3.1: Wage Dummies with GDP Growth

3.2 Tech Presence Variable Investigation

The next investigation involved refining the definition for our variable that would indicate tech presence within a firm, which we would measure to validate our hypothesis. Several potential definitions were considered for analysis:

A. Share of tech workers within a company [continuous variable]

B. Above median share of tech presence within a company (above 6.67% of a company consists of tech workers) [discrete binary variable]

C. Significant influx of tech workers in a firm; a change of at least $x\%$ in the proportion of tech workers at a company's head office, coupled with a net increase of at least y workers [discrete binary variable]

To conduct this analysis, a simplified version of the regressions that would later be applied to the dataset was run on each of the three tech-influx variables to determine the most suitable one for our analysis.

I conducted several modifications of the regression mentioned below, incorporating different combinations of fixed effects and various subsets of the data. Specifically, I performed the regressions on all wage jumps associated with a change in occupation, as well as on the subset consisting solely of positive wage jumps that accompanied a change in occupation. The regression that will be analyzed in this section can be described as follows.

$$y_{ijkt} = \alpha + \beta \text{tech\_presence} + \gamma_t + \gamma_{o_{t-1},o_t,k} + \epsilon$$

y_{ijkt}: Wage growth accompanied by a promotion

tech_presence: One of variables A, B and C

A: Average tech presence in a company while a worker was an apprentice (pre-promotion) [continuous]

B: Whether a company had an above median number of tech workers for the majority of the time a worker was an apprentice [discrete binary]

C: Whether a tech influx had occurred in any of the years in which the worker was an apprentice [discrete binary]

$\gamma_{o_{t-1},o_t,k}$: fixed effect for branch, pre-and-post-promotion occupation

This was run on the subset of the data that only included positive wage jumps occupied by a change in occupation (a promotion).

Table 3.1: Tech Presence Investigation Regression: A

Variable	**Estimate**	**Std. Error**	**Pr(>|t|)**
2006	3.31869	0.59414	2.33e-08 ***
2007	6.33125	0.56648	< 2e-16 ***
2008	9.25483	0.55576	< 2e-16 ***
2009	12.72648	0.56181	< 2e-16 ***
2011	9.98357	0.61854	< 2e-16 ***
2012	3.14088	0.57714	5.28e-08 ***
2013	6.92126	0.53687	< 2e-16 ***
Tech presence A	0.04923	0.05053	0.00314 *

Model Statistics:
Multiple R-squared (full model): 0.4035
Adjusted R-squared: 0.2163

Table 3.2: Tech Presence Investigation Regression: B

Variable	**Estimate**	**Std. Error**	**Pr(>|t|)**
2006	2.9490	0.6665	9.67e-06 ***
2007	6.2983	0.6310	< 2e-16 ***
2008	8.6301	0.6268	< 2e-16 ***
2009	11.4204	0.6481	< 2e-16 ***
2011	9.2674	0.7183	< 2e-16 ***
2012	2.9395	0.6585	8.05e-06 ***
2013	8.1911	0.6062	< 2e-16 ***
Tech presence B	-1.6570	0.6074	0.00638 **

Model Statistics:
Multiple R-squared (full model): 0.3129
Adjusted R-squared: 0.1815

Tables 3.1, 3.2, and 3.3 display the outputs of the regressions for the tech presence variables A, B and C. Tech variable C had a large statistically significant negative coefficient that was robust to changes in fixed effects and subsets of data (positive jumps vs all jumps). We therefore decided to proceed with this definition of tech presence in our development of our model but continued to vary tech variable C's parameters (x and y).

Table 3.3: Tech Presence Investigation Regression: C

Variable	**Estimate**	**Std. Error**	**Pr(>|t|)**
2006	3.0937	0.6627	3.04e-06 ***
2007	5.4847	0.6290	< 2e-16 ***
2008	7.9186	0.6216	< 2e-16 ***
2009	9.5799	0.6487	< 2e-16 ***
2011	6.1881	0.7242	< 2e-16 ***
2012	-0.1376	0.6649	0.836
2013	3.5085	0.6284	2.37e-08 ***
Tech presence C	-20.5509	0.7800	< 2e-16 ***

Model Statistics:
Multiple R-squared (full model): 0.3188
Adjusted R-squared: 0.1885

Chapter 4

Final Regression and Results

As discussed in chapter 2, the final results were obtained by running the following two regressions:

4.1 Model A

4.1.1 Description

$$y_{ijkt} = \alpha + \beta 1(\text{tech\_influx}_{j,\tau} = 1 \text{ for any } \tau < t) + \gamma_t + \gamma_{o_{t-1},o_t,k} + \gamma_p + \gamma_q + \gamma_r + \epsilon$$

The model aims to analyze the impact of a tech influx in a company on wages of employees as a proxy for productivity. We aim to measure the value of the coefficient β, our hypothesis being that this value will be negative. This would signify that the appearance of tech influx within a company would dampen wages of apprentices upon promotion.

- **Dependent Variable** (y_{ijkt}):
  - y_{ijkt} is the measure of wages post-promotion. This is in a few different forms:
    - Log of real wage immediately post-promotion.
    - Log of real wage one year post-promotion.

        * Log of real wage two years post-promotion.

    – Using the log of real wages helps to stabilize variance and interpret coefficients in terms of percentage changes.

- **Independent Variable** ($1(\text{tech\_influx}_{j,\tau} = 1 \text{ for any } \tau < t)$):

    – This is our variable that measures tech influx in company j at any time before the year of promotion t.

    – We are measuring the coefficient β, which captures the impact of this tech influx on post-promotion wages.

- **Fixed Effects** (γ Terms):

    – γ_t: Year

    – $\gamma_{o_{t-1},o_t,k}$: Branch, pre-promotion, and post-promotion occupations

    – γ_p: Number of years between promotions

    – γ_q: Tenure at the firm

    – γ_r: Pre-promotion wage

By including fixed effects for the year, we ensure that macroeconomic trends or events affecting wages in a particular year do not bias the results. Controlling for branch and occupation fixed effects accounts for the fact that different branches and occupations may have distinct wage structures and promotion policies. The fixed effect for the number of years between promotions helps to isolate the effect of tech influx from the effects of promotion timing. Employees' tenure can significantly affect their wages due to accumulated experience and loyalty to the firm. Controlling for tenure ensures that the estimated impact of tech influx is not confounded by tenure effects. Pre-promotion wages are controlled for to ensure that the analysis focuses on wage changes due to promotion, independent of initial wage levels.

The key difference in this regression approach is that we directly measure post-promotion wages while including pre-promotion wages as a fixed effect. In the previous approach, we measured the log difference between pre- and post-promotion wages. We chose the current approach because collapsing the data into a single log difference measure meant losing some information about the absolute levels of pre- and post-promotion wages. This approach allows us to retain that valuable information. While our new model has added complexity because of an extra fixed effect, we have no loss of this information.

4.1.2 Results

Table 4.1: Model A Results for Post-Promotion Wages

Variable	Estimate	Std. Error	t value	p value
tech influx	-0.001739	0.004013	-0.433	0.665
2008	0.061430	0.003585	17.143	<2e-16 ***
2009	0.075445	0.004007	18.828	<2e-16 ***
2010	0.135797	0.003979	34.125	<2e-16 ***
2011	0.252127	0.003965	63.588	<2e-16 ***
2012	0.216174	0.003989	54.186	<2e-16 ***
2013	0.239580	0.004381	54.681	<2e-16 ***
pre	0.662362	0.002230	297.060	<2e-16 ***

Notes:

Residual standard error: 0.1934 on 125926 degrees of freedom

Multiple R-squared (full model): 0.9568, Adjusted R-squared: 0.9166

Multiple R-squared (proj model): 0.5377, Adjusted R-squared: 0.1065

Observations with tech influx = FALSE: 212926; = TRUE: 134595

Table 4.2: Model A Results for 1 Year Post-Promotion Wages

Variable	Estimate	Std. Error	t value	p value
tech influx	-0.019542	0.005669	-3.447	0.000567 ***
2008	0.124596	0.005085	24.502	<2e-16 ***
2009	0.114264	0.005699	20.051	<2e-16 ***
2010	0.194138	0.005658	34.315	<2e-16 ***
2011	0.261714	0.005631	46.477	<2e-16 ***
2012	0.222512	0.005666	39.274	<2e-16 ***
2013	0.276751	0.006263	44.258	<2e-16 ***
pre	0.567667	0.003266	173.837	<2e-16 ***

Notes:

Residual standard error: 0.2629 on 110846 degrees of freedom

Multiple R-squared (full model): 0.9162, Adjusted R-squared: 0.8328

Multiple R-squared (proj model): 0.3222, Adjusted R-squared: -0.3521

Observations with tech influx = FALSE: 190993; = TRUE: 126451

Table 4.3: Model A for Results for 2 Years Post-Promotion Wages)

Variable	Estimate	Std. Error	t value	p value
tech influx	-0.049475	0.006925	-7.145	9.08e-13 ***
2008	0.132189	0.006242	21.178	<2e-16 ***
2009	0.156030	0.006972	22.380	<2e-16 ***
2010	0.207405	0.006972	29.823	<2e-16 ***
2011	0.257581	0.006919	37.228	<2e-16 ***
2012	0.227727	0.006982	32.614	<2e-16 ***

Variable	Estimate	Std. Error	t value	p value
2013	0.292986	0.007899	37.092	<2e-16 ***
pre	0.512623	0.004131	124.096	<2e-16 ***
Notes:				
Residual standard error: 0.301 on 92850 degrees of freedom				
Multiple R-squared (full model): 0.89, Adjusted R-squared: 0.772				
Multiple R-squared (proj model): 0.2299, Adjusted R-squared: -0.5958				
Observations with tech influx = FALSE: 166353; = TRUE: 107953				

4.1.3 Interpretation

The values relevant to the tech influx variable only are reproduced from tables 4.1, 4.2 and 4.3. The value of 'Estimate' corresponds to β, the value we have been trying to measure.

Table 4.4: Impact of Tech Influx on Wages, Post, 1-Year-Post, and 2-Year-Post-Promotion

Model	Estimate	Std. Error	t value	p value
Post-Promotion Wages	-0.001739	0.004013	-0.433	0.665
1 Year Post-Promotion Wages	-0.019542	0.005669	-3.447	0.000567 ***
2 Years Post-Promotion Wages	-0.049475	0.006925	-7.145	9.08e-13 ***

As can be seen in Table 4.4, The tech influx variable does not have a statistically significant effect on immediate post-promotion wages. 1 year after an apprentice has been promoted, the tech influx has a significant negative effect, reducing wages by approximately 1.95%. 2 years post-promotion, the tech influx has a significant negative effect, reducing wages by approximately 4.95%. The impact of the tech influx appears to become more pronounced over time, with the most substantial negative effect observed 2 years after promotion. This could indicate that while the immediate post-promotion wage is unaffected, the longer-term wage trajectory is negatively influenced by the tech influx.

4.1.4 Testing for Robustness

Different Parameters for Tech Influx

As I did within the initial development of the model, I ran this regression with various levels of severity for the tech influx variable. As a reminder, tech influx is a change of at least x% in the proportion of tech workers at a company's head office, coupled with a net increase of at least y workers. By changing these parameters, we can make the thresholds for which a company is considered to have had a tech influx more or less severe.

Table 4.5: Impact of Tech Influx on Post-Promotion Wages for Different Values of X and y

Model	**Estimate**	**Std. Error**	**t value**	**p value**
$x = 5, y = 3$				
Post-Promotion Wages	-0.001739	0.004013	-0.433	0.665
1 Year Post-Promotion Wages	-0.019542	0.005669	-3.447	0.000567 ***
2 Years Post-Promotion Wages	-0.049475	0.006925	-7.145	9.08e-13 ***
$x = 7, y = 5$ (more severe)				
Post-Promotion Wages	-0.009496	0.007503	-1.266	0.206
1 Year Post-Promotion Wages	-0.035220	0.010757	-3.274	0.00106 **
2 Years Post-Promotion Wages	-0.062852	0.013444	-4.675	2.94e-06 ***
$x = 3, y = 2$ (less severe)				
Post-Promotion Wages	-0.013828	0.002654	-6.904	5.06e-12 ***
1 Year Post-Promotion Wages	-0.067534	0.003993	-16.904	<2e-16 ***
2 Years Post-Promotion Wages	-0.041755	0.004954	-8.429	<2e-16 ***

The immediate post-promotion wages show mixed results: insignificant impacts for $x = 5, y = 3$ and $x = 7, y = 5$, but a significant negative impact for $x = 3, y = 2$. The 1 year post-promotion wages consistently show significant negative impacts across all scenarios, with the severity of impact increasing with the severity of x and y values. The 2 years post-promotion wages also consistently show significant negative impacts, with the severity of the impact being highest for the most severe scenario ($x = 7, y = 5$). The tech influx still generally has a more pronounced negative effect on wages as the time post-promotion increases and the severity of X and y values rises.

It is unclear why the immediate post-promotion wages show mixed results, and this may be a point of further investigation. A possible explanation could be that less severe influxes might disrupt wage dynamics immediately due to increased competition or rapid internal shifts. More severe influxes might result in longer-term strategic adjustments, delaying immediate wage impacts.

Head Office Influx vs. Company-Wide Influxes

Additionally, we have been measuring tech-influx within a company's head office (its largest branch) thus far within this version of the model, assuming that the majority of IT workers would be employed there. As an additional robustness check, we ran the regression applying the same definition of a tech influx to the entire company instead of the company's head office. This, a tech influx would be defined as a change of at least $x\%$ in the proportion of tech workers at a company (all branches included), coupled with a net increase of at least y workers.

Table 4.6: Impact of Tech Influx on Post-Promotion Wages for Different Values of x and y: Company-wide

Model	Estimate	Std. Error	t value	p value
$x = 5, y = 3$				
Post-Promotion Wages	-0.040015	0.003257	-12.29	<2e-16 ***
1 Year Post-Promotion Wages	-0.078702	0.004845	-16.24	<2e-16 ***
2 Years Post-Promotion Wages	-0.024334	0.005926	-4.106	4.03e-05 ***
$x = 7, y = 5$ (more severe)				
Post-Promotion Wages	-0.031542	0.005384	-5.859	4.66e-09 ***
1 Year Post-Promotion Wages	-0.025871	0.008157	-3.172	0.00152 **
2 Years Post-Promotion Wages	-0.042222	0.010163	-4.154	3.26e-05 ***
$x = 3, y = 2$ (less severe)				
Post-Promotion Wages	-0.003927	0.004017	-0.978	0.328
1 Year Post-Promotion Wages	-0.016905	0.005673	-2.98	0.00289 **
2 Years Post-Promotion Wages	-0.050266	0.006929	-7.254	4.07e-13 ***

The post-promotion wages show a significant negative impact for $x = 5, y = 3$ and $x = 7, y = 5$, but an insignificant impact for $x = 3, y = 2$. The 1 year post-promotion wages consistently show significant negative impacts across all scenarios, with the severity of the impact increasing with higher X and y values. The 2 years post-promotion wages also consistently show significant negative impacts, with the severity of the impact being highest for the most severe scenario ($X = 7, y = 5$). The tech influx generally has a more pronounced negative effect on wages as the time post-promotion increases and the severity of X and y values rises.

4.1.5 Model A Conclusions

The impact of tech influx on wages is generally more pronounced when applied to the whole company compared to just the head office. This could be due to the broader scope of influence when the entire company is considered, leading to more substantial organizational changes and adjustments in wages.

The negative impact is significant across almost all scenarios when applied to the whole company, whereas for the head office, the immediate post-promotion wages sometimes show insignificant results. This difference could be due to the head office potentially having more stable wage structures and promotion policies that are less sensitive to changes in tech worker proportions and numbers. However, as stated previously, further research must be done to confirm this is the case.

More severe tech influx scenarios $x = 7, y = 5$ tend to show significant negative impacts regardless of whether the scope is the head office or the whole company. However, the magnitude of the impact is larger when the whole company is considered. This indicates that larger influxes disrupt wage dynamics more deeply across the entire organization compared to just the head office. Overall, the robustness checks showed that our findings for post-promotion wages 1 year and 2 years after a promotion are very robust, confirming our initial hypothesis that tech presence within a company negatively affects productivity (therefore wages) upon promotion of an apprentice.

4.2 Model B

This model is a general specification that resembles an event study. The tech influx variable works slightly differently from Model A. In Model A, the tech influx model was cumulative, meaning that if a tech influx occurred within a company, the variable would be set to true in all years thereafter. In contrast, in this model, tech influx is set to TRUE if a tech influx occurred in that year only. Therefore, β_τ measures the coefficient on a tech influx that occurred 2 years before a promotion ($\tau = -2$), 1 year before a promotion ($\tau = -1$), the year of the promotion itself ($\tau = 0$), 1 year after a promotion ($\tau = 1$), and 2 years after a promotion ($\tau = 2$). This model allows us to measure the specific impact of technological influx at different time points relative to a promotion event.

The regression analysis in Model B examines the effect of past tech influx on current log wages and other events at the firm level, considering tech influx for up to three years prior. Missing values for wages and tech influx before the sample start are set to 0, with added dummies to account for these cases. Fictitious observations are added starting four years before the sample to avoid data loss.

In Model B, the dataset is structured differently from Model A, which focused solely on wage jumps upon promotion. Model B uses the full set of data, where each worker has multiple time series—one for each firm they work for. Time series before a worker joins a firm are removed. For example, if a worker's trajectory from 2007-2012 shows they moved from Itaú Unibanco to Banco Bradesco in 2011, the modified dataset would have the data series duplicated: one time series for Itaú Unibanco from 2007-2012, and another for Banco Bradesco from 2011-2012. In general, one time series is created for each separate firm a worker works at, and all the data points for that worker after they started working at that firm are included. This structure ensures that the impact of tech influx and other variables is accurately measured within the context of each firm for a worker.

Year	Worker ID	Firm ID	Occupation Code	Earnings
2007	18700712	2134 Itaú Unibanco	413210 Bank Teller	406.86
2008	18700712	2134 Itaú Unibanco	413210 Bank Teller	432.84
2009	18700712	2134 Itaú Unibanco	253215 Bank Manager	504.95
2010	18700712	2134 Itaú Unibanco	253215 Bank Manager	522.93
2011	18700712	4292 Banco Bradesco	253215 Bank Manager	510.75
2012	18700712	4292 Banco Bradesco	253215 Bank Manager	515.44

Table 4.7: Employee Data Before Manipulation

Year	Worker ID	Firm ID	Occupation Code	Earnings
2007	18700712	2134 Itaú Unibanco	413210 Bank Teller	406.86
2008	18700712	2134 Itaú Unibanco	413210 Bank Teller	432.84
2009	18700712	2134 Itaú Unibanco	253215 Bank Manager	504.95
2010	18700712	2134 Itaú Unibanco	253215 Bank Manager	522.93
2011	18700712	4292 Itaú Unibanco	253215 Bank Manager	510.75
2012	18700712	4292 Itaú Unibanco	253215 Bank Manager	515.44
2011	18700712	4292 Banco Bradesco	253215 Bank Manager	510.75
2012	18700712	4292 Banco Bradesco	253215 Bank Manager	515.44

Table 4.8: Employee Data After Manipulation

The model includes a dummy for promotions, interacting this with tech influx variables from the current and past four years to replicate previous promotion specifications. It incorporates fixed effects such as in model A. The approach ensures no observations are lost at the start of the panel, and the specification is flexible enough to switch the dependent variable to other questions like "are you still working at the firm" or "are you promoted" by adjusting the dependent variable of the specification accordingly.

This approach allows us to measure the precise impact of technological influx on various outcomes, such as wage changes, employment continuity, and promotion likelihood, at different time points relative to a promotion event. By doing so, it provides a more nuanced understanding of how tech influx influences career trajectories and organizational dynamics over time. By including all the data instead of just the promotion data, we can compare individuals who were and weren't promoted, as well as comparing just people who were promoted (as in model A).

4.2.1 Description

$$y_{ijkt} = \alpha + \sum_{\tau=-2}^{0} \beta_\tau \, \text{tech\_influx}_{j(t+\tau)} + \sum_{\tau=-2}^{-1} \phi_\tau \, \text{replaced}_{j(t+\tau)} + \lambda_p + \gamma_w + \gamma_p + \epsilon$$

The model aims to analyze the effects of a tech influx on a worker's likelihood of being promoted or leaving the firm. This analysis helps us understand whether technological advancements lead to increased employee attrition, potentially due to a reduced need for apprentices, and whether certain employees face barriers that prevent them from advancing beyond the apprentice level.

- **Dependent Variable** (y_{ijkt}):
  - y_{ijkt}: worker variable that measures the probability of a certain event occurring upon tech influx. The events include:

    * Leaving the firm
    * Receiving a promotion

- **Independent Variables**
  - $\sum_{\tau=-2}^{2} \beta_\tau \, \textbf{tech\_influx}_{j(t+\tau)}$: The variables that measures the impact of tech influx in company j at time $t + \tau$.
  - $\sum_{\tau=-2}^{-1} \phi_\tau \, \textbf{replaced}_{j(t+\tau)}$: Variables indicating whether a previous tech influx value was replaced by FALSE, to mitigate losing large portions of the data
  - λ_p: Whether a promotion with wage increase occurred. This variable is omitted when the
- **Fixed Effects** (γ Terms):
  - γ_w: Establishment
  - γ_p: Number of years between promotions

4.2.2 Results

Table 4.9: Impact of tech influx on a worker receiving a promotion

Variable	**Estimate**	**Std. Error**	**t value**	**p value**
tech_influx	0.039770	0.004094	9.713195	$< 2.2e-16^{***}$
l1.tech_influx	-0.005035	0.004039	-1.246633	0.2125
l2.tech_influx	0.004219	0.006807	0.619788	0.5354
l3.tech_influx	0.024331	0.006186	3.933463	$8.375e-05^{***}$
l1.replaced	0.002210	0.001764	1.253116	0.2102
l2.replaced	0.000444	0.001552	0.286037	0.7749
l3.replaced	-0.005613	0.001422	-3.947040	$7.914e-05^{***}$

Variable	Estimate	Std. Error	t value	p value
2008	0.036673	0.003645	10.060847	$< 2.2e-16^{***}$
2009	-0.005784	0.003904	-1.481783	0.1384
2010	0.009238	0.004434	2.083721	0.0372^{*}
2011	0.009442	0.004700	2.008890	0.0446^{*}
2012	-0.008574	0.004604	-1.862470	0.0625
2013	-0.004487	0.005527	-0.811870	0.4169

Table 4.9 displays the impact of tech influx on a worker **receiving a promotion**. l1.tech_influx, l2.tech_influx, and l3.tech_influx display the values of tech influx 1, 2, and 3 years before the event being analyzed. The immediate impact of tech influx is positive and significant, indicating that a tech influx is associated with a worker being more likely to be promoted. However, the lagged effects of tech influx show a mixed impact, with the third lagged value also being positive and significant. The conclusion is that it is unclear whether a tech influx results in more apprentices being promoted, but there is some evidence to suggest it may increase likelihood of a promotion.

Table 4.10: Regression Results with Fixed Effects

Variable	Estimate	Std. Error	t value	p value
tech_influx	-0.024007	0.008941	-2.684905	0.007255 **
l1_tech_influx	-0.027584	0.012502	-2.206350	0.027360 *
l2_tech_influx	-0.033134	0.014676	-2.257713	0.023964 *
l3_tech_influx	0.018204	0.018756	0.970607	0.331740
l1_replaced	-0.034751	0.005676	-6.122529	9.2220e-10 ***
l2_replaced	0.041466	0.005825	7.118417	1.0943e-12 ***

Variable	Estimate	Std. Error	t value	p value
l3_replaced	0.028661	0.004487	6.388030	1.6829e-10 ***
promotion	-0.101711	0.002360	-43.104935	$<$ 2.2e-16 ***
2008	-0.066282	0.008698	-7.620562	2.5334e-14 ***
2009	0.162523	0.011039	14.723017	$<$ 2.2e-16 ***
2010	0.216030	0.013899	15.542370	$<$ 2.2e-16 ***
2011	0.144772	0.013978	10.357435	$<$ 2.2e-16 ***
2012	0.156000	0.014521	10.743435	$<$ 2.2e-16 ***
2013	0.190572	0.014265	13.359585	$<$ 2.2e-16 ***

Table 4.10 displays the impact of tech influx on a worker **leaving a firm**. The tech influx variable and its first two lags have negative and significant impacts on leaving a firm, indicating that tech influx tends to reduce the likelihood that an apprentice leaves the firm over the short term. The promotion variable shows a substantial and significant negative impact on leaving a firm, suggesting those who receive promotions are less likely to leave a firm, aligning with common intuition.

4.2.3 Additional General Regression

While developing the linear probability models in Model B, I also coded up a general specification that employed this event study model to measure wages unconditional on promotion. The model is as follows:

$$y_{ijkt} = \alpha + \sum_{\tau=-2}^{0} \beta_\tau \, \text{tech\_influx}_{j(t+\tau)} + \sum_{\tau=-2}^{-1} \phi_\tau \, \text{replaced}_{j(t+\tau)} + \lambda_p + \sum_{\tau=-2}^{0} \eta_\tau \, (\text{tech\_influx}_{j(t+\tau)} * \lambda_p) + \lambda_p + \gamma_w + \gamma_p +$$

- **Dependent Variable** (y_{ijkt}):

- y_{ijkt}: log wages of a worker

- **Independent Variables**

  - $\sum_{\tau=-2}^{2} \beta_\tau \, \textbf{tech\_influx}_{j(t+\tau)}$: The variables that measures the impact of tech influx in company j at time $t + \tau$.
  - $\sum_{\tau=-2}^{-1} \phi_\tau \, \textbf{replaced}_{j(t+\tau)}$: Variables indicating whether a previous tech influx value was replaced by FALSE, to mitigate losing large portions of the data
  - λ_p: Whether a promotion with wage increase occurred. This variable is omitted when the
  - $\sum_{\tau=-2}^{0} \eta_\tau \, (\textbf{tech\_influx}_{j(t+\tau)} * \lambda_p)$: A set of interaction variables that determine the combined effect of a tech influx and a promotion

- **Fixed Effects** (γ Terms):

  - γ_w: Establishment
  - γ_p: Number of years between promotions

This regression is different from that which was run on Model A in a few key ways. Firstly, it specifically investigates relationship between the length between a tech influx occurring and a worker's wage. It is also run on the complete dataset similar to model B, so we can compare those who got promotions to those who did not instead of simply comparing across promotions (comparing firms that underwent a tech influx as compared to those that did not).

Results

Table 4.11: General Specification Result on Log Earnings

Variable	**Estimate**	**Std. Error**	**t value**	**p value**
tech_influx	-0.000120	0.000692	-0.174164	0.8617
l1.tech_influx	-0.020426	0.000709	-28.793240	$< 2.2e-16^{***}$
l2.tech_influx	-0.036411	0.000786	-46.299379	$< 2.2e-16^{***}$
l3.tech_influx	-0.010226	0.001047	-9.764984	$< 2.2e-16^{***}$
l1.replaced	0.000056	0.000391	0.143747	0.8857
l2.replaced	-0.005236	0.000323	-16.228215	$< 2.2e-16^{***}$
l3.replaced	-0.004574	0.000290	-15.750731	$< 2.2e-16^{***}$
promotion	0.060126	0.000349	172.238498	$< 2.2e-16^{***}$
2008	0.169695	0.000531	319.453891	$< 2.2e-16^{***}$
2009	0.331792	0.000807	411.275057	$< 2.2e-16^{***}$
2010	0.506633	0.000964	525.807678	$< 2.2e-16^{***}$
2011	0.714059	0.001026	695.976893	$< 2.2e-16^{***}$
2012	0.912923	0.001102	828.310582	$< 2.2e-16^{***}$
2013	1.102382	0.001192	925.073618	$< 2.2e-16^{***}$
tech_influx:promotion	-0.007142	0.001536	-4.651142	3.301e-06***
l1.tech_influx:promotion	0.016451	0.001764	9.324647	$< 2.2e-16^{***}$
l2.tech_influx:promotion	0.037388	0.001774	21.080193	$< 2.2e-16^{***}$
l3.tech_influx:promotion	0.016956	0.002658	6.380389	1.767e-10***

As this regression has interaction variables, it is a little more difficult to interpret. As seen before, tech influx continues to negatively affect wages of workers. A tech influx that occurred 2 years prior seems to have the most negative effect (a 2% decrease). As expected, a promotion is correlated to wages increasing. The interaction between tech influx and promotion initially shows a negative impact, but the lagged interaction terms show significant positive impacts on log earnings. The reason for the coefficients on interaction terms need to be investigated further, which is why this regression is incomplete and will not be addressed in our conclusion.

Chapter 5

Conclusion and Next Steps

5.0.1 Summary

In this chapter, we analyzed the impact of technological influx on various employee outcomes, specifically focusing on post-promotion wages and the likelihood of promotion or leaving the firm. We employed two regression models, Model A and Model B, to measure these effects.

Model A Findings

Model A examined the effect of tech influx on post-promotion wages. The results indicate that:

- Immediate Post-Promotion Wages: Tech influx did not have a statistically significant effect on immediate post-promotion wages.
- 1 Year Post-Promotion Wages: Tech influx had a significant negative effect, reducing wages by approximately 1.95%. For context, the mean wage jumps upon promotion is around 30%.

- 2 Years Post-Promotion Wages: Tech influx had a more pronounced significant negative effect, reducing wages by approximately 4.95%. The results suggest that while tech influx does not affect immediate post-promotion wages, its negative impact on wages becomes evident one and two years after promotion. This could imply that technological advancements within a company may lead to longer-term wage stagnation or reduction for promoted employees.

Robustness checks were conducted by varying the severity of tech influx parameters and comparing head office-specific influxes to company-wide influxes. The findings consistently showed a negative impact on wages, with more severe tech influx scenarios resulting in more substantial wage reductions. These results were particularly pronounced when considering the entire company rather than just the head office.

Model B Findings

Model B, designed as a general specification similar to an event study, assessed the specific impact of tech influx at different time points relative to a promotion event. The findings include:

- Likelihood of Promotion: Immediate tech influx increased the likelihood of promotion by 3.9%. Lagged effects of tech influx showed mixed impacts, with some positive and significant effects two to three years after the influx. For example, three years after an influx, a worker is 2.4% more likely to be promoted.
- Likelihood of Leaving the Firm: Immediate tech influx and its first two lags reduced the likelihood of an apprentice leaving the firm. Promotions had a significant negative impact on the likelihood of leaving the firm, indicating that promoted employees are less likely to leave - 2.7%, 3.3% and 1.8% less likely for a tech influx that occurred 1, 2, and 3 years before. These results suggest that tech

influx initially increases the chances of promotion and also appears to stabilize employee retention in the short term.

5.1 Overall Conclusion

The analysis demonstrates that tech influx within a company has a nuanced impact on employee outcomes. While immediate post-promotion wages remain unaffected, the longer-term wage trajectory is negatively influenced by technological advancements. Conversely, tech influx can initially promote employee retention and increase promotion likelihood, but its longer-term effects are less clear. The current analysis paints a picture of apprentices being more likely to be promoted and stay at a firm given a technological influx, but being paid less and performing more poorly.

Overall, the findings align with our initial hypothesis that tech influx negatively affects productivity, as reflected in wage reductions over time. It must be noted that these findings are of interest because we may expect a firm to hire a lot of IT professionals when they are doing well. However, since employee's wages are generally decreasing upon promotion as compared to a firm with no tech influx, this suggests that the hiring of tech professionals may be changing the fundamental structure of an organization and the way in which these apprentices gain skills.

5.2 Next Steps

To further validate our hypothesis, we must dismiss competing explanations for the results obtained. This involves a thorough assessment of alternative factors that could influence employment trends, ensuring a clear correlation between ICT adoption and workforce changes. Another area of exploration, given our dataset, is the employment margin. Specifically, we need to determine if firms experiencing a tech influx hire fewer apprentices while maintaining the same number of experts. Examining these aspects will provide a deeper understanding of the impact of ICT adoption on workforce composition.

The next stage of the research will involve implementing AI-focused evaluations by analyzing actual rollouts of generative AI across different teams. This approach will enable us to observe the real-world impact of AI integration on team dynamics, productivity, and skill development. This will help to further substantiate our findings and refine our understanding of its effects on employment structures.

References

[1] D. Autor, F. Levy, and R. Murnane, "The skill content of recent technological change: An empirical exploration," in *Proceedings*, 2003.

[2] E. Brynjolfsson, D. Rock, and C. Syverson, "Artificial intelligence and the modern productivity paradox: A clash of expectations and statistics," *NATIONAL BUREAU OF ECONOMIC RESEARCH*, 2017, Working Paper.

[3] M. C. Higgins and K. Kram, "Reconceptualizing mentoring at work: A developmental network perspective," *The Academy of Management Review*, vol. 26, no. 2, 2001. URL: http://www.jstor.org/stable/259122.