STOCKHOLM SCHOOL OF ECONOMICS
Department of Economics
5330 Master's thesis in Economics
Spring 2026

Not All AI is Created Equal: AI Adoption and Firm-Level Productivity in Sweden

Elise Ramsay Veljanovska (42878)

Abstract This study describes the development of AI adoption in Swedish firms between 2021 and 2025, and explores the associations between AI adoption and total factor productivity. Firm-level survey data from Statistics Sweden are used and three technology-specific AI indicators, text mining, natural language generation, and process automation, are regressed on total factor productivity in a two-way fixed effects model. The productivity analysis shows modest, but consistently positive, associations for all three technologies. Additionally, the different technologies show heterogeneity patterns that align well with the descriptive analysis. The language-based technologies show stronger productivity associations in service sectors, while process automation instead shows larger associations among large goods-producing firms. This suggests that associations are dependent on the specific AI technology being adopted, and on the firm in which it is used. The results therefore also highlight the limitations of measuring effects of AI as a single homogeneous technology.

Keywords: Artificial intelligence, Total factor productivity, Text mining, Natural language generation, Process automation

JEL: D24, O33, L86

Supervisor:	Sampreet Singh Goraya
Date submitted:	18/05/2026
Date examined:	29/05/2026
Discussant:	Simone Michele Bartlett Paul
Examiner:	Patrick Agte

Acknowledgements

I would like to thank my supervisors, Sampreet Singh Goraya at the Stockholm School of Economics, and Samuel Hederén, at Statistics Sweden, for helpful support, feedback and suggestions on this thesis. I would also like to thank the employees at the Innovation, Business sector production and Research section at Statistics Sweden, and its manager, Björn Forssell, for their engagement with this project and for creating a welcoming and stimulating work environment during my time there. Finally, I would like to thank my classmate Smrithi Ganesh Kanna for proofreading this thesis and for being a source of support throughout the writing process.

Contents

List of Abbreviations

AI	Artificial Intelligence
DESI	Digital Economy and Society Index
GMM	Generalized Methods of Moments
GPT	General Purpose Technology
ICT	Information and Communication Technology
LLM	Large Language Model
LP	Levinsohn and Petrin
ML	Machine Learning
NACE	Nomenclature of Economic Activities
NLG	Natural Language Generation
NLP	Natural Language Processing
OLS	Ordinary Least Squares
OP	Olley and Pakes
PA	Process Automation
PSM	Propensity Score Matching
R&D	Research and Development
RCT	Randomized Controlled Trial
RPA	Robotic Process Automation
SCB	Statistiska centralbyrån
SEK	Swedish Krona
SMD	Standardized Mean Difference
SME	Small and Medium-sized Enterprises
TFP	Total Factor Productivity
TM	Text Mining
TWFE	Two-Way Fixed Effects

1 Introduction

Firms are increasingly adopting various Artificial Intelligence (AI) technologies with the goal of increasing efficiency and quality of work, either through the automation of tasks otherwise performed by humans, or letting the technology function as tools to improve the performance of workers. The recent development of AI has resulted in a discussion on if AI truly fulfills its intended purpose on the aggregate level. Brynjolfsson et al. (2017) point out that despite advances in AI, measured productivity gains remain limited, which reflects the lag between technological innovation and its observable impact on economic output, first recognized by Solow (1987) during the computer revolution. It is in this disconnect, between anticipated and observed productivity gains, that this study finds its motivation.

Firm-level productivity associations with AI have been well studied during the 2020s. However, data have been limited, which has led to compromises with regard to how AI is measured, and explored time periods. Most existing evidence relies on aggregated measures that treat AI as a single phenomenon, and on data predating the large expansion of AI observed during the 2020s. As AI capabilities diversify into technologies with fundamentally different purposes and mechanisms, these limitations become increasingly consequential: the productivity implications of specific, recently emerged technologies remain largely undocumented at the firm level. This lack of updated and specific research constrains both managerial decision-making and the design of policies aimed at fostering digitalisation. The purpose of this study is to describe how firm level AI adoption in Sweden, one of Europe's leading countries in regard to technological development and digitalisation, has evolved during the 2020s, and what it looks like today. Additionally, it explores how adoption of specific technologies is associated with differences in productivity. Specifically, it aims to document total factor productivity (TFP) associations with three technologies, one of the most widely adopted AI technologies in Sweden five years ago, AI-based process automation (PA), and two that have evolved significantly and seen the largest increases in adoption since then, text mining (TM) and natural language generation (NLG). These objectives can be summarized into two research questions:

(1) How has AI adoption developed among Swedish firms between 2021 and 2025, and which firms, industries, and technologies characterize the Swedish AI landscape?

(2) How does the association between AI adoption and firm-level total factor productivity vary across technologies, firm sizes, and sectors?

The study follows a structure where the descriptive analysis of the Swedish AI landscape provides the foundation for the productivity analysis by identifying which technologies are most relevant to study and how and where they are deployed in practice.

These questions are answered using high quality firm-level survey data from 2021–2025, and detailed financial data, provided by Statistics Sweden (SCB), the national statistical agency of Sweden. Access to these data allows this study to document AI adoption and explore productivity associations during a period of rapid technological development that most previous studies predate. In addition, by distinguishing between specific AI technologies, the analysis provides actionable insights for firms by identifying which technologies are more consistently associated with productivity increases and under which conditions these patterns are most pronounced. The findings also provide policy-makers with more granular empirical evidence on how productivity associations vary across specific technologies, sectors, and firm types, which can motivate more targeted digitalisation initiatives.

For the descriptive analysis, all available survey data on AI are used and weighted to resemble the true population with regard to firm sizes, industries and regions. In the productivity analysis, a smaller sample of firms between 2021-2024 is utilized to estimate the TFP which is then used as the primary productivity measure of this study. The three AI technologies are found to have modest, but heterogeneous associations with productivity, which correlate differently with the size and sector of the adopting firm.

Limitations to these results include potential measurement errors in self reported survey data, as well as a focus on short term productivity effects, while long term effects are left unexplored. Moreover, the timing of estimated effects is uncertain due to a mismatch between the reference periods of the AI adoption survey and the financial data. Additionally, despite including controls and robustness checks to minimize it, the estimations are affected by residual endogeneity, as unobserved firm level factors could influence both AI adoption and productivity changes. The estimates should therefore be interpreted as conditional associations, rather than causal effects. Documenting these associations is nonetheless informative, as consistent patterns across technologies, sectors, and model specifications can indicate where productivity differences are most pronounced, even in the absence of causal identification.

The rest of this study is structured as follows: Section 2 elaborates on AI and the specific technologies examined in this study, as well as potential productivity mechanisms. Section 3 summarizes the existing literature in the relevant field of study, and identifies the gaps relevant to this analysis. Section 4 introduces the data. Section 5 provides the empirical framework and discusses its challenges. Section 6 shows the results related to both the descriptive and inferential parts of the assessment, as well as the heterogeneity analysis and robustness checks. Section 7 discusses the findings, their implications, and their limitations, and Section 8 concludes.

2 Background

This section defines AI and the three technologies examined in this study, elaborates on how they have evolved during the 2020s, and discusses the theoretical motivations behind their potential associations with productivity and how these may differ between technologies. Additionally, it provides the broader context of the AI climate in Sweden, which is the setting of this study.

2.1 AI and its Applications

AI refers to the ability of machines to act rationally, reason from information, or take actions designed to achieve the best expected outcome (Russell & Norvig, 2010). Rather than a single unified technology, AI is a broad family of methods and applications, each designed to automate different aspects of human cognitive work. This study focuses on three such applications: TM and NLG, both of which fall under the broader umbrella of natural language processing (NLP), and PA.

NLP is a field within AI that focuses on machines understanding and extracting information from human language text or speech (Russell & Norvig, 2010). Two of its central functions are the ability to extract meaning and insight from natural language text (here referred to as TM), and the ability to produce natural language text (here referred to as NLG) (Russell & Norvig, 2010). These technologies are often applied in business contexts such as assisting customer service in written customer communication (Brynjolfsson et al., 2025), searching for information in internal documents (Cambon et al., 2023), and monitoring news flows for market-relevant insights (Supriyono et al., 2024).

Process automation refers to a technology that automates rule-based tasks, often by transferring information between systems (Afrin et al., 2025). This has been applied to perform tasks such as loan processing, registering and posting invoices in account payable systems, and automating administrative functions in health services (Afrin et al., 2025).

In 2020, OpenAI released GPT-3, a large language model (LLM) capable of performing a broad range of language tasks based on human instructions (Brown et al., 2020). Since then, the capabilities and adoption of AI have accelerated. Particularly, recent advances in text mining have been driven by transformer based language models such as BERT and OpenAI's GPT-models, which enable improved extraction of contextual meaning from large volumes of unstructured text (Gardazi et al., 2025). Generative AI tools such as ChatGPT and its competitors have significantly lowered the cost of AI outputs and increased adoption among firms (Maslej et al., 2025; Microsoft, 2023; OpenAI, 2025). Additionally, process automation has evolved from more rule based robotic process automation (RPA), to intelligent process automation that supports more adaptive decision making workflows (Afrin et al., 2025). The recent development of AI across sectors and functions has led researchers to characterize it as a general purpose technology (GPT) in the sense of Bresnahan and Trajtenberg (1995). That is, a technology with broad applicability that has the potential to improve over time and creates innovation opportunities where it is adopted (Brynjolfsson et al., 2017).

2.2 AI and Productivity

In the empirical literature, the relationship between AI and productivity often translates to AI as an intangible capital asset that expands a firm's productive capacity beyond what is explained by traditional inputs (Brynjolfsson et al., 2017; Czarnitzki et al., 2023). This theoretical framework motivates the focus on TFP as a relevant outcome to examine in relation to AI adoption. This framework is applied in this study as well. In this view, firms that adopt AI should, all else equal, achieve a higher output per unit of input, that is, higher TFP.

Agrawal et al. (2018) suggest that AI fundamentally affects a company's productivity by decreasing the cost of prediction, and as a result, decision making under uncertainty improves, it enables automation of tasks that previously needed human judgment and allows for restructuring of workflows to further take advantage of predictive capacity. These mechanisms are all applicable to AI in general, but each of them might be more or less relevant to individual technologies. While process automation is most closely related to pure automation, where routine, rule-based tasks are automated with limited human intervention (Afrin et al., 2025), NLP technologies are relevant for automation, but also play an increasingly assistive role, where AI provides information that humans use as input for decision making (Brynjolfsson et al., 2025; Cambon et al., 2023). This distinction suggests that productivity associations may vary systematically across technologies and the contexts in which they are deployed, which motivates studying them separately rather than as an aggregate measure of AI adoption.

However, the productivity effects of AI adoption may not be immediately observable. Brynjolfsson et al. (2017) argue that AI, as a GPT, is likely to follow the pattern of previous GPTs such as electrification and information technology, where productivity gains materialized only after substantial delays. They identify two main sources of this delay: the time required to build a sufficient stock of the new technology, and the time needed to develop and implement complementary investments in processes, human capital, and organizational practices. During this transition period, measured productivity may even decline as resources are devoted to building intangible capital that is not yet captured in output statistics.

2.3 AI in Sweden

Sweden has been a country of quick technological development during the 2020s. The country ranks among the top countries in Europe on the EU Commission's Digital Economy and Society Index (DESI) and is a global leader in tech entrepreneurship, producing one of the highest numbers of unicorns per capita (Business Sweden, 2025; European Commission, 2022). Institutional support for AI is also strong, for example through the government-appointed AI Commission, which provides guidance on policies to foster AI development in terms of competitiveness, security, and ethics (International Trade Administration, U.S. Department of Commerce, 2024). As a result, Sweden represents one of the more digitally mature economies in Europe, with a business environment characterized by high levels of technology adoption and digital infrastructure.

3 Literature Review

This section provides an overview of the literature on AI and productivity, structured according to how AI adoption is measured across studies. This framework highlights the gaps most relevant to this study, and motivates how access to high quality and detailed recent firm-level panel data contributes to filling them.

3.1 Existing Literature

In the empirical study of AI adoption, there has been a lack of firm-level data on the use of AI, and scholars have emphasized the importance of comprehensive data (Brynjolfsson et al., 2017; Raj & Seamans, 2019). Because of these limitations, researchers have taken a few different approaches to measuring the use of AI: (i) AI patent data, (ii) job posting and resume based indicators of AI skill demand, (iii) firm level surveys of self reported adoption, where each of these methods capture slightly different aspects of AI activity. Further, the productivity effects of specific AI technologies are not as commonly studied at scale, and evidence here comes primarily from (iv) randomized controlled trials (RCT) and (v) observational case studies, which offer more precise identification, but in narrower settings. Below, previous literature is discussed structured based on these different approaches.

AI Patent Data

One of the most common approaches of measuring the level of AI adoption in companies is to use patent applications as a proxy for AI innovation. Damioli et al. (2021) analyse a global dataset of Small and Medium-sized Enterprises (SME) with a System Generalized Methods of Moments (GMM) estimator and conclude that doubling a firm's number of AI patent applications is associated with a 3% labour productivity increase. They emphasize the low economic maturity of AI and the need for more time to assess productivity gains, especially for larger firms. Alderucci et al. (2020) apply machine learning (ML) to identify the level of AI-relatedness of U.S. patent grants between 1990–2018, and report that a firm's first AI patent relates to 40% revenue growth during the five subsequent years. The authors note that these findings should be interpreted primarily as descriptive rather than causal. This standpoint is common in the field of AI and productivity, as an endogeneity problem is present in most specifications where AI is not instrumented or part of a natural experiment.

Job Posting and Resume-based Indicators of AI Skill Demand

Babina et al. (2024) introduce a new way of measuring firm AI adoption by identifying AI related skills in resumes and job postings, which they match to companies in the U.S.. They find that AI adoption is associated with an increase in sales growth, employment and market valuation. They also conclude that positive effects are concentrated among larger companies, which stands in contrast to Damioli et al. (2021), who exclusively study SMEs. Engberg et al. (2025) study Sweden, and show that firms that are highly exposed to AI increase employment of both workers with and without AI skills. However these firms do not see the corresponding TFP increase that would be expected. Engberg et al. (2025) therefore conclude that AI restructures the way work is organised, rather than increases TFP. This finding might sound contradictory to Babina et al. (2024), but the two studies use different outcome variables and these are associated to increased employment in different ways. If AI adoption gives increases in employment which stand in proportion to increases in output, TFP does not change even if the firm experiences increases in sales growth.

Firm Level Surveys of Self Reported Adoption

Firm level surveys on AI use exist but are relatively recent, which means that any recurring panel surveys are still short in duration. Acemoglu et al. (2022) cross-sectionally analyse the labour productivity effects of AI, robotics, dedicated equipment, specialized software and cloud computing in U.S. firms, based on the 2019 Annual Business Survey. They conclude that all technologies together are associated with an 11.4% increase of labour productivity, but find no significant effect of AI alone. Acemoglu et al. (2022) themselves note that the individual technology coefficients are difficult to interpret given that many of the tested technologies are often adopted jointly. They also comment that AI adoption in firms in 2019 may still have needed more time to take effect. Czarnitzki et al. (2023) use the 2019 German Innovation Survey to estimate the productivity effects of AI with OLS cross-sectionally with an AI dummy, and with an Instrumental Variable. Depending on different specifications, they find a 5.5–33% marginal effect on sales of adopting AI, and 14.6% on value added. Lee et al. (2022) use survey data on 300 Korean high tech ventures' reported AI intensity to find that sufficiently large investments in AI are required to get positive revenue effects, and that this increase is strengthened by complementarities such as data base systems, cloud computing, and R&D. They find positive associations for three individual AI technologies: NLP, computer vision, and machine learning (ML), where NLP gives the largest revenue growth effects, and is the least dependent on complementary technologies. More recently, Aldasoro et al. (2026) use data on over 12,000 European and US firms and find that AI adoption increases labour productivity by 4% on average, with the effect being concentrated among medium and large firms.

Employee Level RCTs

Given the recent developments of Large Language Models (LLMs), especially for technologies like TM and NLG, and the limited availability of firm level data, employee level RCTs have produced some of the most recent empirical findings in the field. Brynjolfsson et al. (2025) explore the use of a conversational assistant among customer service workers, and find a 15% increase in issues resolved per hour, with larger effects for less experienced workers. A similar study is carried out by Noy and Zhang (2023), who assign writing tasks to college educated workers, and randomly give half of them access to ChatGPT. This results in a 40% decrease in task completion time, and an 18% increase in the quality of the work. These examples show clear employee level productivity gains from NLP assistance.

Observational Case-Study Evaluations

In a similar way as with NLP, firm level studies have not been conducted in the area of PA to the same extent as studies of general AI. However, simpler evaluations have been made within firms. Hyun et al. (2021) study the application of a business document RPA, which is used to revise government project proposals. The technology outperformed ten experienced office workers in both time consumption and document quality. Moving beyond pure RPA, Lievano-Martínez et al. (2022) document productivity increases from integrating AI optimization algorithms with PA in a manufacturing context. The combined system reduced the order processing time by roughly 65% and material waste by 30%. This shows how AI additions to rule-based automation can generate efficiency improvements in routine, high-volume processes.

3.2 Main Findings

The use cases presented among the RCT and case study findings align well with the mechanisms discussed by Agrawal et al. (2018). NLP can act as automation if the user lets it perform tasks independently, but also functions as an assistive tool when humans remain in control of decision making. With PA however, tasks are fully executed by the system that replaces humans in repetitive tasks, which is typical for automation. These smaller scale studies find positive productivity effects of both NLP and PA. However, the larger productivity increases that are documented in employee RCTs and case studies compared to firm level studies highlight the need for both perspectives when evaluating a technology. Firm-level studies capture the aggregate effect of AI adoption across the entire organization, including both direct effects in specific tasks and indirect effects through reallocation of labour, complementarities, and organizational adjustments. Employee RCTs and case studies instead isolate settings where AI has its strongest immediate impact, which makes them less suited for understanding how these effects translate into overall firm performance.

The majority of firm level studies document positive associations between AI adoption and productivity, labour productivity, or revenue (Babina et al., 2024; Czarnitzki et al., 2023; Damioli et al., 2021), however magnitudes differ. This variation in itself is informative and reflects differences in what is measured, as well as when and where it is measured, and how AI is defined, and does not necessarily indicate that the studies are contradictory to one another. Some studies report null findings (Acemoglu et al., 2022; Engberg et al., 2025), but the nature of these findings differs. Acemoglu et al. (2022) attribute their null effects to the possibility that gains had not yet materialized by 2019. Engberg et al. (2025) find no corresponding TFP gains despite increased hiring among AI-exposed establishments, which they interpret as consistent with AI augmenting rather than replacing workers. Since their measure captures occupational exposure to AI rather than actual adoption, it remains an open question whether firms that actively adopt specific AI technologies show similar tendencies.

One of the strongest explanations for variation across studies is that productivity associations with AI may only be observable after adoption reaches a certain intensity threshold (Lee et al., 2022), or with a time lag of several years (Babina et al., 2024). The technology may also take time to mature in the economy before effects aggregate to the firm level, which may explain why Damioli et al. (2021) find effects only in the more recent subperiod of their data. The apparent contradiction between Damioli et al. (2021), who find effects concentrated among SMEs, and Babina et al. (2024), who find effects concentrated among larger firms, might reflect differences in what is measured rather than a genuine disagreement: Damioli et al. (2021) capture AI innovation activity through patenting, while Babina et al. (2024) measure adoption intensity through hiring, and these two dimensions of AI engagement need not produce the same size heterogeneity.

3.3 Gaps in the Literature and the Contributions of this Study

Firm-level studies of AI and productivity face three limitations. First, most existing evidence relies on data that predate the large expansion of AI observed during the 2020s, with the majority of firm-level studies exploring data up until 2019. Second, data availability has led to compromises in how AI is measured, and most firm-level studies examine AI as a broad, aggregated concept rather than distinguishing between specific technologies. Third, survey-based evidence is mostly cross-sectional, which limits the ability to exploit within-firm variation over time and exclude time-invariant confounders. Together, these three factors limit this field of study: aggregated measures on older data say little about how specific modern technologies affect productivity, and without panel data it is difficult to separate the effects of adoption from underlying firm characteristics.

Recent studies have begun to individually address some of the limitations mentioned above. Bick et al. (2026) document gaps in AI adoption between Europe and the U.S. and find a significant positive association between industry-level AI adoption rates and productivity growth using data extending up to 2025, but their productivity analysis relies on aggregated national statistics rather than firm-level financial data. Aldasoro et al. (2026) provide causal firm-level evidence using data up to 2024. However, their analysis treats AI as a single aggregated measure and relies on pooled cross-sectional data. Lee et al. (2022) differ from most firm-level studies by exploring effects of three specific AI technologies, but their data are from 2019 and apply to a narrow sample of 300 high-tech ventures in Korea.

As the range of available AI technologies expands and their capabilities diverge, the value of technology-specific evidence on recent data only increases. This study addresses all three limitations simultaneously. It draws on firm-level panel data spanning 2021–2024, which covers the period of quick AI expansion that most existing evidence predates. Furthermore, it distinguishes between three specific AI technologies rather than treating adoption as one homogeneous technology. Lastly, it uses within-firm variation over time through a fixed effects framework, which reduces the influence of time-invariant firm characteristics.

4 Data

To link productivity to AI, two datasets are used. The data on AI use in Swedish firms are obtained from Statistics Sweden's (SCB) annual survey ICT Usage in Enterprises, and the firm-level financial data are collected from SCB's Structural Business Statistics. This chapter elaborates on each of the datasets separately below.

4.1 AI Use Data

From ICT Usage in Enterprises this study uses survey responses on firms' ICT use, as well as registered information about firms' industry classification (NACE Rev.2 codes). Enterprises are legally obliged to answer the survey, and for the years covered in this study, response rates range between 77–83% (Statistics Sweden, 2021, 2023b, 2024, 2025c), which is higher than any comparable studies to my knowledge. As the survey cannot be submitted without answering all questions, item non-response is minimal, and remaining missing data are therefore primarily due to unit non-response (Statistics Sweden, 2025c). SCB does not provide documentation on potential systematic uncertainty related to this missingness. Given the high response rates, no further analysis of missing data has been conducted within the scope of this study. The only item non-responses observed for the variables used in this study are firms that in a filter question answered that none of their employees use a device that gives access to the internet. For these firms, questions about AI and other technologies are filtered out. Over all four years, this goes for less than 20 rows and they were removed by listwise deletion.

The survey is sent to active firms with a Swedish address under the NACE Rev.2 codes 10–63, 68–75, 77–82, and 95.1 (Statistics Sweden, 2025d). This excludes firms in primary industries and mineral extraction, financial and insurance companies, as well as firms in public administration, education, health and social services, culture and leisure, and other service activities (Statistics Sweden, 2025a). The exclusion is a result of prioritization of sectors that are considered more relevant in the context, a focus on private firms, and quality considerations. For firms with 0–199 employees, a yearly stratified sampling is made with Neyman allocation, based on firm size, industry and region, while a census is conducted for firms with 200+ employees (Statistics Sweden, 2025d).

The survey includes questions about AI use in firms for the years 2021, 2023, 2024 and 2025. The ones relevant to this study are multiple choice questions related to which AI technologies the firm uses, in which business areas AI technologies are used, and the acquisition methods used to access AI. The questions refer to the beginning of the year of the survey. This study focuses on three of the AI technologies: TM, NLG, and PA. Their corresponding AI indicator variables are listed and explained in Table 1.

The use of survey data in this study is considered necessary to explore AI in Swedish firms, as alternative methods introduce systematic biases that are particularly problematic in the Swedish context. AI patent data would systematically exclude the majority of Swedish AI-using firms who buy their AI systems commercially or use open source AI instead of in-house development (Statistics Sweden, 2025e). AI skill demand instead risks including firms that aim to start using AI soon, which may be especially harmful in Swedish firms where the most commonly given reason for not using

AI is insufficient expertise (Statistics Sweden, 2025e). Given these circumstances, survey data likely contributes to the least systematic bias in their measurements, despite its potential sources of error.

Table 1: Definition of Variables

Variable	**Definition**
AI adoption indicators	
TM	1 if a firm used text mining in the beginning of the survey year, 0 otherwise. Text mining refers to AI technologies that analyse written language.
NLG	1 if a firm used natural language generation in the beginning of the survey year, 0 otherwise. Natural language generation refers to AI technologies that generate written or spoken language.
PA	1 if a firm used process automation in the beginning of the survey year, 0 otherwise. Process automation refers to AI technologies that automate workflows or assist in decision-making.
Production inputs and output	
ln(Output)	Log of production value of the year, in SEK.
ln(Employees)	Log of number of employees.
ln(Capital)	Log of assets, in SEK.
ln(Material)	Log of purchases of raw materials and other operating and external costs of the year, in SEK.
TFP (ω)	Log total factor productivity of the year, estimated using the Wooldridge (2009) proxy variable approach.

Notes: AI adoption indicators refer to the beginning of the survey year. All monetary variables are measured in SEK and expressed in logarithmic form.

4.2 Production Function Variables

The firm-level financial data used in this study are obtained from SCB's Structural Business Statistics, covering all active Swedish enterprises with NACE Rev.2 codes 01–63, 68–82, and 85–96, fully overlapping the population of ICT Usage in Enterprises (Statistics Sweden, 2025b). The main source of the data is the administrative tax data (SRU) from Skatteverket, but direct data collection is also conducted for larger companies (Statistics Sweden, 2025b). All firms are legally obliged to provide the relevant information to Skatteverket and SCB (Statistics Sweden, 2023a, 2025b), and no imputations have been made. The variables used in this study are the log of production value, used as the outcome variable in a production function, and the logs of material, capital and the number of employees, used as inputs. The variables are listed and explained in Table 1. As of 2026, only data up until 2024 is published, which gives this study a three year panel of overlapping financial and AI use data, spanning over four years, 2021, 2023, and 2024.

4.3 Data handling

The dataset used in the descriptive analysis contains the 15 674 observations (7 842 firms) of AI data that have been sampled through the years 2021-2025. To ensure that the descriptive statistics are representative of the true population of firms, all descriptive analyses are weighted using survey weights provided by SCB. These weights are constructed through stratified random sampling across firm size, industry, and region, and are designed to scale the sample to population proportions (Statistics Sweden, 2025d) .

The dataset for the productivity analysis is limited to the firms and years where AI data and financial data overlap, which leaves 7 239 observations (2 914 firms) for the main analysis. The variables used in the main analysis are presented in Table 1. Because of the sampling methods used for the AI data, the sample consists of two subsamples:

1. An unbalanced panel of 3 262 observations containing mostly smaller companies (<200 employees) that have been part of two of the yearly stratified samples, as well as larger companies that have not been covered in all three years.
2. A balanced panel of 3 977 observations containing mainly larger companies (200+ employees) that have been part of the yearly census, as well as smaller companies that have been part of the yearly stratified samples for all three years.

The firms that make up the unbalanced panel are only those that are observed in more than one year during 2021-2024, to ensure identifying variation. This introduces selection that the survey weights are not designed to account for. The regression analysis is therefore conducted without weights. Results from the regression analysis should thus be interpreted as conditional associations for the observed panel of firms, rather than population-representative estimates.

In the final dataset of 7 239 observations, rows containing logically implausible values have been deleted, and further outliers are adjusted through Winsorizing at the 0.5th and 99.5th percentiles, so that observations outside this range were replaced by the respective percentile values. All firm-level data used in this study are treated as confidential and reported exclusively in aggregated form. As a result, individual firms cannot be identified in any published or shared output.

5 Methodology

The empirical framework can be described as a two step procedure, where the second step is the central part of this study. First, TFP, which is not directly observed, must be estimated for every firm and year in the data. A simultaneity bias complicates this estimation and requires a structural approach, which is addressed using the proxy variable method of Wooldridge (2009). Second, the estimated TFP is used as the outcome variable in a Two-Way Fixed Effects (TWFE) regression, where it is regressed on binary indicators of AI adoption.

5.1 Estimating the Total Factor Productivity

The empirical analysis begins by estimating the TFP of each observation in the data. To do this, a standard Cobb-Douglas production function is assumed at the firm level, where output is produced as a multiplicative function of inputs with constant elasticities (Cobb & Douglas, 1928):

$$Y_{it} = A_{it} K_{it}^{\beta_K} L_{it}^{\beta_L} \tag{1}$$

where Y_{it} is output, K_{it} capital, L_{it} labour, and A_{it} represents total factor productivity. The parameters β_K and β_L represent the output elasticities of capital and labour respectively, and their sum measures returns to scale in production. Taking logs gives the standard linear specification:

$$\ln Y_{it} = \beta_K \ln K_{it} + \beta_L \ln L_{it} + \omega_{it} + \varepsilon_{it} \tag{2}$$

where ω_{it} is unobserved firm-level productivity known to the firm but not to the econometrician, and ε_{it} is a classical error term. Firms are assumed to observe ω_{it} before choosing their inputs, making the input choices correlated with the unobserved productivity, which causes OLS estimates of the inputs to be inconsistent Ackerberg et al. (2015). There are several proposed estimation approaches to avoid this simultaneity bias, some of the most common ones being Olley and Pakes (1996) (OP) and Levinsohn and Petrin (2012) (LP). The Wooldridge (2009) proxy-variable approach takes these methods a step further by estimating the production function in a single-step GMM framework, which is particularly useful in short panels with limited within-firm variation in labour (Ackerberg et al., 2015; Wooldridge, 2009).

In the framework of Wooldridge (2009), intermediate inputs (in this case materials) serve as the proxy for unobserved productivity, ω_{it}, as they are inputs that are chosen after ω_{it} is observed:

$$m_{it} = f(k_{it}, \omega_{it}) \tag{3}$$

By inverting this demand function of observable inputs, ω_{it} can be recovered non-parametrically as a function of observable inputs (Wooldridge, 2009):

$$\omega_{it} = h(k_{it}, m_{it}) \tag{4}$$

This recovered productivity is then incorporated into the production function along with the other inputs, and combined with moment conditions that exploit the dynamic structure of productivity (Wooldridge, 2009). Thereby, the proxy function and production function parameters are jointly estimated at once, which distinguishes Wooldridge from OP and LP (Wooldridge, 2009).

Given this framework, the model relies on four key assumptions: (i) material can serve as a proxy for unobserved productivity, (ii) strict monotonicity, meaning that the material demand function is strictly increasing in ω_{it}, which results in invertibility, (iii) ω_{it} is a scalar unobservable, meaning it is the only hidden factor entering the proxy equation, and (iv) capital is decided one period before output, while materials are chosen after ω_{it} is realized, ensuring that the proxy is conditionally independent of the innovation u_{it} (Levinsohn & Petrin, 2012; Wooldridge, 2009). The four key assumptions listed above imply the following moment conditions: the residual in the production function is uncorrelated with current inputs given the proxy, and the productivity innovation is uncorrelated with past capital (Wooldridge, 2009):

$$E[\varepsilon_{it} \mid k_{it}, m_{it}] = 0, \qquad E[u_{it} \mid k_{i,t-1}] = 0 \tag{5}$$

These conditions are required for consistent estimation of the production function parameters.

To confirm assumptions (i) and (ii), it is checked if materials contain variation related to output beyond that explained by capital and labour, and if the relationship between materials and output is monotone. Specifically, the residuals of $\ln(\text{mat})$ are regressed on $\ln(\text{prod})$ to show residual variation, and $\ln(\text{prod})$ is visualized against $\ln(\text{mat})$ by year to show monotonicity. The results, presented in Appendix A, show a positive and approximately monotone relationship in Figure A1, and indicate that there is residual variation left in material, even after controlling for the observed productivity, as presented in Table A1. This supports the use of materials as a proxy for unobserved productivity. Assumption (iii) is supported by the standard Cobb-Douglas specification adopted here, where all relevant inputs are explicitly included in the production function. Regarding assumption (iv) about the timing of the different inputs, the irregularity of the data should be discussed. The time interval between year 1 (2021) and year 2 (2023) is two years, while the interval between year 2 (2023) and year 3 (2024) is only one year. This means that one period does not refer to a consistent time interval, which could potentially weaken the interpretation of capital as a strictly predetermined state variable. Though, capital is nonetheless a slow-moving input, suggesting that the predetermined timing assumption is unlikely to be severely violated even over the longer interval. Because materials are flexible inputs that can adjust quickly to observed productivity, the irregular timing of the panel does not compromise the assumption that materials are set after ω_{it} is realized.

Estimations are made in R with the `wooldridge()` function, which can be found in the package `estprod`. The model is run with bootstrapped standard errors, 200 repetitions, to increase the robustness of the estimations, given the short and unbalanced panel. Bootstrap sampling is made on the firm level, which results in standard errors effectively clustered at the firm level. Wooldridge (2009) recommends estimation of $\omega_{it} = h(k_{it}, m_{it})$ with polynomials of lower degrees that allow for non-linearities but still avoid overfitting. In this study, the capital function is approximated with a cubic polynomial, and material with a quadratic polynomial, as materials are often more proportional to output while capital might be a more complex relationship.

5.2 Estimating the TFP associations with AI

After estimating the production function, the log TFP, $\hat{\omega}_{it}$, is extracted as the residual of the first-stage Wooldridge (2009) estimator:

$$\hat{\omega}_{it} = \ln Y_{it} - \hat{\beta}_K \ln K_{it} - \hat{\beta}_L \ln L_{it} \tag{6}$$

The estimated TFP measure $\hat{\omega}_{it}$ contains not only the firm-level productivity component ω_{it} but also the idiosyncratic error term ε_{it} from the first-stage production function estimation. This measurement error in the dependent variable does not bias the TWFE coefficients, but introduces additional noise in the outcome variable that may contribute to imprecision of estimated AI associations. TFP is then regressed on AI adoption in a TWFE specification. The estimated baseline specification is:

$$\hat{\omega}_{it} = \alpha_i + \alpha_{jt} + \beta \cdot \mathrm{AI}_{it} + \nu_{it} \tag{7}$$

The firm fixed effects α_i absorb all time-invariant firm characteristics, such as managerial ability or organizational culture, that may simultaneously influence AI adoption and productivity. Industry$\times$year fixed effects $\alpha_{j(i)t}$ control for industry-specific trends and common shocks within each sector and year. AI_{it} is a binary indicator equal to one if firm i has adopted the relevant AI technology in period t. The coefficient of interest β is identified from within-firm variation in AI adoption over time, net of industry-specific trends. As $\hat{\omega}_{it}$ relates to the full survey year, and the AI indicator refers to the start of the same year, β will relate to productivity associations which can be observed approximately within a year of the first reported adoption, net of fixed effects. Estimations are conducted in R using the function `feols()` in the `fixest` package. Standard errors are clustered at the firm level to account for serial correlation in the error term within firms over time.

For the TWFE coefficient to have a causal interpretation, the identifying assumption requires that there are no firm-specific, time-varying shocks that simultaneously influence both AI adoption and TFP growth. A related requirement is the parallel trends assumption, which requires that treated and control firms would have followed similar TFP trajectories in the absence of AI adoption. From this follows the no-anticipation assumption, which holds that firms do not alter their behavior prior to the measured adoption date. The plausibility of these assumptions, and the steps taken to examine potential violations, are discussed in the following section.

A further consideration concerns the composition of the treatment and control groups over time. Firms that already adopt AI at their first observation in the panel, hereafter referred to as early adopters, contribute to no within-firm variation to the AI coefficient. Rather than being treated observations, they enter the estimation as part of the control group for firms that adopt during the panel period, hereafter referred to as late adopters. This is a concern specific to TWFE estimators with staggered adoption, where already-treated units in the control group can bias estimated associations, with the direction depending on how their productivity evolves relative to non-adopters during the panel (Callaway & Sant'Anna, 2021). For this reason, results are presented both for the full sample and for a restricted sample excluding early adopters. The latter provides a cleaner comparison between late adopters and firms that never adopt during the panel period.

As three AI technologies are examined simultaneously, the probability of obtaining at least one statistically significant result by chance increases. Bonferroni-corrected p-values are therefore reported for the main specification in the notes of Table 4, where all technologies are estimated on identical samples. Such corrections are not applied to the restricted or matched specifications, as these rely on partially different samples across technologies. More generally, inference is based primarily on the consistency of coefficient patterns across technologies, specifications, and outcome measures rather than on individual significance thresholds alone.

5.3 The Endogeneity of AI Adoption

The AI adoption variable is not instrumented, and the estimated TWFE coefficients may therefore be subject to endogeneity bias arising from the non-random selection into AI adoption. As documented in Table 3, firms that adopt AI during the panel period are systematically larger and more productive already at baseline, confirming that selection into AI adoption is non-random. Specifically, if more productive firms are both more likely to adopt AI and experience faster productivity growth for reasons unrelated to AI, the estimated coefficient will partly capture pre-existing differences between adopters and non-adopters.

These concerns are partially addressed in two ways. First, a placebo test is conducted by replacing the AI indicator with its one-period lead. An insignificant lead coefficient provides partial reassurance that the timing of adoption is not systematically driven by pre-existing productivity trends or anticipatory effects. Second, the baseline model is re-estimated using propensity score matched control groups (Rosenbaum & Rubin, 1983), restricting the comparison to firms that are observably similar to adopters at baseline. This reduces concerns about level differences in productivity between treated and control firms driving the results. Both specifications are described in further detail in the following section.

Despite these measures, unobserved time-varying factors cannot be ruled out. Firms undergoing broader digital transformation or organizational restructuring may simultaneously adopt AI and experience productivity gains for reasons not captured by the fixed effects structure. Moreover, management quality and other factors that likely drive both AI adoption and productivity are unobservable in the data. As previously mentioned, the estimated coefficients should therefore not be interpreted as causal effects but rather as conditional associations between AI adoption and TFP, net of firm-invariant characteristics and industry-specific time trends.

5.4 Model Variations and Robustness Checks

The specifications estimated in this study are listed and explained below.

Baseline Model

The baseline models are TWFE specifications where log TFP is regressed on the different AI technologies. It is run both without and with fixed effects on the full sample. Additionally, as stated above, it is run with a restricted sample of only late adopters and non-adopters, to achieve a cleaner control group with no early adopters.

Baseline Model with Propensity Score Matching (PSM)

To run the baseline model with a matched control group, the sample is restricted to late adopters and non-adopters. Firms which had already adopted AI at the start of the panel, early adopters, must be excluded as baseline values are needed to equalize the groups before adoption. Using all firms' first observation (in 2021 or 2023), non-adopters, the control groups, are matched to the late adopters, the treatment groups. The matching is made with regard to propensity scores, which represent the probability of a firm to adopt AI, given observable variables (Rosenbaum & Rubin, 1983). Each treated firm is matched to up to three control firms with the most similar estimated propensity scores. The choice of three matches per treated firm was made because this number yielded the best covariate balance in the matched sample, while still retaining a sufficient number of matched observations for reliable estimation.

Propensity scores are calculated given industry, $\ln(\text{kap})$, $\ln(\text{emp})$ and $\ln(\text{mat})$. For the matching to be successful, there must be an overlap between the treatment and control groups in the distribution of estimated propensity scores. This means that for each treated firm, there must be at least one comparable control firm with a similar probability of adoption. The quality of the matching is evaluated by comparing summary statistics between treated firms, the full control group, and the matched control group in Appendix section A.2.

The purpose of this specification is to reduce any selection bias on observable variables by restricting the comparison to firms that were observably similar before adoption. This specification tests whether the associations observed in the baseline model persist when adopters are compared to firms with similar pre-adoption characteristics. The resulting estimates should still be interpreted as conditional associations, but with a more comparable control group than in the baseline specification.

Heterogeneity Analysis

To explore which types of firms drive the associations observed in the main analysis, heterogeneity analyses are conducted along two dimensions: firm size and broader sector classification. Firms are classified as large if they have 200 or more employees at their first observation, and small otherwise. As this threshold aligns with SCB's sampling design, where firms with 200 or more employees are subject to a census while smaller firms are drawn in a stratified random sampling, splitting the analysis along this dimension ensures that each subgroup is internally more homogeneous in terms of sampling structure. This mitigates concerns that the pooled estimates in the main analysis are disproportionately driven by large firms, which are observed in all years by design.

The sector classification distinguishes between goods-producing industries (manufacturing, energy, water and waste, and construction) and service industries (retail and wholesale, transport, hotels and restaurants, IT and media, real estate, professional services, support services, and other services). A finer industry-level heterogeneity analysis was considered but not pursued, as the resulting subsamples would be too small to support reliable estimation, and because the industry fixed effects already absorb industry-specific productivity trends in all main specifications.

For each subgroup, results are presented for the full sample and for the restricted sample excluding early adopters. PSM results are not reported for the subgroup analyses, as PSM is estimated separately within each subgroup and the quality of the matching cannot be verified without full balance diagnostics. Given the smaller subsamples involved, there is a non-trivial risk that matching quality decreases, which could make PSM results misleading rather than informative. The results should therefore be interpreted as exploratory rather than confirmatory evidence.

Robustness Checks

Placebo tests are conducted by inserting a lead AI indicator into the base model. The purpose of this is to reveal any positive pre-trends or anticipation effects that might suggest that firms were already on a higher productivity path before adopting AI. If the lead AI variable is positive and significant, this would indicate that productivity increases came before the observed adoption of AI. Such a pattern implies that the estimated AI coefficient may partly capture underlying trends or investments correlated with future adoption. Because the panel is irregular (2021, 2023, 2024), the lead variable refers to future adoption at different time horizons: in some cases one year ahead and in others two years ahead. This means that anticipation effects may be captured over slightly different time spans across observations, which may attenuate the estimated lead coefficient. Nevertheless, the placebo specification remains informative as a test for systematic pre-trends preceding AI adoption.

As a second robustness check, AI is regressed on an alternative output measure, log labour productivity, $lp_{it} = \ln Y_{it} - \ln L_{it}$. This measure requires no first step Wooldridge estimation, and is recovered indirectly though measures already available in the data. As a result, this estimation is not dependent on the assumptions about one single scalar observable, monotonicity or material as a valid proxy. Seeing a similar relationship between the AI indicators and a measure that is based on different assumptions would strengthens confidence that the documented patterns reflect genuine productivity differences and are not only a result of the production function specification. Capital is not included as a control variable in this specification. In a fixed effects setting, firm fixed effects absorb time-invariant differences in capital intensity which makes an explicit capital control redundant. Additionally, capital may itself be endogenous to AI adoption, as firms investing in AI often make complementary capital investments simultaneously. This approach follows Acemoglu et al. (2022), who estimate labour productivity regressions with fixed effects but without explicit capital controls. The Wooldridge (2009) TFP estimator in the main analysis addresses capital endogeneity more directly.

The baseline specification, PSM, and placebo test form a descriptive framework for documenting productivity associations that are robust to several alternative explanations, while acknowledging that causal identification requires variation in AI adoption that is independent of firm-level productivity trajectories, which is a requirement that cannot be fully satisfied in the present observational setting.

6 Results

The results are presented in four sections. The first describes AI adoption in Swedish firms and motivates the choice of the three technologies explored in the regression analysis. Descriptive figures are shown over time, between the years 2021-2025, but also sometimes cross-sectionally for 2024 as that is the most recent year of the panel used for the productivity analysis. The remaining sections present the main model estimations, a heterogeneity analysis, and robustness checks.

6.1 Descriptive Statistics: the Swedish AI Landscape

Figure 1 shows the yearly adoption rates of all AI technologies that are included as options in the survey. In 2021, about 10% of Swedish firms adopted AI, and in 2025, this number had increased to 35%. The concentration of adoption in the final years of the sample implies a very short post-treatment period and limits the ability to capture long run effects of AI. Even though the general trend is that adoption increases with time, different AI technologies show different adoption patterns. TM and NLG have made steep increases in 2024 and 2025, and speech recognition has experienced a similar but smaller increase. On the other hand, technologies like ML and PA were already relatively highly adopted in 2021, and have not seen as dramatic an increase. The general pattern seems to be that language based technologies have made a quick increase in the last years, while more process based technologies have been around for longer, but have seen more modest and gradual growth over the study period. Due to this clear distinction, this study focuses on and compares technologies from both groups. TM and NLG in their current form represent a largely unexplored area in the productivity literature, as most previous studies use data from before 2021. Simultaneously, PA was the most commonly adopted technology in 2021 and represents the more established end of the spectrum. Including technologies from both groups allows for a comparison of technologies at different stages of maturity on the market.

Figures 2 and 3 show the adoption patterns of different technologies over time, but split on firm size, and sector, in a similar way in which the sample is split in the heterogeneity analysis. From

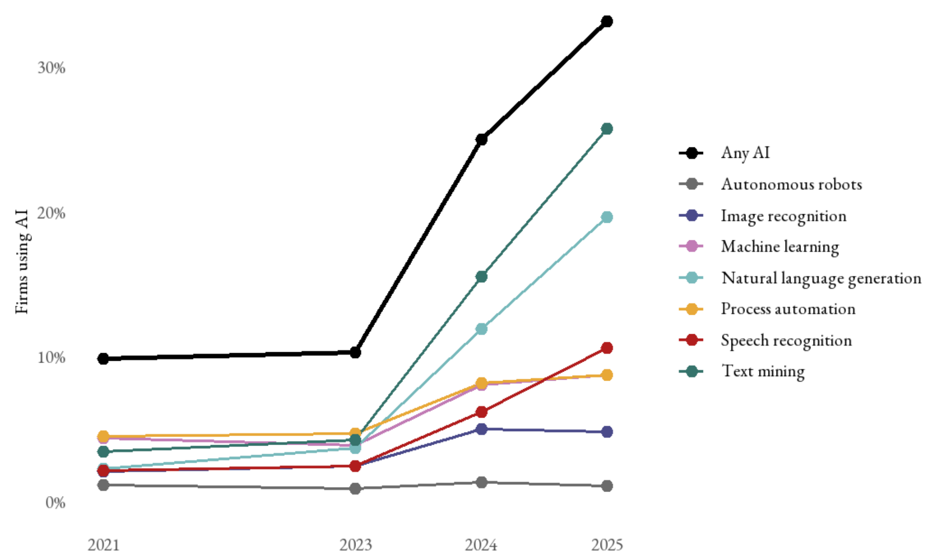

Figure 1: Share of Swedish Firms Using each AI Technology, 2021–2025
Notes: Statistics are weighted using weights provided by SCB, constructed through stratified random sampling across firm size, industry, and region, to produce population-representative estimates. Any AI refers to firms reporting use of at least one listed technology. *Source:* Author's rendering of Statistics Sweden, ICT Usage in Enterprises (2021-2025).

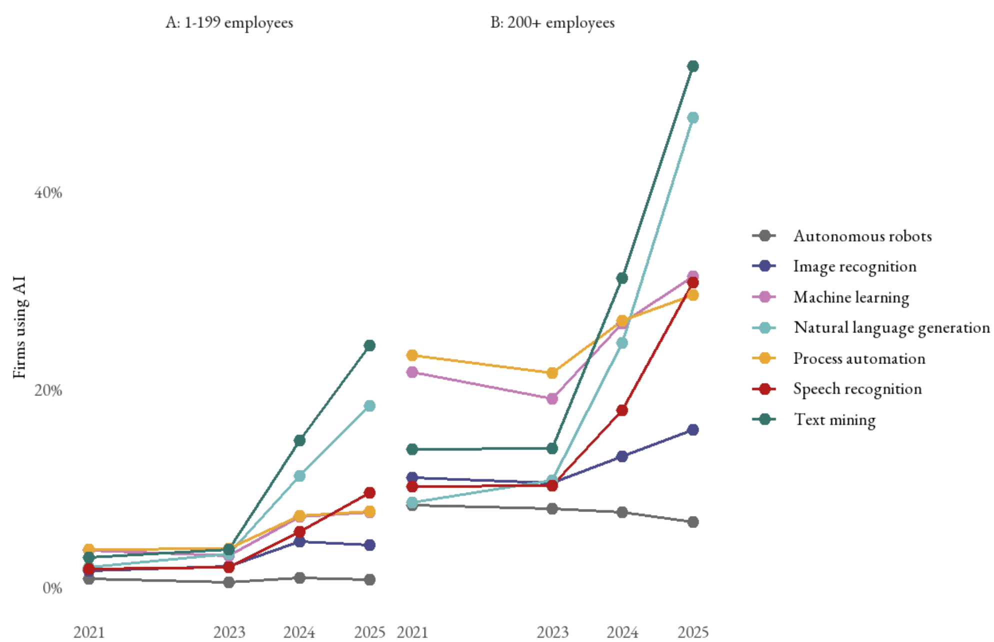

Figure 2: Share of Swedish Firms Using each AI Technology by Firm Size, 2021–2025
Notes: Firms with 200 or more employees are considered large. Statistics are weighted using weights provided by SCB, constructed through stratified random sampling across firm size, industry, and region, to produce population-representative estimates. *Source:* Author's rendering of Statistics Sweden, ICT Usage in Enterprises (2021-2025).

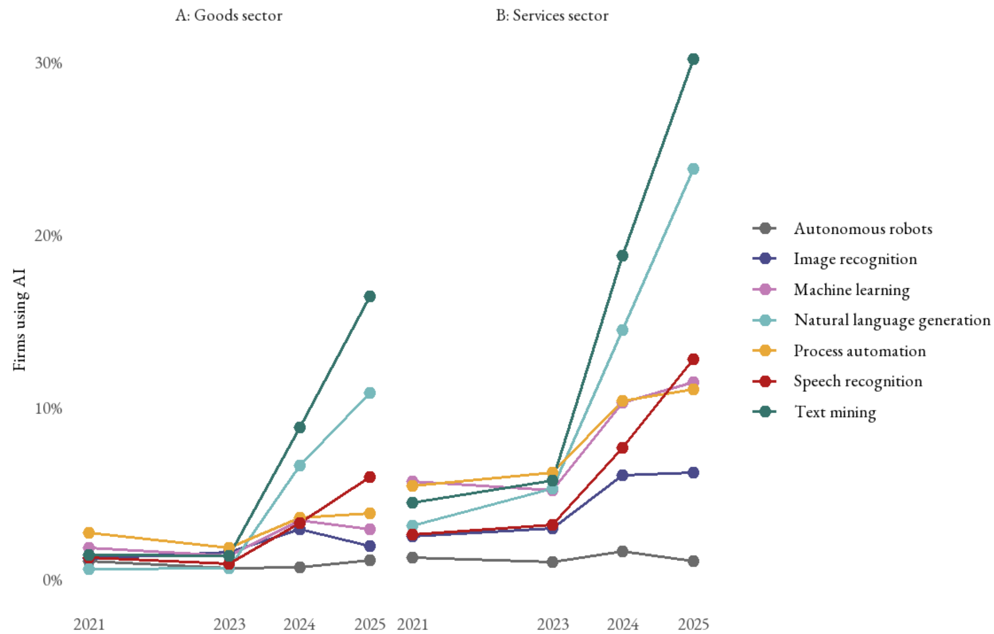

Figure 3: Share of Swedish Firms Using each AI Technology by Firm Sector, 2021–2025

Notes: Goods industries: manufacturing, energy, water and waste, construction. Service industries: retail and wholesale, transport, hotels and restaurants, IT and media, real estate, professional services, support services, other services. Statistics are weighted using weights provided by SCB, constructed through stratified random sampling across firm size, industry, and region, to produce population-representative estimates. *Source:* Author's rendering of Statistics Sweden, ICT Usage in Enterprises (2021-2025).

Figure 2 it is clear that large firms have stood for the majority of adoption, especially before 2023, but also after, and that the quick increase in language based models was steeper in this group. This comparison also highlights a possible difference in sample composition depending on if large or small firms are studied, where a sample of large firms from 2021 would to a higher degree relate to AI as mainly ML and PA, while a sample of small firms would rather have a more mixed group of AI technologies. Figure 3 shows a similar, but smaller difference in adoption between goods and service producing firms, where service firms have a higher adoption rate and steeper increase. This means that early adopters refers to an unproportionate amount of service firms. Another notable difference is that the curves of ML and PA among goods producing firms are close to flat, compared to the development in service firms. Generally, the relative composition of adoption of the different technologies in 2024-2025 looks similar across both the size and sector split, with language based technologies being the absolute most adopted, followed by ML and PA.

To further describe the AI adoption in Swedish firms, Figure 4 shows adoption in each industry in 2024. Adoption varies substantially between different sectors, where adopting firms are concentrated

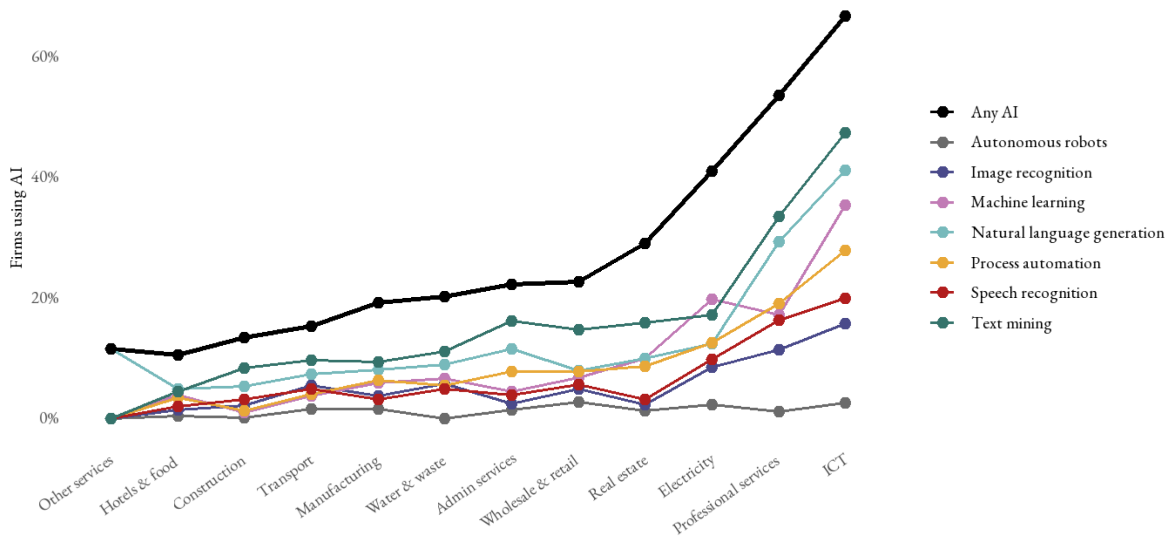

Figure 4: Share of Swedish Firms Using Each AI Technology by Industry, 2024
Notes: Industries classified according to SNI 2007. Statistics are weighted using weights provided by SCB, constructed through stratified random sampling across firm size and region, to produce population-representative estimates. *Source:* Author's rendering of Statistics Sweden, ICT Usage in Enterprises (2024).

in more knowledge-intensive service sectors, such as Information and Communications Technologies (ICT), professional services, and real estate, but also electricity. Non adopters are more common in labour intensive sectors such as hotel, food, construction, and transport. However, the relative ordering of technologies is fairly consistent across sectors. In all sectors but electricity, the most common technology is TM, or NLG. This might suggest that these technologies are applicable over many different areas and functions. After the top two technologies, ML and PA are the third and fourth most common technologies, and ML stands out in being the most used technology in electricity. Less applied technologies in most sectors are speech recognition and image recognition, but most of all, autonomous robots which is close to 0 in all sectors.

Figure 5 shows the acquisition forms used for AI in firms. It depicts a clear shift in the composition of acquisition modes over the study period, where pre-built options such as off-the-shelf and open source AI account for an increasing share, while custom developed solutions made in house or by an external contractor make up a declining share of total AI acquisition.

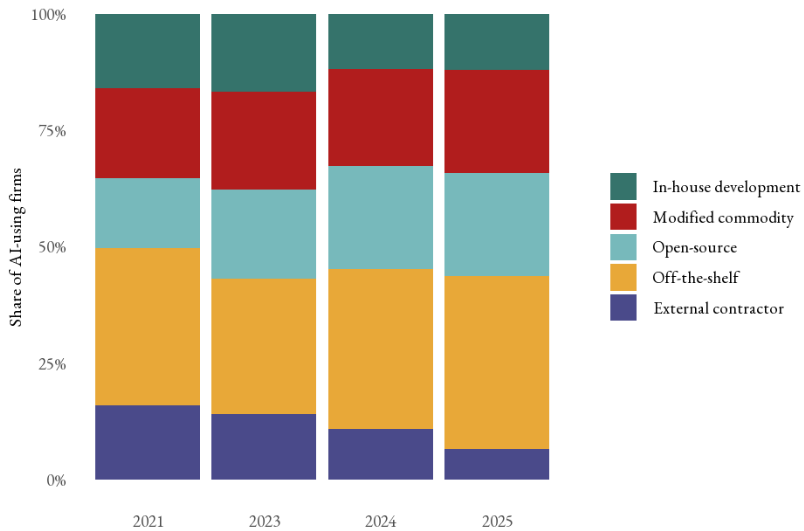

Figure 5: Acquisition Modes among AI-Adopting Swedish Firms, 2021–2025
Notes: Acquisition modes are defined as follows: In-house dev.: developed internally by the firm; Modified comm.: commercially available software modified by the firm; Open-source: open-source software used without modification; Off-the-shelf: commercially available software used without modification; Ext. contractor: developed by an external contractor. Sample restricted to firms reporting use of at least one AI technology. Statistics are weighted using weights provided by SCB, constructed through stratified random sampling across firm size, industry, and region, to produce population-representative estimates. *Source:* Author's rendering of Statistics Sweden, ICT Usage in Enterprises (2021-2025).

Figure 6 visualises in what business areas the adopters of the different technologies use AI. It shows that the most common business functions to use AI in are marketing and sales, administration, production and R&D, while less common areas are logistics, IT security and finance. There is also a relative difference in where the users of each specific technology choose to apply AI. Relative to the other technologies, PA users apply AI in logistics, production and finance more intensively, while TM and NLG users apply AI relatively more especially in marketing and sales, but also R&D. This distinction between language-based and process-based technologies is further reflected in Table 2. It shows that adoption of TM, NLG and speech recognition are more strongly correlated with each other than with process-oriented technologies, while PA and ML show the highest mutual correlation with each other. This suggests that the two groups tend to be adopted by overlapping but distinct sets of firms. The patterns depicted in these descriptive figures motivate the focus on specific technologies rather than aggregate AI adoption in the productivity analysis, as where and how they are adopted differs notably.

Table 2: Pairwise Correlations Between AI Technologies

	TM	SR	NLG	PA	IR	ML	AR
TM		0.489	0.567	0.406	0.324	0.444	0.231
SR			0.517	0.360	0.358	0.372	0.269
NLG				0.354	0.333	0.379	0.241
PA					0.330	0.541	0.339
IR						0.393	0.331
ML							0.317
AR							

Notes: Correlations computed on the full analysis sample. TM = Text mining, SR = Speech recognition, NLG = Natural language generation, PA = Process automation, IR = Image recognition, ML = Machine learning, AR = Autonomous robots. *Source:* Author's rendering of Statistics Sweden, ICT Usage in Enterprises (2021-2024).

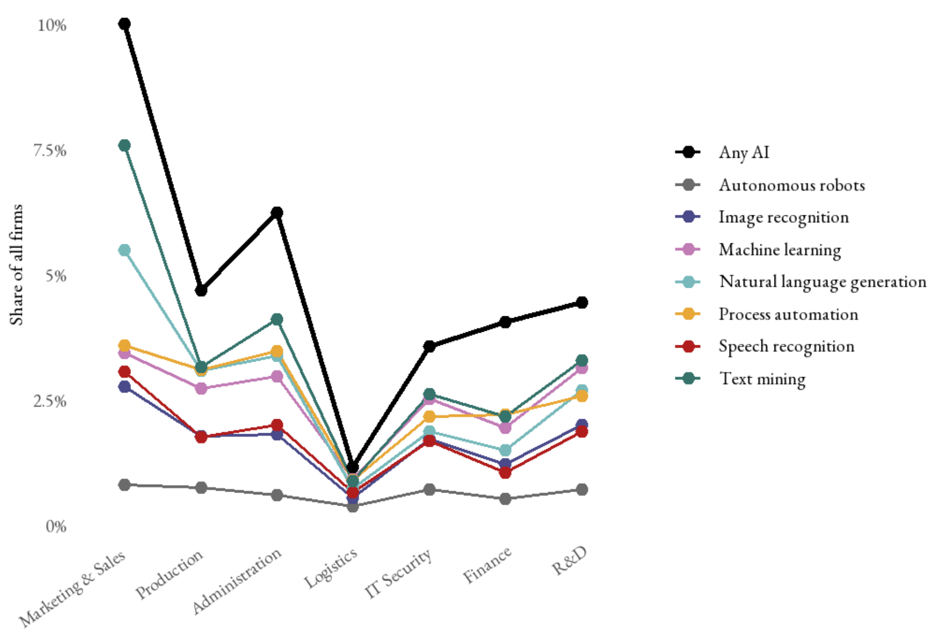

Figure 6: AI Use by Business Function and Technology in Swedish Firms, 2024
Notes: Bars show the share of all firms in the sample that both adopt a given technology and report AI use within each business function. Statistics are weighted using weights provided by SCB, constructed through stratified random sampling across firm size, industry, and region, to produce population-representative estimates. *Source:*Author's rendering of Statistics Sweden, ICT Usage in Enterprises (2024).

6.2 AI Adoption and Total Factor Productivity

Table 3 presents summary statistics for AI adopters and non-adopters at baseline. It shows that late adopters were larger than non-adopters in all measured aspects, even before adopting, averaging more than twice the number of employees, and three times the production value, capital and material inputs. Figure 4 highlights another imbalance, where adopters make up significantly larger fractions of some sectors, and smaller of others. These systematic pre-adoption differences confirm that AI adoption is not random. Larger, more productive firms in knowledge-intensive industries are more likely to adopt AI, regardless of any causal effect of the technology itself. This is precisely the selection problem discussed in Section 5.3. A naive comparison of adopters and non-adopters would mix up the effect of AI adoption with the underlying characteristics of firms that select into adoption.

Table 3: Summary Statistics by AI Adoption Status at First Observation

	Already adopted		Adopts later		Non-adopters	
Variable	Mean	SD	Mean	SD	Mean	SD
N firms	605		582		1727	
Employment	634	1 563	310	1 014	120	241
Production value	2 570	10 281	947	4 105	291	682
Capital	1 516	9 367	390	1 682	100	437
Materials	1 666	7 355	627	3 585	177	491

Notes: Statistics are reported at each firm's first observed year in the panel. Already adopted refers to firms reporting AI use at first observation. Adopts later refers to firms that report no AI use at first observation but adopt during the panel. Non-adopters never report AI use during the panel. All monetary variables are reported in thousand SEK. Employment is reported in number of employees. *Source:* Author's rendering of Statistics Sweden, ICT Usage in Enterprises (2021-2024); Statistics Sweden, Structural Business Statistics (2021-2024).

The results of the production function estimation can be found in Table A2. The labour elasticity is estimated to 0.622, while the capital elasticity is 0.325, and both are highly significant. This results in slightly below constant returns to scale of 0.947, consistent with estimates typically reported in the production function literature.

Figure 7 shows that differences between late adopters and non-adopters persist even when standard inputs are controlled for. Late adopters of all the technologies of interest have a higher average TFP even before adopting. Adopters of the different technologies start off with similar TFPs, but the most extreme differences can be seen between the TM late adopters and non adopters, while NLG late adopters differ the least from their control group. Looking at the change in TFP, it is clear that

AI adopters see larger TFP growth during 2023-2024 than non adopters. This is notable as most of the AI adoption in this panel occurs between 2023 and 2024. No conclusions can be drawn from this raw comparison on its own, as the differences in TFP growth may reflect an underlying selection rather than a causal effect of AI adoption. However, the consistent pattern provides descriptive motivation for the more controlled analysis that follows.

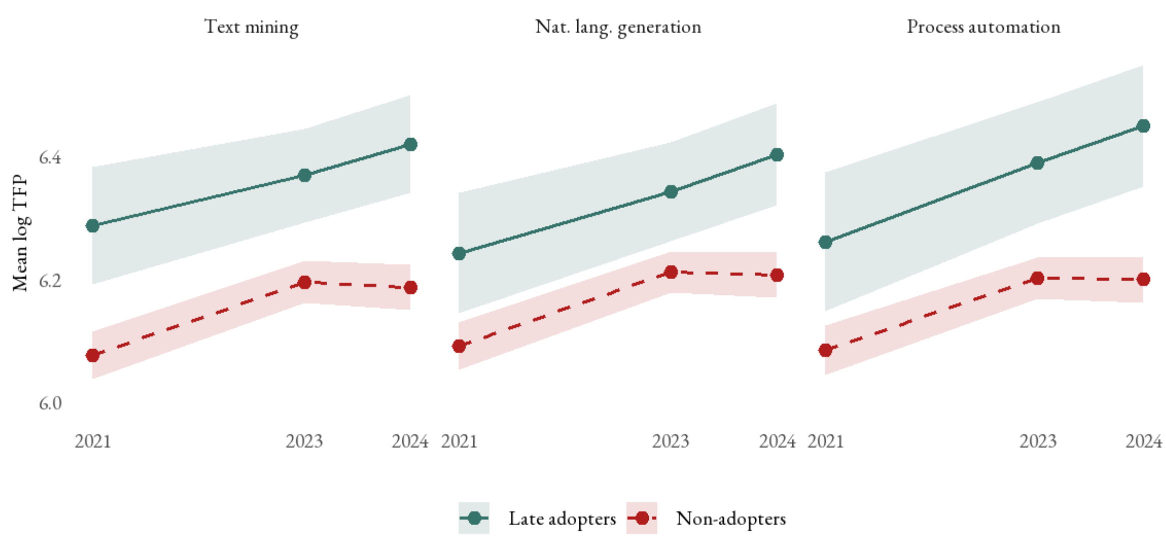

Figure 7: Mean Log TFP for Late Adopters and Non-Adopters by AI Technology, 2021–2024
Notes: Late adopters are defined as firms that adopt the relevant technology at any point during the panel but were not already using it at first observation. Non-adopters are firms that never adopt the technology. Early adopters are excluded. Shaded areas indicate 95% confidence intervals around group means. *Source:* Author's rendering of Statistics Sweden, ICT Usage in Enterprises (2021-2024); Statistics Sweden, Structural Business Statistics (2021-2024).

Given the short panel, the concentration of adoption in the final year of the data, and the relatively small number of adopters per technology, the estimated associations are expected to be imprecisely estimated. Where associations are statistically insignificant, this should be interpreted in light of the wide confidence intervals surrounding the estimates which reflects imprecision rather than necessarily indicating that associations are close to zero. The analysis therefore focuses on the direction and magnitude of associations across technologies, sample compositions, and model specifications, and treats consistency across these dimensions as informative even where individual estimates do not reach conventional significance levels.

The main results from the second part of the estimation are presented in Table 4. Column (1) reports the model estimates for the full 2021-2024 sample without fixed effects. Columns (2)-(4) report the estimation results of the same model with fixed effects, but with three different sample compositions. Column (2) shows the full sample, column (3) shows the step in between the baseline specification and PSM specification, where all early adopting firms have been removed, and column

(4) shows the results where the early adopters have been removed and PSM has been conducted. The purpose of presenting the results like this is to decompose which changes are associated with restricting the sample to only late adopters (change from column (2)-(3)), and which are associated with matching the control group (change from column (3)-(4)). The former reveals how the presence of already-treated firms in the control group affects estimated associations, as discussed in Section 5.2. The latter reflects the degree to which observable selection bias affects the estimates, as the matched control group is restricted to firms with similar baseline characteristics to the adopters.

The substantial reduction in coefficients from column (1) to column (2) confirms the presence of strong positive selection into AI adoption, as documented in the descriptive analysis, and is expected. In column (2), TM has a coefficient of 0.036 and a p-value of 0.016. The Bonferroni-corrected p-value of 0.050 indicates that the association remains just significant at the 5% level after correction for multiple testing. The coefficient decreases from 0.036 to 0.027 when early adopters are excluded in column (3). As discussed in Section 5.2, this shift reflects a change in control group composition rather than stronger associations among early adopters. The association remains positive but is estimated with greater uncertainty, which is consistent with fewer adopters in the restricted sample. The change between columns (3) to (4) is a smaller decrease in coefficient and significance level which might reflect the elimination of some positive selection bias, or greater uncertainty in the estimates due to a further restricted sample.

NLG shows a positive association with TFP across all specifications, with coefficients ranging from 0.030 to 0.036. The estimates are imprecisely estimated throughout, and even more so after Bonferroni correction of the column (2) estimate. The direction and magnitude are nonetheless consistent across sample compositions. When the sample is further restricted with PSM, the coefficient is barely affected, which is an indication that selection on observable baseline characteristics does not substantially drive the estimates.

PA shows a small positive but imprecisely estimated association in the full sample in column (2). However, in column (3), the estimated coefficient grows substantially. This shift reflects the change in control group composition when early adopters are excluded. The estimate does not change considerably when the sample is further restricted with PSM, and amounts to 0.044 with p=0.022. This would indicate that conditional on firm fixed effects, industry-specific time trends, and observable baseline firm characteristics, late-adopting firms that adopt PA are associated with a 4.5% higher TFP compared to observationally similar non-adopting firms.

Table 4: TWFE Estimates of AI Adoption on TFP

	(1)	(2)	(3)	(4)
Text Mining (TM)				
AI indicator	0.266***	0.036*	0.027	0.023
	(0.033)	(0.015)	(0.017)	(0.019)
Natural Language Generation (NLG)				
AI indicator	0.273***	0.030·	0.035·	0.036·
	(0.037)	(0.017)	(0.018)	(0.019)
Process Automation (PA)				
AI indicator	0.220***	0.013	0.042*	0.044*
	(0.034)	(0.016)	(0.019)	(0.019)
Firm FE	No	Yes	Yes	Yes
Industry×Year FE	No	Yes	Yes	Yes
Early adopters excl.	No	No	Yes	Yes
PSM	No	No	No	Yes
Observations	7239	7239	6672–6882	2897–4490

Notes: Dependent variable is log TFP estimated by the Wooldridge (2009) method. Column (1) reports OLS without fixed effects. Column (2) adds firm and industry×year fixed effects. Column (3) additionally excludes firms already adopting AI at first observation (TM: 251, NLG: 162, PA: 356 firms excluded). Column (4) additionally applies PSM with nearest-neighbour matching ($k = 3$) on inputs and industry. Standard errors clustered at the firm level in parentheses. Bonferroni-corrected p-values for the main specification (column (2)): TM 0.050, NLG 0.206, PA 1.000. Significance levels: *** 0.001, ** 0.01, * 0.05, · 0.1. *Source:* Statistics Sweden, ICT Usage in Enterprises (2021-2024); Statistics Sweden, Structural Business Statistics (2021-2024).

6.3 Heterogeneity Analysis

To learn more about the types of firms that drive the associations documented in previous sections, a heterogeneity analysis is conducted with respect to firm size, and broader sector. The purpose of this is not to establish precise effect sizes for each subgroup, but rather to explore which types of firms drive the associations observed in the main analysis. Given the smaller subsamples and the increased number of tests, the results should be interpreted with caution and treated as exploratory rather than confirmatory evidence.

Table 5 shows the regression results with the same decomposition as columns (2)-(3) in Table 4, but split on size. It should be kept in mind that due to how the firms were sampled, this split makes the subsample of large firms an almost perfectly balanced sample, which might contribute to lower standard deviation of the coefficients compared to the sample of small firms which is to a much larger extent unbalanced. However, the sample of smaller firms is larger, with almost 4912 observations in baseline, compared to the 2324 observations of larger firms, which also affects the certainty of the estimates. In Table 5, the coefficient of TM is positive for both firm sizes, but larger for the full sample of small firms (0.062) than for large firms (0.011). Among small firms, the TM coefficient decreases and becomes insignificant once early adopters are excluded. However, for large firms, the pattern is reversed and the coefficient increases when early adopters are removed, though it remains imprecisely estimated. For NLG, the coefficients are all positive, with the clearest pattern among larger firms when early adopters are excluded. Lastly for PA, the coefficient increases from 0.014 to 0.053 for large firms when early adopters are excluded. Among small firms the coefficient also increases but remains imprecisely estimated.

Moving on, Table 6 shows the regression results split by broader sector, where each firm is classified as part of the goods or service industry. For TM and NLG, the positive productivity associations are more apparent in the service industry, while PA has larger coefficients in the goods industry. The patterns observed in the main analysis persist: TM coefficients are larger when early adopters are included, while PA coefficients increase when they are excluded. However the coefficients of NLG do not seem to be as sensitive to the sample restrictions in the service industry as they are in the goods industry.

Table 5: Heterogeneity by Firm Size: TWFE Estimates of AI Adoption on TFP

	Large		**Small**	
	(1)	**(2)**	**(3)**	**(4)**
Text Mining (TM)				
AI indicator	0.011	0.030	0.062*	0.034
	(0.017)	(0.022)	(0.025)	(0.028)
Natural Language Generation (NLG)				
AI indicator	0.028	0.041·	0.033	0.035
	(0.020)	(0.022)	(0.028)	(0.030)
Process Automation (PA)				
AI indicator	0.014	0.053*	0.012	0.040
	(0.020)	(0.027)	(0.025)	(0.028)
Firm FE	Yes	Yes	Yes	Yes
Industry×Year FE	Yes	Yes	Yes	Yes
Early adopters excl.	No	Yes	No	Yes
Adopters (TM)	354	236	349	246
Adopters (NLG)	285	208	275	209
Adopters (PA)	391	176	250	130
Observations	2324	1745–2018	4912	4613–4754

Notes: Dependent variable is log TFP estimated by the Wooldridge (2009) method. Large firms are defined as firms with 200 or more employees at first observation. Column (1) uses the full size-group sample. Column (2) excludes firms already adopting AI at first observation. Adopters refer to unique firms adopting at any point during the panel. Standard errors clustered at the firm level in parentheses. Significance levels: *** 0.001, ** 0.01, * 0.05, · 0.1. *Source:* Statistics Sweden, ICT Usage in Enterprises (2021-2024); Statistics Sweden, Structural Business Statistics (2021-2024).

Table 6: Heterogeneity by Sector: TWFE Estimates of AI Adoption on TFP

	Goods		**Services**	
	(1)	**(2)**	**(3)**	**(4)**
Text Mining (TM)				
AI indicator	0.033	0.014	0.038*	0.034
	(0.026)	(0.029)	(0.019)	(0.022)
Natural Language Generation (NLG)				
AI indicator	0.021	0.034	$0.034^{\cdot}$	$0.036^{\cdot}$
	(0.034)	(0.039)	(0.019)	(0.020)
Process Automation (PA)				
AI indicator	0.013	0.055*	0.013	0.036
	(0.024)	(0.026)	(0.020)	(0.025)
Firm FE	Yes	Yes	Yes	Yes
Industry×Year FE	Yes	Yes	Yes	Yes
Early adopters excl.	No	Yes	No	Yes
Adopters (TM)	198	144	505	338
Adopters (NLG)	147	118	413	299
Adopters (PA)	196	91	445	215
Observations	3085	2803–2939	4154	3558–3878

Notes: Dependent variable is log TFP estimated by the Wooldridge (2009) method. Goods industries: manufacturing, energy, water and waste, construction. Service industries: retail and wholesale, transport, hotels and restaurants, IT and media, real estate, professional services, support services, other services. Column (1) uses the full sector-group sample. Column (2) excludes firms already adopting AI at first observation. Adopters refer to unique firms adopting at any point during the panel. Standard errors clustered at the firm level in parentheses. Significance levels: *** 0.001, ** 0.01, * 0.05, $\cdot$ 0.1. *Source:* Statistics Sweden, ICT Usage in Enterprises (2021-2024); Statistics Sweden, Structural Business Statistics (2021-2024).

6.4 Robustness

As a first robustness check, the current AI indicators in the model specifications are exchanged for lead AI indicators, to test for pre-trends or anticipation effects. The results can be observed in Table 7 and are presented both for the full sample, where potential productivity effects of TM were more evident in the main analysis, and for the restricted sample, where potential positive productivity effects were indicated by the coefficient of PA. None of the lead coefficients are significant. Among the three technologies, NLG has the largest lead coefficient in absolute terms with $p = 0.179$. Importantly, the coefficient is negative, which is not consistent with a pre-existing upward productivity trend that would inflate the main estimates. If anything, it suggests that the main NLG estimates are conservative rather than overstated. With that being said, the overall results provide no clear evidence of pre-trends or anticipation effects.

Table 8 shows the main regression results, but with the alternative productivity measure log labour productivity. The estimated coefficients follow a similar pattern to the main results where all AI indicator coefficients are positive, but significance levels are generally weaker than in the main analysis. TM shows the strongest association with productivity in the full sample, and the coefficient decreases when early adopters are excluded. However, both estimates are insignificant. NLG shows consistently positive but imprecisely estimated associations across both specifications, and for PA, the relationship is positive and weak in the full sample, but increases in magnitude and certainty when early adopters are removed from the sample. The smaller and less significant coefficients are consistent with labour productivity being a noisier productivity measure than TFP. Labour productivity conflates efficiency gains with changes in capital intensity within firms over time, whereas the Wooldridge (2009) estimator explicitly separates the efficiency component from input contributions. Overall, the consistently positive direction of associations is preserved under this alternative productivity measure, which does not rely on the assumptions of the Wooldridge (2009) production function estimator. This consistency across measures that rely on different assumptions strengthens confidence that the documented patterns reflect genuine productivity differences rather than artifacts of the main specification.

Table 7: Placebo Test: Lead AI Adoption on TFP

	(1)	(2)
Text Mining (TM)		
Lead AI indicator	0.018	0.007
	(0.029)	(0.033)
Natural Language Generation (NLG)		
Lead AI indicator	−0.026	−0.047
	(0.029)	(0.035)
Process Automation (PA)		
Lead AI indicator	0.003	0.023
	(0.031)	(0.032)
Firm FE	Yes	Yes
Industry×Year FE	Yes	Yes
Early adopters excl.	No	Yes
Observations	2822	2390–2592

Notes: Dependent variable is log TFP estimated by the Wooldridge (2009) method. Column (1) uses the full sample. Column (2) excludes firms already adopting AI at first observation. Firm and industry×year fixed effects are included in all specifications. Standard errors clustered at the firm level in parentheses. Significance levels: *** 0.001, ** 0.01, * 0.05, · 0.1. *Source:* Statistics Sweden, ICT Usage in Enterprises (2021-2024); Statistics Sweden, Structural Business Statistics (2021-2024).

Table 8: Labour Productivity as Alternative Outcome

	(1)	(2)
Text Mining (TM)		
AI indicator	0.022	0.013
	(0.014)	(0.016)
Natural Language Generation (NLG)		
AI indicator	0.018	0.021
	(0.015)	(0.017)
Process Automation (PA)		
AI indicator	0.022	0.030·
	(0.014)	(0.018)
Firm FE	Yes	Yes
Industry×Year FE	Yes	Yes
Early adopters excl.	No	Yes
Observations	7239	6361–6882

Notes: Dependent variable is log labour productivity, defined as log production value minus log employment. This specification does not rely on the Wooldridge (2009) production function estimator and serves as a robustness check for the main results, following the approach of Acemoglu et al. (2022), who control for industry, size, and age fixed effects without explicit capital controls. In the present study, firm fixed effects absorb time-invariant differences in firm size, age, and capital intensity. Column (1) uses the full sample. Column (2) excludes firms already adopting AI at first observation. Standard errors clustered at the firm level in parentheses. Significance levels: *** 0.001, ** 0.01, * 0.05, · 0.1. *Source:* Statistics Sweden, ICT Usage in Enterprises (2021-2024); Statistics Sweden, Structural Business Statistics (2021-2024).

7 Discussion

Given the short panel and limited variation, the productivity associations with AI presented in this study are generally modest. Nevertheless, the consistent positive direction of the estimates across technologies, sample compositions, and outcome measures is informative in itself. The following discussion evaluates the results based on two criteria: whether associations are consistent in direction and magnitude across technologies, sample compositions, and outcome measures, and whether the heterogeneity in associations across firm types and sectors is consistent with the patterns documented in the descriptive analysis. These criteria provide a basis for interpreting what the documented patterns suggest about the relationship between AI adoption and productivity in Swedish firms.

7.1 Interpretation and Implications of Results

Firm level AI adoption in Sweden has seen a dramatic increase between 2023 and 2025. The development has mainly been driven by two language-based technologies: TM and NLG, which seem to be highly relevant in most sectors, while process based technologies that were the most popular in 2021 have seen a slower increase. Additionally, a shift towards ready-made acquisition solutions suggests that AI adoption is increasingly driven by the availability of accessible general purpose tools rather than firm specific development. The development towards simpler acquisition forms and language based models aligns well with the introduction of generative AI chatbots in 2023, and is consistent with the notion that the barrier to AI adoption has decreased substantially in recent years.

AI adoption is generally larger among larger and more productive firms, and is also disproportionally concentrated among knowledge-intensive service sectors. The consistent dominance of TM and NLG across sectors and business functions aligns with the belief that these technologies function as GPTs, applicable across many different contexts rather than being specific to any particular industry. PA adopters, by contrast, show a more concentrated adoption pattern, particularly in logistics, finance, and production, which indicates a more sector-specific use case. These patterns motivate the focus on specific technologies in the productivity analysis. TM, NLG, and PA represent three technologies with distinct adoption trajectories and use cases, making them well suited for examining whether heterogeneous productivity associations exist across technologies, firm sizes, and sectors.

TM and NLG are relatively understudied at the firm level due to their low adoption prior to 2023, and the NLG estimates are the most imprecise of the three technologies. The significance of the NLG coefficient does not survive Bonferroni correction in the full sample specification. However, the coefficient is notably stable across all specifications, ranging between 0.030 and 0.036, which is more consistent with a genuine but imprecisely estimated association than with a null effect. General analysis of the NLP technologies' coefficients point towards a consistent pattern: the estimates of both technologies indicate productivity increases, and the clearest associations are present in the service industry. However, the two technologies differ in how sensitive their estimates are to sample composition. For TM, the coefficient decreases when early adopters are excluded, while NLG is more stable across specifications. This divergence likely reflects differences in the composition of early adopters across technologies rather than a genuine difference in productivity dynamics. Given this, the full-sample TM coefficient and its significance should be interpreted with caution. More broadly, for all three technologies, the restricted sample specifications in columns 3 and 4 provide cleaner estimates by removing already-treated firms from the control group, and should be considered the primary basis for interpretation.

For NLG, the coefficient is more stable across all specifications regardless of sample composition, which suggests that control group composition is less consequential for this technology. The differences in magnitude, significance, and heterogeneity patterns between TM and NLG are difficult to attribute to a single underlying firm characteristic driving both adoption and productivity, given that late adopters of the two technologies are similar types of firms.

PA shows the most certain and largest positive associations. Contrary to the language based technologies, estimates are clearer in goods producing firms. This aligns well with the descriptive analysis, where adopters of PA disproportionately report using AI in logistics, production, and finance which are functions characterized by repetitive, rule-based tasks where automation can directly speed up operations in a goods producing firm. The stronger associations among large firms may reflect an additional mechanism: larger firms can spread the fixed implementation costs of process automation across a greater number of transactions, which may be associated with stronger productivity patterns in high-volume operational settings. The stronger associations when early adopters are excluded are consistent with the mechanism discussed in Section 5.2: early PA adopters, who constitute a disproportionately large share of all PA adopters, tend to be more productive at baseline and their presence in the control group attenuates the full-sample coefficient. The restricted sample specifications, and especially the PSM specification, therefore likely provide cleaner estimates.

The general patterns of PA stand in contrast to the language-based technologies. As shown in the descriptive analysis, NLP adopters disproportionately apply AI in marketing & sales and R&D, which are functions where human judgment and creativity are central to value creation, and where AI can act as a complement to rather than a replacement for the worker. This assistive role fits naturally with service firms, where human expertise is itself the primary source of value. It follows that language-based technologies show their strongest productivity associations precisely in that context. The diverging patterns are more consistent with the technologies operating in genuinely different contexts than with a uniform selection story. This highlights the limitations of using aggregated measurements of AI, where heterogeneous effects are hidden and only the net results are shown. This heterogeneity in itself is informative. For firms considering adoption, the relevant question does not seem to be whether to adopt AI but which technology fits their operational context. The documented patterns suggest that process automation may be particularly relevant for larger goods producing firms, while language-based technologies show stronger associations in knowledge-intensive service sectors. For policymakers, the findings imply that digitalisation initiatives may be more effective when differentiated by technology type and sector rather than targeting AI adoption broadly.

7.2 Alignment with Previous Literature

This study's findings are broadly consistent with previous evidence of positive associations between AI adoption and productivity. No previous estimates are directly comparable to those reported here, given the focus on specific individual technologies and TFP as outcome variable. The preferred specifications here suggest TFP differences in the range of 2.3–4.5%, depending on the technology. These align well with previous measures of AI adoption increasing labour productivity, sales or value added with most commonly around 0-12%. The much larger productivity gains reported in within-firm RCTs and case studies are also consistent with the positive direction of results found in all three technologies explored in this study. However, the magnitudes of the effects are incomparable, as they measure work per time unit, and thereby exclude any implementation costs, organizational adjustment, and the parts of the organization that are not directly affected by the AI adoption. An important caveat when comparing this study to previous firm level literature is that most earlier studies were conducted when AI adoption in Sweden was dominated by PA and ML. To the extent that these adoption patterns were similar across countries, NLP technologies were considerably less common at the time analysed by most previous studies. This suggests that the aggregate AI measures in previous studies may primarily reflect PA and ML associations, and that evidence on NLP technologies specifically is more limited.

The descriptive analysis further shows a shift towards off-the-shelf and externally sourced AI solutions over the study period. This is relevant in relation to Lee et al. (2022), who find that internal R&D strategy strengthens the productivity associations with AI adoption. One interpretation is that if standardized AI solutions reduce the need for firm-specific R&D, the complementarity between AI and internal R&D documented by Lee et al. (2022) may simply reflect an earlier phase of the technology's maturity rather than a persistent pattern. As modern off-the-shelf tools require little internal adaptation, Lee's finding may be less applicable in the current period. An alternative interpretation is less optimistic: if firms are moving away from internally developed solutions and investing less in firm-specific R&D as a result, and if such R&D is complementary to realizing productivity gains from AI, the shift towards external acquisition may contribute to weaker productivity associations than what would otherwise be observed.

The findings offer mixed evidence on size heterogeneity. In the restricted sample specifications, associations tend to be somewhat clearer among larger firms across technologies, which is more consistent with Babina et al. (2024) than Damioli et al. (2021). However, as the descriptive analysis suggests, size differences in previous literature may partly reflect differences in technology composition rather than a direct effect of firm size. The descriptive evidence in Figure 2 shows that in 2021, large firms had already adopted ML and PA at substantially higher rates, while small firms showed a more mixed and limited adoption profile. This means that studies finding size-differentiated productivity effects of aggregated AI may partly capture differences in which technologies large and small firms were using, rather than a direct effect of firm size.

7.3 Limitations

Firstly, the scope of the study should be addressed. The reported results apply to firms in Sweden, which is a country with high digital maturity and technological development. Similar patterns may therefore not be observed in settings that differ in these aspects. Firms included in the explored population have 10 or more employees and belong to the 2007 statistical classification of economic activities (NACE Rev.2) codes 10-63, 68-75, 77-82, and 95.1 (covers manufacturing, construction, trade, transport, accommodation, information and communication, real estate, professional services, and selected other service activities, while excluding agriculture, mining, financial services, and public administration). Additionally, it covers the years 2021-2025 which is a period defined by quick development in the studied area, and may therefore not apply for future adoption waves when the technologies are more mature and widespread. Furthermore, the productivity analysis is estimated on a subset of firms that either adopt AI during the panel or serve as a matched control group of observably similar non-adopters. The estimated associations therefore capture productivity differences among firms that are plausible adopters, rather than being representative of the full survey population. Firms that are structurally very different from adopters, such as smaller firms in less knowledge-intensive sectors, are underrepresented in the estimation sample, and the documented patterns may not extend to these firms.

The estimated models are limited to only immediate productivity effects of AI, where the effects must be observed in the same year as adoption to show in the estimates. The reason for this limitation is the short panel of only three years, as well as the fact that most of the firms in the data adopt in the very last year, which leaves no room for a lagged AI indicator. This is a relevant limitation given that several previous studies find that productivity effects of AI only materialize with a delay of two to three years, as firms need time to reorganize workflows, train employees, and integrate new tools into existing processes. If such adjustment costs are present, long-run associations between AI adoption and productivity may be larger than what is captured in the present short-term analysis.

As another time related limitation, the inconsistent timing of the measures of TFP and AI adoption should be mentioned. AI adoption variables relate to the firms' AI use in the beginning of each year, while TFP relates to the full year. Because of this, any firm that adopts AI later in the year will be partially treated, but remains in the control group until the following year. This introduces bias, as partially treated firms may experience productivity effects due to AI while they are still in the control group. The bias makes the interpretation of the timing of effects uncertain. If the productivity gains of AI take time to materialize, firms classified as new adopters in January may have already passed the initial adjustment phase, having adopted mid-way through the previous year. The estimated associations should therefore be interpreted with some caution regarding how quickly returns to AI adoption arise, where the maximum timing discrepancy is one year, but for most firms it is likely a matter of a few months.

Lastly, the largest limitation of this study is the potential endogeneity of the estimates due to non-random treatment assignment. As documented in the descriptive analysis, firms that adopt AI are systematically larger and more productive already before adoption, which confirms that selection into treatment is not random. This study partially addresses the endogeneity concern in several ways: firm and industry-year fixed effects absorb time-invariant firm characteristics and sector-specific trends, propensity score matching restricts the comparison to observationally similar firms at baseline, and placebo tests provide no evidence of systematic pre-existing productivity trends. However, residual endogeneity that vary over firm and time, may simultaneously drive both AI adoption and productivity in ways that are not captured by the structure of the model. A related concern is that firms adopting one AI technology are more likely to simultaneously adopt others, meaning that individual technology coefficients may partly capture the productivity associations of co-adopted technologies. Moreover, if complementary investments such as cloud computing and data infrastructure moderate the relationship between AI and productivity, as documented by Lee et al. (2022), the estimated coefficients may also reflect variation in these unmeasured complements rather than the effect of the technology itself. The estimated AI coefficients should therefore not be interpreted as causal effects, but rather as conditional associations between adoption of the technology and TFP, net of firm-invariant characteristics, industry specific time trends, and observable baseline differences between adopters and non-adopters.

8 Conclusions

This study aims to describe AI adoption in Swedish firms, and examine its association with firm-level total factor productivity, with an emphasis on how associations between AI adoption and productivity vary across specific technologies. The descriptive analysis shows that AI adoption remains concentrated among larger, more productive firms in knowledge-intensive sectors, and that selection into adoption is strong. The main findings paint a nuanced picture of AI adoption and productivity, where estimated associations are modest but consistently positive over all model specifications and technologies. The question of which AI and for whom matters more than whether AI adoption increases productivity in the aggregate. The analysis reveals that language based AI is more used in knowledge-intensive business functions and show stronger productivity associations in service industries. However, in the case of process automation AI, it seems to be more highly adopted in routine-based areas, and shows the strongest associations in large goods producing firms.

These results depict AI as a heterogeneous set of technologies whose productivity implications depend on the context in which they are deployed. The consistency of documented associations across technologies, specifications, and outcome measures strengthens confidence that the observed productivity differences reflect genuine patterns rather than artifacts of any single modeling choice. For firms considering AI adoption, the documented estimates suggest that matching AI technology to firm characteristics such as size and sector is relevant for where productivity associations are strongest. For policymakers, the heterogeneity across sectors and technologies provides an empirical basis for more differentiated digitalisation initiatives than aggregate AI measures alone could motivate.

Several of the limitations of this study are due to or somewhat related to the short panel of data, which is also one of the key challenges of this field of study in general. As more firm-level panel data on specific AI technologies is accumulated over time, more robust documentation of associations will be possible, and stronger causal identification may become feasible with longer panels and more variation in adoption timing. Additionally, many AI technologies are so new in their current form that it will take a few more years until the total short- and long term effects are fully observed. As this data becomes available, future research should keep studying AI as different individual technologies, and explore the role of complementary factors, such as digital infrastructure, employee skills, and the simultaneous adoption of multiple AI technologies, to keep research relevant and informative to firms and policymakers. Furthermore, as AI adoption spreads to an increasing share of firms, the binary adoption dummy used in this and many previous studies will become a less informative measure. Therefore, it will be increasingly important to develop continuous measures that capture not only the breadth, but also the depth of how AI is integrated in a firm.

References

Acemoglu, D., Anderson, G. W., Beede, D. N., Buffington, C., Childress, E. E., Dinlersoz, E., Foster, L. S., Goldschlag, N., Haltiwanger, J., Kroff, Z., Restrepo, P., & Zolas, N. (2022). *Automation and the workforce: A firm-level view from the 2019 annual business survey* (NBER Working Paper No. 30659). National Bureau of Economic Research. https://doi.org/10.3386/w30659

Ackerberg, D. A., Caves, K., & Frazer, G. (2015). Identification properties of recent production function estimators. *Econometrica*, *83*(6), 2411–2451. https://doi.org/10.3982/ECTA13408

Afrin, S., Roksana, S., & Akram, R. (2025). Ai-enhanced robotic process automation: A review of intelligent automation innovations. *IEEE Access*, *13*, 173–197. https://doi.org/10.1109/ACCESS.2024.3513279

Agrawal, A., Gans, J., & Goldfarb, A. (2018). *Prediction machines: The simple economics of artificial intelligence*. Harvard Business Review Press.

Aldasoro, I., Gambacorta, L., Pal, R., Revoltella, D., Weiss, C., & Wolski, M. (2026, January). *Ai adoption, productivity and employment: Evidence from European firms* (BIS Working Papers No. 1325). Bank for International Settlements.

Alderucci, D., Branstetter, L., Hovy, E., Tambe, P., Baviskar, S., Goldschlag, N., Runge, A., & Zolas, N. (2020). *Quantifying the impact of AI on productivity and labor demand: Evidence from U.S. census microdata* [Unpublished manuscript. Carnegie Mellon University. Presented at AEA meetings.].

Babina, T., Fedyk, A., He, A., & Hodson, J. (2024). Artificial intelligence, firm growth, and product innovation. *Journal of Financial Economics*, *151*, 103745. https://doi.org/10.1016/j.jfineco.2023.103745

Bick, A., Blandin, A., Deming, D. J., Fuchs-Schündeln, N., & Jessen, J. (2026, March). *Mind the gap: AI adoption in europe and the US* (Discussion Paper No. DP21337). Centre for Economic Policy Research. https://fuchsschuendeln.com/wp-content/uploads/2026/04/2026-AI-Adoption_CEPR-DP.pdf

Bresnahan, T. F., & Trajtenberg, M. (1995). General purpose technologies 'engines of growth'? *Journal of Econometrics*, *65*(1), 83–108. https://doi.org/https://doi.org/10.1016/0304-4076(94)01598-T

Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., Agarwal, S., Herbert-Voss, A., Krueger, G., Henighan, T., Child, R., Ramesh, A., Ziegler, D. M., Wu, J., Winter, C., … Amodei, D. (2020). Language models are few-shot learners. *Proceedings of the 34th International Conference on Neural Information Processing Systems*.

Brynjolfsson, E., Li, D., & Raymond, L. (2025). Generative ai at work. *The Quarterly Journal of Economics*, *140*(2), 889–942. https://doi.org/10.1093/qje/qjae044

Brynjolfsson, E., Rock, D., & Syverson, C. (2017). *Artificial intelligence and the modern productivity paradox: A clash of expectations and statistics* (NBER Working Paper No. 24001). National Bureau of Economic Research. https://doi.org/10.3386/w24001

Business Sweden. (2025). *Swedish tech continues to defy the economic downturn* [Accessed 2026-02-05]. https://www.business-sweden.com/about-us/media/press-releases/press-releases/2025/swedish-tech-continues-to-defy-the-economic-downturn/

Callaway, B., & Sant'Anna, P. H. C. (2021). Difference-in-differences with multiple time periods. *Journal of Econometrics*, *225*(2), 200–230. https://doi.org/10.1016/j.jeconom.2020.12.001

Cambon, A., Hecht, B., Edelman, B., Ngwe, D., Jaffe, S., Heger, A., Vorvoreanu, M., Peng, S., Hofman, J., Farach, A., Bermejo-Cano, M., Knudsen, E., Bono, J., Sanghavi, H., Spatharioti, S., Rothschild, D., Goldstein, D. G., Kalliamvakou, E., Cihon, P., … Teevan, J. (2023). *Early llm-based tools for enterprise information workers likely provide meaningful boosts to productivity* (tech. rep.). Microsoft Research. https://www.microsoft.com/en-us/research/wp-content/uploads/2023/12/AI-and-Productivity-Report-First-Edition.pdf

Cobb, C. W., & Douglas, P. H. (1928). A theory of production. *The American Economic Review*, *18*(1), 139–165. http://www.jstor.org/stable/1811556

Czarnitzki, D., Fernández, G. P., & Rammer, C. (2023). Artificial intelligence and firm-level productivity. *Journal of Economic Behavior and Organization*, *211*, 188–205. https://doi.org/10.1016/j.jebo.2023.05.008

Damioli, G., Roy, V. V., & Vertesy, D. (2021). The impact of artificial intelligence on labor productivity. *Eurasian Business Review*, *11*(1), 1–25. https://doi.org/10.1007/s40821-020-00172-8

Engberg, E., Hellsten, M., Javed, F., Lodefalk, M., Sabolová, R., Schroeder, S., & Tang, A. (2025). Artificial intelligence, hiring and employment: Job postings evidence from sweden. *Applied Economics Letters*, 1–6. https://doi.org/10.1080/13504851.2025.2497431

European Commission. (2022). *Index för digital ekonomi och digitalt samhälle (desi) 2022: Sweden* (DESI country profile (SV). Accessed 2026-02-05). European Commission. https://digital-strategy.ec.europa.eu/en/library/digital-economy-and-society-index-desi-2022

Gardazi, N. M., Daud, A., Malik, M. K., Bukhari, A., Alsahfi, T., & Alshemaimri, B. (2025). Bert applications in natural language processing: A review. *Artificial Intelligence Review*, *58*, 166. https://doi.org/10.1007/s10462-025-11162-5

Hyun, Y., Lee, D., Chae, U., Ko, J., & Lee, J. (2021). Improvement of business productivity by applying robotic process automation. *Applied Sciences*, *11*(22), 10656. https://doi.org/10.3390/app112210656

International Trade Administration, U.S. Department of Commerce. (2024). *Sweden - digital economy* [Accessed 2026-02-05]. https://www.trade.gov/country-commercial-guides/sweden-digital-economy

Lee, Y. S., Kim, T., Choi, S., & Kim, W. (2022). When does ai pay off? ai-adoption intensity, complementary investments, and r&d strategy. *Technovation*, *118*, 102590. https://doi.org/10.1016/j.technovation.2022.102590

Levinsohn, J., & Petrin, A. (2012). Measuring aggregate productivity growth using plant-level data. *The RAND Journal of Economics*, *43*(4), 705–725. http://www.jstor.org/stable/41723351

Lievano-Martínez, F. A., Fernández-Ledesma, J. D., Burgos, D., Branch-Bedoya, J. W., & Jimenez-Builes, J. A. (2022). Intelligent process automation: An application in manufacturing industry. *Sustainability*, *14*(14), 8804. https://doi.org/10.3390/su14148804

Maslej, N., Fattorini, L., Perrault, R., Gil, Y., Parli, V., Kariuki, N., Capstick, E., Reuel, A., Brynjolfsson, E., Etchemendy, J., Ligett, K., Lyons, T., Manyika, J., Niebles, J. C., Shoham, Y., Wald, R., Walsh, T., Hamrah, A., Santarlasci, L., … Oak, S. (2025). *The ai index 2025 annual report* (AI Index Steering Committee). Institute for Human-Centered AI, Stanford University. Stanford, CA. https://doi.org/10.48550/arXiv.2504.07139

Microsoft. (2023). *Announcing microsoft copilot, your everyday ai companion* [Accessed 2026-02-05]. https://blogs.microsoft.com/blog/2023/09/21/announcing-microsoft-copilot-your-everyday-ai-companion/

Noy, S., & Zhang, W. (2023). Experimental evidence on the productivity effects of generative artificial intelligence. *Science*, *381*(6654), 187–192. https://doi.org/10.1126/science.adh2586

Olley, G. S., & Pakes, A. (1996). The dynamics of productivity in the telecommunications equipment industry. *Econometrica*, *64*(6), 1263–1297. https://doi.org/10.2307/2171831

OpenAI. (2025). *Chatgpt* [Accessed 2026-02-05]. https://openai.com/sv-SE/index/chatgpt/

Raj, M., & Seamans, R. (2019). Artificial intelligence, labor, productivity, and the need for firm-level data. In A. Agrawal, J. Gans, & A. Goldfarb (Eds.), *The economics of artificial intelligence: An agenda* (pp. 553–565). University of Chicago Press. https://www.nber.org/books-and-chapters/economics-artificial-intelligence-agenda/artificial-intelligence-labor-productivity-and-need-firm-level-data

Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, *70*(1), 41–55. https://doi.org/10.2307/2335942

Russell, S. J., & Norvig, P. (2010). *Artificial intelligence: A modern approach* (3rd ed.). Pearson Education.

Solow, R. M. (1987). We'd better watch out. *New York Times Book Review*, 36.

Statistics Sweden. (2021, November). *Kvalitetsdeklaration: It-användning i företag 2021* (tech. rep. No. NV0116_KD_2021) (Accessed: 2026-02-09). Statistics Sweden (Statistiska centralbyrån). https://www.scb.se/contentassets/9cfc4627e5954d109ab39dda9df1ba69/nv0116_kd_2021_gb_20211125.pdf

Statistics Sweden. (2023a, May). *Statistikens framställning: Företagens ekonomi (nv0109) – slutlig statistik 2021* (tech. rep. No. NV0109_ST_slut2021_v1_20230510) (Accessed: 2026-02-

09). Statistics Sweden (Statistiska centralbyrån). https://www.scb.se/contentassets/9dd20ce462644cc19f6f04eb2edbbe28/nv0109_st_slut2021_v1_20230510.pdf

Statistics Sweden. (2023b, November). *Kvalitetsdeklaration: It-användning i företag* (tech. rep. No. NV0116_KD_2023) (Accessed: 2026-02-09). Statistics Sweden (Statistiska centralbyrån). https://www.scb.se/contentassets/9cfc4627e5954d109ab39dda9df1ba69/nv0116_kd_2023.pdf

Statistics Sweden. (2024, November). *Kvalitetsdeklaration: It-användning i företag* (tech. rep. No. NV0116_KD_2024) (Accessed: 2026-02-09). Statistics Sweden (Statistiska centralbyrån). https://www.scb.se/contentassets/9cfc4627e5954d109ab39dda9df1ba69/nv0116_kd_2024.pdf

Statistics Sweden. (2025a). *Sni 2025 – standard för svensk näringsgrensindelning* [Accessed: 2026-02-09]. https://www.scb.se/dokumentation/klassifikationer-och-standarder/standard-for-svensk-naringsgrensindelning-sni/sni-2025/

Statistics Sweden. (2025b, May). *Statistikens framställning: Företagens ekonomi (nv0109) – slutlig statistik 2023* (tech. rep. No. NV0109_ST_slut2023_v1_20250509) (Accessed: 2026-02-09). Statistics Sweden (Statistiska centralbyrån). https://www.scb.se/contentassets/9dd20ce462644cc19f6f04eb2edbbe28/nv0109_st_slut2023_v1_20250509.pdf

Statistics Sweden. (2025c, November). *Kvalitetsdeklaration: It-användning i företag* (tech. rep. No. NV0116_KD_2025) (Accessed: 2026-02-09). Statistics Sweden (Statistiska centralbyrån). https://www.scb.se/contentassets/9cfc4627e5954d109ab39dda9df1ba69/nv0116_kd_2025.pdf

Statistics Sweden. (2025d, November). *Statistikens framställning: It-användning i företag, nv0116* (tech. rep. No. NV0116_STAF_2025) (Accessed: 2026-02-09). Statistics Sweden (Statistiska centralbyrån). https://www.scb.se/contentassets/9cfc4627e5954d109ab39dda9df1ba69/nv0116_staf_2025.pdf

Supriyono, Wibawa, A. P., Suyono, & Kurniawan, F. (2024). Advancements in natural language processing: Implications, challenges, and future directions. *Telematics and Informatics Reports*, *16*, 100173. https://doi.org/10.1016/j.teler.2024.100173

Wooldridge, J. M. (2009). On estimating firm-level production functions using proxy variables to control for unobservables. *Economics Letters*, *104*(3), 112–114. https://doi.org/https://doi.org/10.1016/j.econlet.2009.04.026

A Appendix

A.1 Figures & Tables

Table A1: Material Input as Proxy Variable: Residual Diagnostic

	ln(Output)
$\mathrm{Mat}_{\perp K,L}$	0.606***
	(0.022)
Intercept	11.952***
	(0.021)
Observations	7,239
R^2	0.091

Notes: $\mathrm{Mat}_{\perp K,L}$ denotes material inputs orthogonalized with respect to capital and labour by regressing log material on log capital and log labour. A significant coefficient confirms that material inputs carry information about unobserved productivity ω beyond standard inputs, supporting its use as a proxy variable in the Wooldridge (2009) estimator. Standard errors in parentheses. *** $p < 0.001$. *Source:* Statistics Sweden, Structural Business Statistics (2021-2024).

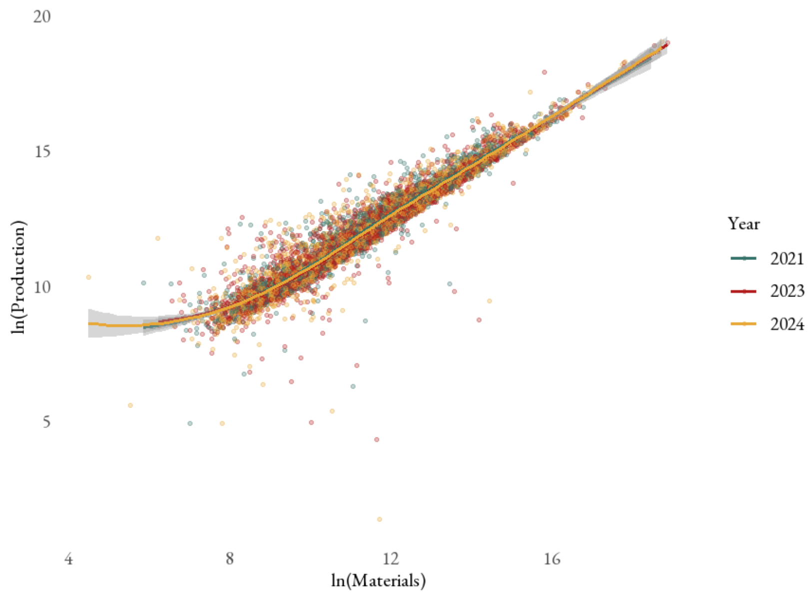

Figure A1: Relationship Between Log Material Inputs and Log Output by Year

Notes: Lines show locally weighted smoothing curves (LOESS) fitted separately by year. The monotonically increasing relationship supports the strict monotonicity assumption required for material inputs to serve as a valid proxy variable in the Wooldridge (2009) estimator. Shaded areas indicate 95% confidence intervals. *Source:* Author's rendering of Statistics Sweden, Structural Business Statistics (2021-2024).

Table A2: Production Function Estimates

	ln(Output)
ln(Labour)	0.622^{***}
	(0.024)
ln(Capital)	0.325^{***}
	(0.010)
Returns to scale	0.947
Observations	7,239
Bootstrap reps	200

Notes: Wooldridge (2009) production function estimator. Materials serve as proxy variable for unobserved productivity. Bootstrapped standard errors in parentheses (200 repetitions). Returns to scale is the sum of estimated input elasticities. *** $p < 0.001$. *Source:* Statistics Sweden, ICT Usage in Enterprises (2021-2024); Statistics Sweden, Structural Business Statistics (2021-2024).

A.2 Propensity Score Matching Diagnostics

PSM is used to reduce observable selection bias, since firms adopting AI tend to be larger and more productive even before adoption. Each treated firm is matched to the three control firms with the closest predicted probability of treatment. The matching is based on the variables ln_kap, ln_emp, ln_mat and industry. In Table 2, the means of all variables used in the regression model are presented for treated firms, the full control groups, the matched control groups. The matching is only done based on late adopters, as the matching must be done before the AI technology in question is adopted. Early adopters are therefore also excluded from the model estimation.

As already stated, the treatment and control groups differ substantially in their mean input values, already before AI adoption occurs. Table A3 shows that matching substantially reduces the imbalance on all covariates across all three technologies. For NLG and PA, the post-matching standardized mean difference (SMD) values are below or close to 0.10 for most covariates, indicating good balance. TM shows somewhat larger residual imbalance, particularly for ln(Employees) and ln(Material), where SMD remains at 0.17 and 0.15 respectively, which falls within the acceptable range but suggests that the matched control group is not perfectly comparable to the treated group. Importantly, baseline TFP is not included in the matching model, yet the SMD on this variable also decreases substantially after matching, which provides additional confidence that the matched samples are comparable in terms of pre-adoption productivity.

Looking at Tables A4, A5, and A6, it is clear that the matched control groups also become significantly closer to the treatment groups in terms of industry composition. Though similar to the numerical variables, the matched control groups do not perfectly replicate the treatment groups, and the industry shares often lie between those of the full control group and the treatment group, rather than fully aligning with the treated. This is especially clear for one industry that in the previous chapter was identified as underrepresented among non AI adopters: IT & Media. Because of these remaining differences, the treatment groups and matched control groups should not be viewed as fully comparable, but the matched control group provides a more similar comparison group than the full control group.

Even though the matched control groups are closer to the treatment groups than the full control groups, they are not perfect matches. This is an indication that even though the full control groups are much larger than the treatment groups, they do not overlap sufficiently for the matching to achieve perfect balance. Though by bringing the control group substantially closer to the treatment group, it can be assessed whether the documented associations persist when adopters are compared to observably similar firms, or whether they diminish substantially, which would suggest that observable selection drives the baseline associations

Table A3: Covariate Balance Before and After PSM Matching

Technology	Covariate	Treated	Before matching		After matching	
			Control	SMD	Control	SMD
Text Mining (TM)						
	ln(Capital)	9.92	8.90	0.33	9.62	0.10
	ln(Employees)	4.85	3.94	0.54	4.57	0.17
	ln(Material)	11.78	10.78	0.46	11.46	0.15
	TFP (baseline)	6.31	6.15	0.20	6.21	0.13
Natural Language Generation (NLG)						
	ln(Capital)	9.65	9.00	0.22	9.56	0.03
	ln(Employees)	4.85	4.00	0.49	4.69	0.11
	ln(Material)	11.69	10.86	0.38	11.53	0.07
	TFP (baseline)	6.31	6.16	0.18	6.27	0.05
Process Automation (PA)						
	ln(Capital)	10.11	8.75	0.47	9.98	0.04
	ln(Employees)	5.08	3.86	0.76	4.88	0.14
	ln(Material)	11.94	10.71	0.57	11.73	0.10
	TFP (baseline)	6.31	6.16	0.18	6.21	0.12

Notes: SMD denotes the standardized mean difference. $|\text{SMD}| < 0.10$ indicates good balance, $<$ 0.25 acceptable. TFP (baseline) is not part of the matching model. Matching uses nearest-neighbour propensity score matching ($k = 3$) with industry included as a covariate. *Source:* Statistics Sweden, ICT Usage in Enterprises (2021-2024); Statistics Sweden, Structural Business Statistics (2021-2024).

Table A4: Industry Distribution Before and After PSM Matching: Text Mining (TM)

Industry	**Treated**	**Control before**	**Control after**
Manufacturing (C)	0.226	0.328	0.277
Energy (D)	0.031	0.023	0.033
Water & waste (E)	0.008	0.028	0.008
Construction (F)	0.033	0.065	0.038
Retail & wholesale (G)	0.164	0.158	0.158
Transport (H)	0.046	0.052	0.051
Hotels & rest. (I)	0.042	0.075	0.048
IT & media (J)	0.156	0.056	0.095
Real estate (L)	0.054	0.043	0.056
Professional serv. (M)	0.166	0.096	0.155
Support services (N)	0.075	0.070	0.082
Other services (S)	0.000	0.005	0.000

Notes: Share of firms in each industry. Treated refers to late adopters, Control before to the unmatched control group, and Control after to the matched control group. Industry is included as a covariate in the propensity score model. *Source:* Statistics Sweden, ICT Usage in Enterprises (2021, 2023, 2024); Statistics Sweden, Structural Business Statistics (2021-2024).

Table A5: Industry Distribution Before and After PSM Matching: Natural Language Generation (NLG)

Industry	**Treated**	**Control before**	**Control after**
Manufacturing (C)	0.223	0.322	0.249
Energy (D)	0.031	0.024	0.032
Water & waste (E)	0.014	0.027	0.019
Construction (F)	0.014	0.066	0.013
Retail & wholesale (G)	0.178	0.156	0.188
Transport (H)	0.043	0.053	0.041
Hotels & rest. (I)	0.041	0.073	0.050
IT & media (J)	0.149	0.066	0.120
Real estate (L)	0.043	0.045	0.048
Professional serv. (M)	0.178	0.097	0.156
Support services (N)	0.084	0.068	0.082
Other services (S)	0.002	0.004	0.002

Notes: Share of firms in each industry. Treated refers to late adopters, Control before to the unmatched control group, and Control after to the matched control group. Industry is included as a covariate in the propensity score model. *Source:* Statistics Sweden, ICT Usage in Enterprises (2021, 2023, 2024); Statistics Sweden, Structural Business Statistics (2021-2024).

Table A6: Industry Distribution Before and After PSM Matching: Process Automation (PA)

Industry	**Treated**	**Control before**	**Control after**
Manufacturing (C)	0.226	0.320	0.245
Energy (D)	0.020	0.022	0.021
Water & waste (E)	0.016	0.027	0.014
Construction (F)	0.036	0.063	0.032
Retail & wholesale (G)	0.173	0.154	0.178
Transport (H)	0.043	0.051	0.045
Hotels & rest. (I)	0.020	0.079	0.022
IT & media (J)	0.177	0.064	0.131
Real estate (L)	0.056	0.042	0.064
Professional serv. (M)	0.150	0.105	0.157
Support services (N)	0.085	0.067	0.093
Other services (S)	0.000	0.004	0.000

Notes: Share of firms in each industry. Treated refers to late adopters, Control before to the unmatched control group, and Control after to the matched control group. Industry is included as a covariate in the propensity score model. *Source:* Statistics Sweden, ICT Usage in Enterprises (2021, 2023, 2024); Statistics Sweden, Structural Business Statistics (2021-2024).

A.3 AI Disclosure

This thesis is written with assistance from Claude Sonnet 4.6 (Anthropic) and Copilot GPT-5 (Microsoft). The tools were used in a few different ways. I used them for coding assistance, where I let AI debug or correct my own code, or suggest smaller bits of syntax which were always reviewed before use. I also used them as a sounding board to think through model choices, assumptions, and how results can and interpreted. Lastly in the writing process, they were used for text revision, where I used AI to translate words, search for synonyms, and look for typos and mistakes in already written text. No raw or confidential data was shared with any AI tool at any point.

The tools contributed to the quality of the thesis in two ways. As AI often gives quicker and more contextualized answers than a traditional search engine, it freed up more time for me to focus on more complex aspects of the thesis that required more thorough research or reasoning. Second, AI helped to keep a consistent and easy to read language quality in the full thesis.

Two risks stood out to me during this process. The first is hallucination. I never accepted factual claims from AI without going back and checking with the original sources directly. The second is that AI tools have a tendency to validate the user rather than being critical, which is not always what is needed when working through a research problem. I tried to avoid this by framing prompts to invite criticism of whatever I was asking about. Small differences in how I phrased a prompt had a larger effect on the response than I initially expected. These tools work best when used with good knowledge of their limitations. They are built for accessibility rather than as research assistants, and as a user you need to adjust how you use them depending on what you actually need.