Lisbon School of Economics & Management
Universidade de Lisboa

MASTER

APPLIED ECONOMETRICS AND FORECASTING

MASTER'S FINAL WORK

DISSERTATION

# THE ROLE OF CEO SALARY IN FIRM PRODUCTIVITY: SECTORAL EVIDENCE FROM PORTUGAL VIA PENALIZED REGRESSION

JOANA ISABEL LOPES GROMICHO

SUPERVISION:

JOÃO ANTÓNIO MENDES DA CRUZ

DOCUMENT SPECIALLY WRITTEN OBTAINING THE MASTERS DEGREE

JULY - 2025

*To Mom and Dad,*
*to whom I owe*
*everything.*

# GLOSSARY

**CEA** Classification of Economic Activities. 9

**CEO** Chief Executive Officer. ii, iii, 1–11, 14, 17, 19, 21, 23–25, 38

**EC** European Commission. 9

**FE** Fixed Effects. 10–12, 15

**FWL** Frisch-Waugh-Lovell. 11–13, 15

**GMD** Groupwise-Majorization-Descent. 15

**ISCO** International Standard Classification of Occupations. 8

**JEL** Journal of Economic Literature. ii, iii

**LASSO** Least Absolute Shrinkage and Selection Operator. ii, iii, v, 2, 3, 12, 14–18, 20–24

**MBA** Master of Business Administration. 6

**NACE** Statistical Classification of Economic Activities. 9

**OLS** Ordinary Least Squares. 12, 23, 37

**QP** Quadros de Pessoal. iii, 7, 8

**R&D** Research & Development. 18, 19, 21, 23, 24, 34

**RPE** Relative Performance Evaluation. 3, 5

**USA** United States of America. 5, 6

ABSTRACT

This study aims to identify the economic sectors in Portugal in which the Chief Executive Officer (CEO) salary affects firm productivity. The analysis focuses primarily on estimation using a matched employer-employee dataset within a high-dimensional setting, encompassing various Portuguese economic sectors. To address the issue of overfitting common in these scenarios, the framework draws on machine learning and variable selection research to apply data-driven methods for choosing economic sectors. Specifically, we utilize Least Absolute Shrinkage and Selection Operator (LASSO) techniques to select relevant economic sectors in which CEO salary has an impact on firm productivity. The persistently selected economic sectors are real estate promotion, specialized construction, telecommunications, accommodation and food services, which reveals the prevalent business paradigm in Portugal. Among the selected 45 economic sectors, we highlight two sectors that exhibit a positive relationship between CEO salary and firm productivity: telecommunications sector and residential care within the non-profit sector.

KEYWORDS: productivity; CEO salary; economic sectors; machine learning; LASSO.

JOURNAL OF ECONOMIC LITERATURE (JEL) CODES: C52; C53; C55; J24; M12.

## RESUMO

Este estudo pretende identificar os setores económicos em Portugal, nos quais o salário dos CEOs afecta a productividade da empresa. A análise foca-se principalmente na estimação utilizando os Quadros de Pessoal (QP) num contexto de multi-dimensionalidade, assim como considerando vários setores económicos portugueses. De modo a corrigir o problema de sobreajuste comum nestes cenários, a metodologia baseia-se em *machine learning* e seleção de variáveis para aplicar metódos orientados por dados para escolher os setores económicos. Especificamente, utilizamos os métodos *LASSO* para selecionar os setores económicos relevantes nos quais o salário do CEO afeta a produtividade da empresa. Os setores económicos persistentemente selecionados são a promoção imobiliária, a construção especializada, as telecomunicações, o alojamento e a restauração, revelando o paradigma empresarial prevalecente em Portugal. Entre os 45 setores económicos selecionados, destacamos dois sectores que exibem uma relação positiva entre o salário do CEO e a produtividade da empresa: o sector das telecomunicações e o sector de apoio social com alojamento.



PALAVRAS-CHAVE: produtividade; salário de CEO; setores económicos; machine learning; LASSO.

CÓDIGOS JEL: C52; C53; C55; J24; M12.

TABLE OF CONTENTS

## LIST OF FIGURES

## List of Tables

## ACKNOWLEDGEMENTS

I want to sincerely thank Professor João Cruz for his encouragement and expert advice during the course of this thesis. His valuable input and constant support have been essential in shaping the direction and quality of this research.

I am also deeply indebted to Professor Isabel Proença and Dr. Jorge Oliveira, Director and Deputy Director of *Unidade Técnica de Avaliação de Políticas Tributárias e Aduaneiras (U-TAX)*, respectively, as well as to my work colleagues. Their exceptional understanding and support proved indispensable as I endeavored to balance professional responsibilities with academic pursuits. The team constantly demonstrated remarkable empathy, flexibility, and a generous willingness to share their extensive knowledge, thereby significantly contributing to the refinement of my own expertise. I must especially thank Dr. Hugo Morão for his incredible help and constructive comments that significantly improved my research, and with whom I am constantly learning.

My sincere appreciation extends to my fellow Master's colleagues, especially those who worked closely with me over these past two years. Our discussions, problem-solving endeavors, and mutual support have been invaluable.

I am profoundly grateful to my professors, whose instruction and mentorship substantially enriched my knowledge and equipped me with the methodologies employed in this research. I also wish to acknowledge their encouragement to cultivate critical thinking and to strive for excellence, which has not only enhanced my skills as an econometrician but also strengthened my role within society.

I offer heartfelt thanks to my dearest friends who have been a constant source of encouragement and support throughout this journey. Their presence and reassurance during challenging moments have been truly invaluable.

Finally, my deepest gratitude is reserved for my parents, whose infinite patience, emotional support, and unwavering encouragement have sustained me throughout the process of finishing this thesis. Their faith in my abilities, even during the most demanding times, has been a source of strength and motivation. This work is as much a reflection of their steadfast love and support as it is of my own commitment and perseverance.

# 1 INTRODUCTION

It is widely acknowledged that Chief Executive Officer (CEO) salaries contribute to firm productivity; however, how they influence productivity across economic sectors has received less attention in the literature. The purpose of this research is to identify the economic sectors in which the CEO salary affects the firm productivity. The results address a significant gap in the literature by providing a micro-level perspective on the relationship between these. This study uncovers that CEO salary has an impact on firm productivity in the food and garment industries, real estate promotion, construction and related services, telecommunications, accommodation and restaurant services, as well as consulting, financial and legal services. The sectors of education, health, non-profit support activities and associative organizations also exhibit a relationship between CEO salary and firm productivity.

Firm productivity is a major instrument to strength the firm stability and long-term financial performance, incrementing firm value. In parallel, CEO salary serves as a key mechanism to motivate executives to make strategic choices that align with firm objectives and influence firm's exposure to external factors that may affect outcomes. Empirical theory supports the evidence that CEO salary does impact the firm productivity, with firm size being a prominent factor in this relationship (Gabaix & Landier 2008). Still, prior studies have predominantly examined this relationship at the aggregate economy level, neglecting the sector-specific dynamics.

Moreover, Gopalan et al. (2010) state that the optimal wage contract is influenced by sector movements, highlighting a positive relationship between CEO salary and firm performance within operating sectors. This positive relationship is often attributed to the motivational effect of high salaries, which encourages CEOs to outperform by effectively regulating market exposure through firm strategy. Such dynamics may explain the rise in CEO salary over the last decades, driven by the scarcity of qualified CEOs and the increasing demand for general managerial skills rather than firm-specific expertise (Gopalan et al. 2010). Larger firms, with greater financial capacity, are better positioned to attract the scarce CEO talent, which leads to the increment of firm productivity and firm growth. This conclusion uncovers the importance of firm size to the relationship between CEO salary and firm productivity.

However, the identification of specific economic sectors in which this relationship is significant has not been sufficiently studied in the literature. This research provides empirical evidence for the Portuguese market, which is further under-researched. Importantly, the methodological framework used here can be extended to other markets for

future contributions to the literature.

The persistence of substantial productivity differences between sectors, which are not attributed to production factors, is extensively documented in the literature (e.g., Fukui et al. 2023). However, the determination of the impact of the CEO salaries on firm productivity in certain economic sectors faces difficulties due to the large number of sectors, creating overfitting, and reducing estimation efficiency. A traditional linear regression model generally struggles to deal with such amount of variables. This paper leverages recent advances in machine learning and variable selection literature to apply data-driven methods identifying the most important economic sectors and measure the respective impact of the CEOs' salary on firm productivity. The method applied offers substantial improvements in the determination of the most relevant economic sectors by reducing the enormous number of potential explanatory variables and overcoming the problem of overfitting.

We start by considering the group-Least Absolute Shrinkage and Selection Operator (LASSO) approach by Yuan & Lin (2006). We further complement this analysis with the LASSO method for variable selection (see Tibshirani 1996). Specifically, we apply LASSO to the linear model and demonstrate how to select its turning parameter using cross-validation. Firstly, the predictors were selected by groups, and then within those, we selected the more persistent individual sectors. We apply our LASSO estimators to an unique and rich matched employer-employee dataset made available by the Portuguese Ministry of Employment and Solidarity. Notably, the dataset offers detailed information at both the firm and employee levels.

This thesis makes a primary contribution to the expanding literature on CEO salary, and further adds to research on its effects on firm productivity. By employing these methods, we are able to identify the economic sectors. This contribution allows to use the CEO salaries as instruments to manage firm productivity in the selected economic sectors. Technically, this thesis conjugates the group LASSO with grouped variables and, within the chosen groups, the baseline method with disaggregated sectors. To our knowledge, this methodology has not been used in the literature and allows an improved adjustment of the results. The estimation strategy consisting on the consecutive application of the two machine learning methods - LASSO and group LASSO - to refine the results constitutes an additional contribution. Finally, the method is versatile, which means that it can be used in multiple datasets, both national and international, and improve different topics of research.

The remainder of this thesis is organized as follows. Section 2 reviews the relevant literature on CEO salaries, explores the causes driving its significant increase, as well as

its relation with firm productivity. Section 3 presents the employer-employee matched dataset at firm and employee level and the descriptive analysis. Section 4 introduces the model and explains the implementation of linear-LASSO estimation. Section 5 applies the methods described above to our employer-employee database and presents the variable selection and estimation results. Section 6 concludes the thesis, and technical details are provided in the Appendix.

## 2 LITERATURE REVIEW

The literature widely documents persistent productivity disparities across countries (Dai et al. 2025) and within specific industries (Fukui et al. 2023, Kaplan & Pérez-Cavazos 2021, Demiraj et al. 2024), which remain unexplained by variations in traditional production factors. A potential explanation for these productivity differentials may be the disparities in managerial ability. Managers play a critical role in overseeing production, public services, and project execution across for-profit, non-profit and public firms. Consequently, a thorough understanding of top management - particularly CEOs - is essential for examining the fundamental sources of productivity disparities within the economy. Additionally, firms allocate significant time and resources in recruiting and developing their managerial personnel. Notably, a significant portion of the observed increase in wage inequality over the past two to three decades is frequently attributed to the incrementing share of CEO salary to that of the broader workforce. Hence, it is crucial to verify if CEO salary can be used as an instrument to increment firm productivity.

A key insight from compensation theory is that incentive remuneration should be calculated accordingly on factors that the CEO can directly influence, rather than those outside their control. Therefore, when a CEO cannot affect the performance of the sector of the firm, the ideal incentive scheme should reward the firm's results relative to the performance of the sector, known as Relative Performance Evaluation (RPE). Still, when extensively tested, empirical studies generally find limited support for RPE among typical CEO (e.g. Garvey & Milbourn 2003). Instead, CEO salary and the performance of the firm's operating sector exhibit evidence of a strong positive relationship between them. The CEO has some influence over the exposure of the firm to the market forces, beyond the CEO's direct control, through the choice of the strategy of the firm. Gopalan et al. (2010) proves that the optimal wage contract is influenced by the sector movements and those will vary cross-sectionally. Although, the increase in CEO salary may not primarily reflect inefficient rent extraction, but rather the absence of RPE and the enduring positive correlation between CEO salary and sector performance. The relationship between CEO salary and firm productivity is equivalent, which means that an increment in CEO

salary will improve the firm strategy and its results and, given the boost in productivity, translated in firm outcome, the CEO salary may increase, especially if it has performance-based incentives in its structure.

Empirical evidence of the relationship between CEO salary and firm productivity across economic sectors remains very scarce. Generally, the literature studies the impact of CEO salary on firm productivity as a generalized phenomenon within the broader economy, dismissing the sector-specific details of this relationship. Although, if we deepen our analysis, researchers studied several sectors within particular countries, proving that CEO salary affects firm productivity, with firm size being a major factor (e.g. Bhuyan et al. 2020). Sectoral analysis corroborates the conclusions obtained from aggregate economic studies. However, the effects of CEO salary on firm productivity may be more predominant in certain sectors. For instance, Gopalan et al. (2010) proves that the impact of CEO salary on productivity is stronger in multi-segment firms, as well as in firms where the CEO has more flexibility to modify firm's strategy through its exposure to external factors.

Hence, it is necessary to uncover the sectoral effect of CEO salary on productivity. The structure of CEO salary varies across sectors due to market dynamics, regulatory environments and competitive pressures, which influence how salary translate into firm performance. This observed relationship offers suitable incentives for the CEO to shape the firm's strategy by effectively managing its production and project execution, influencing firm productivity. Moreover, the existence of a relationship between CEO salary and firm performance serves as an encouragement to their CEOs to adopt strategies that improve sector-specific opportunities and minimize risks that arise from the economic paradigm. Aside from short-term improvements, this association ensures a sustainable competitive position within the firm's sector.

Below we will discuss the reasons executive compensation increased significantly in the last decades. Scholars have proposed economic explanations to help understand this rise in salary. The literature argues that the expected firm performance depends directly on the CEO's qualifications and the market forces beyond their control. The optimal contract is designed to motivate the CEO to exert effort at maximizing firm value. One explanation for the rise in CEO salary points to the broad adoption of performance-based incentive packages since the late 1980s. Such contracts reward the CEO based on firm performance affected by sector trends. Instead, it should encourage the CEO to anticipate sector trends and strategically decide the firm's optimal exposure to them. Both academics and shareholder activists advocated for stronger, market-oriented managerial incentives throughout the 1990s, with the expectation that this would align the interests of CEOs and sharehold-

ers (e.g. Jensen & Murphy 1990). Bulan et al. (2010) contend that option-based contracts effectively align CEO incentives with shareholders interests, provided there is a strong positive relationship between productivity and firm value. Nonetheless, shareholders are not primarily concerned about the CEO interests, but rather whether the CEO possesses the foresight to select the optimal strategy for utilizing the firm's assets. Shareholders are profit-driven, while CEOs may be also concerned with the social and environmental effects of the firm. Given this, the CEO is therefore focused on guiding the firm's strategic direction given the market conditions in which it finds itself. According to Dow & Raposo (2005), an increment in the business environment volatility has made it optimal for firms to offer higher performance incentives. Consistently, Cuñat & Guadalupe (2009) demonstrate a causal link between heightened market competition and the increased sensitivity of United States of America (USA) CEO salary to firm performance.

An increment in productivity induces an increase in firm value, which encourages CEOs, even risk-averse ones, to pursue projects that allow the former. However, increasing performance incentives in the presence of risk-aversion or limited liability requires higher overall compensation to motivate CEOs' participation. Although, Gopalan et al. (2010) demonstrate that the optimal contract for a risk-averse CEO provides greater rewards during periods of strong sector performance than punishments during downturns, indicating an asymmetric sensitivity to sector conditions. Holmstrom & Kaplan (2001) associate the growth in overall compensation levels with the increasing use of stock-based compensation, which may be included in CEOs' salary. Nevertheless, the relationship between the level and the sensitivity of compensation has not been fully explored. It is unclear why, instead of the substantial increment of the CEOs salaries, it was not implemented the exchange of salaries for securities, maintaining the overall salary constant. Similarly, the substantial existing wealth held by CEOs likely reduces their risk aversion, meaning that only modest increases in overall compensation are needed for them to accept a greater emphasis on performance-based pay.

Taking into consideration the wave of corporate scandals and the notoriety of the shortcomings in the USA corporate governance framework, the theory that CEOs extract more money from firms through the increment of compensation gained more attention (Bertrand & Mullainathan 2001, Bebchuk & Fried 2003, 2004). Advocates of the skimming perspective demonstrate the causal relationship between the rise of CEO salary with exclusively managerial entrenchment. This could consist in adjusting the costs of stock options or camouflaging excessive pay as performance-based rewards, and thus without encountering shareholder resistance (Bebchuk et al. 2002). Bertrand & Mullainathan (2001) even explain that the CEOs' salary contracts are based in "pay for luck" rather than "pay for performance", exhibiting the lack of RPE for the average CEO.

A more moderate version of the skimming perspective, as proposed by Hall & Murphy (2003), argues that the surge in stock-option compensation can be attributed to boards' limited ability to accurately assess its true cost. Most certainly these factors influence compensation, but their relevance is unclear for the majority of firms. For example, Rajan & Wulf (2006) dispute the notion that perks represent mere managerial excess, demonstrating that firms tend to provide substantial perks when they are expected to boost productivity. Furthermore, Gabaix & Landier (2008) proposes a purely competitive benchmark accounts for the increase in USA CEO salary without attributing it to changes in rent extraction, viewing this rise instead as an equilibrium outcome of significant growth in firm size.

The third explanation links the rise in CEO salary to the transformation in the nature of the CEO position. Garicano & Rossi-Hansberg (2006) propose a model in which advancements in communication technology have transformed the nature of the CEO's responsibilities, which in turn affects their compensation. Hermalin (2005) contends that the increase in CEO salary is indicative of a tighter corporate governance. In order to remunerate the increase in the probability of termination, CEO salary must be higher. Frydman (2019) and Murphy & Zábojník (2004) present evidence indicating that CEO roles have progressively shifted focus towards general skills rather than firm-specific expertise, consequently incrementing their job opportunities and contributing to salary increases. This phenomenon may be due to the explicit relationship between CEO ability and managing the firm's place in the broad market. Despite this, the explanation faces challenges when evaluated quantitatively. Moreover, given the increase in the number of Master of Business Administration (MBA)s among executives and the widespread availability of executive education, it is unlikely that a shortage of general skills significantly accounts for the rise in CEO salary. However, elevated compensation levels may stem from firms competing for limited managerial talent (Artige & Bignandi 2023, Green et al. 2021, Tervio 2008), especially since Gopalan et al. (2010) proved that more talented CEOs have a greater pay for sector performance. This results in increased salaries within bigger firms, which implies that salary must rise accordingly to the average firm size within the market (Gabaix & Landier 2008). The expansion of firm size is due to the result of competitive market forces.

This analysis of the underlying factors driving the increase in salary highlights the corresponding relationship between CEO salary and firm productivity. Nonetheless, the primary focus of this study is on the effects of CEO salary on firm productivity, allowing salary to be employed as a strategic tool to enhance performance. An increase in CEO salary can serve to attract and retain highly qualified executives who possess the strategic vision necessary for the specific economic sector in which the firm operates. Moreover,

higher salary may incentivize CEOs to undertake riskier decisions that, if successful, can lead to productivity gains. However, due to the structure of CEO salary - especially when it includes performance-based incentives - firm productivity may also influence salary, reinforcing a bidirectional relationship between the two.

## 3 DATA & SAMPLE SELECTION

This section outlines the variables and sample used in the analysis, and further presents relevant summary statistics. Our analysis is based on a rich matched employer-employee dataset from Portugal that allows us to identify in which economic sectors the CEO salary impacts firm productivity, while controlling the characteristics of the firm.

Quadros de Pessoal (QP) is a yearly survey of the private sector firms carried out by the Portuguese Ministry of Employment and Solidarity, making the data highly reliable. It is compulsory by Portuguese law for firms to make this information available, providing information relative to firms with at least one paid employee, referring to the survey week conducted in October. The nature of QP is longitudinal, with anonymous identifiers for each firm, presenting information at firm and employee level. Being the firm our sample unit, we have data about the industry, location, founding date, gross revenue generated in the prior calendar year and number of establishments and workers. We considered only Portuguese firms, which have more than one employee. Our working sample will have 291,509 firm-year observations. The sample excludes government employees and self-employed individuals. In order to construct the CEO variable, we need the employer's monthly salary (normal and extra), salary bonuses and allowances (regular and irregular), hours of work (normal and overtime) and professional category. Starting from the universe of observations in QP, we considered CEOs in full time, salaried workers, with full salary remuneration. We use QP data for the years 2022 and 2023, which are the latest years available (at the time of writing), demonstrating a stable and more regular activity of the firms after the Covid-19 pandemic. Our outcome of interest is firm productivity, using the gross value-added per worker of the firm as a proxy. This measure of productivity represents the value of goods produced or services provided in a firm divided by its number of workers, delivering a straightforward measure of labor productivity and allowing the identification of the sectors in which firms are more productive (Mensah et al. 2022, Lanau & Topalova 2016). The gross value-added per worker data are frequently available at firm, sector, regional, and national level, conceding comparisons at different levels.

The objective of this paper is to determine in which economic sectors the CEO salary impacts firm productivity. However, does not exist a unique definition of CEO in the scientific context. Hence, we need to identify each firm's CEO using a classification

process that applies different criteria at each stage.[1]

Firstly, given the variable *job title* in QP data set, we considered the practical definition postulated by the current International Standard Classification of Occupations (ISCO).[2] The selected ISCO-08 code was 1120, which characterizes CEOs as the ones who

> formulate and review the policies and plan, direct, coordinate and evaluate the overall activities of enterprises or organizations (except special-interest organizations and government departments) with the support of other managers, usually within guidelines established by a board of directors or a governing body to whom they are answerable for the operations undertaken and results.
>
> In: International Labour Organization.

In cases of a tie, we select the CEO with the highest salary among the candidates.[3] Once all possibilities based on *job title* have been exhausted, we assign the CEO label to the employee with the highest hierarchical level according to the *qualification* variable to identify the remaining firm pairs.[4] If a tie still occurs, we again choose the highest-paid individual among those with the top qualification level. Lastly, for any remaining firm without an identified CEO, we choose the highest-paid employee as CEO. This final step results in no unresolved ties and all firms have an identified CEO.

We construct the CEOs salary by applying the logarithm to each salary, including normal and extra income, and regular earnings supplements. Normal income refers to the fixed portion of a CEO's salary, which is reviewed yearly. Usually, it is determined relative to sector standards and shows a strong correlation with firm size, specifically in firms with solid equity performance. Extra income represents a portion of the regular salary and is frequently linked to the firm's short-term performance. Given this, CEO's cash compensation may exhibit a correlation with firm productivity. Additionally, the salary levels are closely associated with firm performance in a competitive labour market for CEOs (Leone et al. 2006). Conversely, cash compensation is believed to be linked to risk. Coles et al. (2006) and Berger et al. (1997) employ CEO cash compensation as a proxy for CEO risk aversion, based on the premise that CEOs receiving elevated cash remuneration are probably more entrenched and hence be more risk-averse. Another argument

[1] This classification process, along with sample selection method, is based on the approach used in Queiró (2021).

[2] It is available in `http://www.ilo.org/public/english/bureau/stat/isco/docs/groupdefn08.pdf`.

[3] The variable *job title* corresponds to "prof_4d" in the QP data set.

[4] The variable *qualification* corresponds to "nqual1" in the QP data set.

is that CEOs with substantial cash compensation possess greater wealth to invest externally, which may lead them to exhibit lower risk aversion (Guay 1999). We excluded the salaries below 705 euros in 2022 (0.02%), which corresponds to the minimum salary in the respective year. In 2023, all the salaries in the sample were above the minimum salary in that year. To create the sectoral relationship between CEO salary and firm productivity, we must associate the salary to the economic sector in which the firm operates.

The economic sectors are aggregated according to the Classification of Economic Activities (CEA).[5] As the name indicates, this classification aims to classify and aggregate the firms, which produce a good or provide a service, by economic activities. This allows the organization of the social-economic statistical information and the comparison of the data at national and international levels due to the harmonization with the Statistical Classification of Economic Activities (NACE) by European Commission (EC). We excluded residual economic sectors, and agricultural and fishing sectors, residual professional categories and residual regions. The CEO salary is associated with the dummy representing the economic sector in which the firm operates, allowing the correspondence between the CEO salary and firm productivity in each sector.

To assess the impact of CEO salaries on productivity, it is essential to control for other significant variables. Consistent with prior studies, our analysis incorporates the following additional factors: share of public social capital, firm size, number of establishments, antiquity of the firm (Haltiwanger et al. 1999, Huergo & Jaumandreu 2004), share of unionized and salaried workers, as well as the economic sectors. Further details on the measurement and construction of these variables are provided in Appendix A.1.

Table I presents the summary statistics for our sample. The average value of the productivity is 121,338.1 euros, which means that one worker produces, in average, 121,338.1 euros in gross value-added. The average public capital of the firms is 0.1%, hence the firms in the sample are mostly private. We also conclude that most of the firms in Portugal are small and medium sized, having, in median, one establishment and are 16 years old. The workers are mainly salaried and do not belong to an union. The average CEO salary is 1,471.03 euros, which reinforces the medium and small size of the Portuguese firms. Moreover, the minimum CEO salary is 705 euros, which corresponds to the minimum salary in 2022, and the maximum CEO salary is 690,383.2 euros.

[5]The description of the Portuguese economic sectors is available in `https://www.ine.pt/ine_novidades/semin/cae/CAE_REV_3.pdf`.

TABLE I: SUMMARY STATISTICS (2022-2023)

| | Mean | Median | S.D. | Min. | Max. |
|---|---|---|---|---|---|
| Productivity | 121338.1 | 58654.5 | 741984 | 2.5 | 162452816 |
| Public Capital | 0.001 | 0 | 0.04 | 0 | 1 |
| Dimension | 16.89 | 5 | 142.92 | 2 | 26370 |
| Antiquity | 18.85 | 16 | 16.90 | 0 | 525 |
| Unionized Workers | 1.95 | 0 | 11.26 | 0 | 200 |
| Salaried Workers | 87.97 | 100 | 20.84 | 4.44 | 783.33 |
| Number of Establishments | 1.29 | 1 | 4.30 | 1 | 698 |
| CEO Salary | 1471.03 | 1052.5 | 2091.20 | 705 | 690383.2 |

Descriptive statistics of the dependent variable and the control variables. The productivity is the gross value-added per worker, and is represented in euros. The public capital is the share of public social capital. Dimension is the number of employees. Antiquity is the number of years since the firm started. Unionized and salaried workers is the ratio of the workers with the respective characteristic and the total number of workers. The CEO salary is the normal and extra income, including earnings supplements, in euros. The economic sector dummies are not represented due to their large number (81 dummies). The data presents the firm-CEO pairs for the firms represented in QP for the period 2022-2023. The size of the sample is $n = 291,509$. Please refer to Appendix A.1 for variable definitions.

## 4 METHODOLOGY

We now explain the methodology used to estimate and to identify in which economic sectors the CEO salary impacts firm productivity. To preview our approach, we will use a linear regression model where firm productivity is regressed on the covariates of interest (here, the interaction between the economic sector in which the firm is inserted and its CEO salary), the control variables, plus the firm and time Fixed Effects (FE). Building upon this standard framework, we subsequently address the estimation challenges inherent from this specification. Addressing estimation challenges requires the extent of recent computational techniques in high-dimensional estimation based on penalized procedures for variable selection.

### *4.1 Linear Regression Model*

Consider the following regression model to estimate, in which economic sectors, the CEO salary has a relevant influence on firm productivity

$$y_{it} = \sum_{j=1}^{p} \beta_j x_{j,it} + \boldsymbol{\gamma}^{\boldsymbol{T}} \mathbf{Controls_{it}} + \eta_i + \delta_t + u_{it} \tag{1}$$

where:

- $y_{it}$ is the log-productivity of firm $i$ at time $t$,

- $x_{j,it}$ is the $j$-th covariate of interest representing the interaction between the CEO salary and a dummy variable indicating whether firm $i$ belongs to the $j$-th economic sector,
- $\mathbf{Controls_{it}}$ is a vector of firm's non-penalized control variables,
- $\eta_i$ is firm FE,
- $\delta_t$ is time FE.
- $u_{it}$ is the error term.

To examine the relationship between productivity and CEO salary, we identify two potential estimation challenges related to the construction of the variables: the endogeneity of CEO salary and the lasting effects of productivity disturbances. The well-documented endogeneity issue of CEO salary in this context typically arises from unobserved firm heterogeneity or variations in the firm's contractual environment. When a firm is more productive, its gross value increases and, as mentioned before, if the CEO has performance-based incentives, his/hers salaries will increase. The same occurs inversely: if the firm under produces, the CEO salary will decrease. To correct for the endogeneity of CEO salaries, we add to the regression firm FE, which will control the unobservable firm heterogeneity.

A second source of endogeneity may stem from the influence of past firm performance on current CEO salary. More specifically, CEO salary partly reflects the firm's recent performance to some extent. Given the period of our sample, we assume that the past shocks of productivity present a null expected value conditioned on the other variables considered and the FE. The time frame considered (2022-2023) is a period of stabilization and consolidation due to the COVID-19 pandemic, hence we do not assume significant productivity shocks. Having thoroughly addressed the endogeneity concerns, we are assured that potential estimation biases have been substantially mitigated.

Although the model presented has non-penalized control variables and two sets of FE, estimation is feasible due to the FE within transformation, explained by Cameron & Trivedi (2005). Additionally, given the implementation of the chosen method that we will explain later, we will need to apply the Frisch-Waugh-Lovell (FWL) theorem.

Our analysis centers on the set of covariates, $x_{j,it}$. Hence, besides the challenges mentioned before, a key econometric challenge specific to our context is the high-dimensional framework arising from the number of sectors being potentially very large. For this reason, in order to guarantee a feasible estimation and avoid overfitting, we will rely on methods from penalized regression literature, as presented below.

### *4.2 Model Transformations*

Taking Equation 1 into account, $\eta_i$ denotes firm-specific effects, capturing unobserved heterogeneity that may be correlated with the regressors $x_{it}$.

As noted by Cameron & Trivedi (2005), a FE model accounts for unobserved firm heterogeneity that may exhibit correlation with the regressors, leading to an omitted variable bias. For this model, we must assume strict exogeneity,

$$E[u_{it}|\eta_i, x_{i,2022}, x_{i,2023}, \mathbf{Controls_{i,2022}}, \mathbf{Controls_{i,2023}}] = 0 \tag{2}$$

which means that the error term has an expected value of zero conditional on past, current and future values of the regressors.

The presence of $N$ parameters $\eta_1, ...\eta_N$ are a challenge for estimation due to their increment in number as $N \to \infty$, meaning they are nuisance parameters that are not of intrinsic interest. However, they are required for a robust estimation of the parameters $\beta$, which are of interest. In order to present consistent estimators, we need to perform a within transformation, which is obtained by subtracting the time-averaged model $\bar{y}_i = \bar{x}_i'\beta + \boldsymbol{\gamma^T}\overline{\mathbf{Controls}}_i + \eta_i + \bar{\delta}_t + \bar{u}_i$ from Equation 1. Then,

$$\begin{aligned} y_{it} - \bar{y}_i = (x_{it} - \bar{x}_i)'\beta + \boldsymbol{\gamma^T}(\mathbf{Controls}_{it} - \overline{\mathbf{Controls}}_i) + (\delta_t - \bar{\delta}_t) + (u_{it} - \bar{u}_i) \Leftrightarrow \\ \ddot{y_{it}} = \ddot{x_{it}}\beta + \boldsymbol{\gamma^T}\ddot{\mathbf{Controls}}_i + \ddot{\delta}_t + \ddot{u_{it}} \end{aligned} \tag{3}$$

the fixed effect $\eta_i$ and all regressors that do not vary over time are removed.

The FE estimator corresponds to the application of the Ordinary Least Squares (OLS) estimator to the Equation 3. In short panels (large $N$ and small $T$), the estimates $\hat{\eta}_i$ are inconsistent. Nonetheless, $\hat{\beta}_{FE}$ is consistent for $\beta$, due to the fact that $\eta_i$ are eliminated in Equation 3. Given the high-dimensionality mentioned above, the standard FE estimator may be unfeasible, lack efficiency and lead to overfitting if the number of covariates is very large. Penalized methods such as LASSO provide an appropriate complement. The combination of within transformation and LASSO methods is used by Belloni et al. (2016). Given the computational implementation of the chosen method, we will need to further transform the Equation 3 through the FWL theorem explained next.

The Frisch & Waugh (1933) and Lovell (1963) theorem provides an alternative methodology to compute the OLS estimators in a linear regression model. This theorem allows

to transform the Equation 3 in a model which writes the (transformed) dependent variable just as a function of the (transformed) regressors, $x_{it}$, which we will penalize in the next section.

Consider the Equation 3 and the FWL theorem, we will denominate our equation as followed to simplify the explanation:

$$z = \mathbf{W^T}_1\boldsymbol{\beta}_1 + \mathbf{W^T}_2\boldsymbol{\beta}_2 + e \tag{4}$$

where $z$ is the dependent variable as in Equation 3, $\mathbf{W}_1$ is the vector of the variables of interest, $\ddot{x_{ij}}$, $\mathbf{W}_2$ is the vector of the variables $\ddot{\mathbf{Controls}}_i$, and $\ddot{\delta}_t$, and $e$ is the error term.

In order to obtain the parameters of interest, $\boldsymbol{\beta}_1$, we must:

1. Regress $z$ on $\mathbf{W}_2$ and save the residuals $\mathbf{M}_2\mathbf{z}$,
2. Regress $\mathbf{W}_1$ on $\mathbf{W}_2$ and save the residuals $\mathbf{M}_2\mathbf{W}_2$,
3. Regress the residuals obtained from the first step on those derived in the second step.

In this context, $\mathbf{M}_2 = \mathbf{I} - \mathbf{W}_2(\mathbf{W}_2'\mathbf{W}_2)^{-1}\mathbf{W}_1'\mathbf{M}_2$ represents the projection operator onto the orthogonal complement of the column space spanned by $\mathbf{W}_2$.

The estimated coefficient of interest is then presented as

$$\hat{\boldsymbol{\beta}}_1 = (\mathbf{W}_1'\mathbf{M}_2\mathbf{W}_1)^{-1}\mathbf{W}_1'\mathbf{M}_2\mathbf{z} \tag{5}$$

Due to the computational implementation in *R software* of the method used further, the FWL theorem will be a fundamental step to its application, since it allows to transform our within model, Equation 3, to a reduced form, where the (transformed) dependent variable, $\ddot{y}_{it}$, is explained just by the (transformed) variables of interest, $\ddot{x}_{ij}$, expunged of the control variables and time fixed effects. The model used for variable selection is the following

$$\mathbf{M}_2\mathbf{z} = \mathbf{M}_2\mathbf{W}_1\boldsymbol{\beta} + \mathbf{e} \tag{6}$$

### *4.3 Least Absolute Shrinkage and Selection Operator*

The base framework for the variable selection methodology used here considers a substantial number of predictors (high-dimensionality), to which only a few affect the

variable of interest (sparsity). More exactly, we have $p = 81$ (two digit) economic sectors in Portugal, coded as dummies, multiplied by the CEO salary of each firm, of which a subset $s < p$ is presumed to have non-zero effects. In order to identify those economic sectors, is crucial the application of statistical methods to the model described in Equation 6.

Due to the high dimensionality of $x_{ij}$, testing multiple subsets of economic sectors to find the one with the best performance is not practical nor straightforward. Instead, we employ a penalized regression approach that consists in an additional penalty component incorporated into the least squares loss function that would be used to estimate the un-penalized linear model. The penalty term induces shrinkage of all estimated coefficients toward zero, potentially setting some of them exactly to zero. Increasing the penalty strength results in the selection of fewer variables with non-zero coefficients. Keeping the design of the methodological technique in mind, the selected variables should be the ones that have major influence in the fit of the model; variables with less influence should be shrunken to exactly zero.

In order to select the economic sectors in which the CEO salary matters to firm productivity, we will apply the LASSO method, due to its characteristics of prediction and variable selection when we have a significant number of predictors, which may cause a problem of overfitting. The LASSO method produces predictions of a response with high-dimensional data, coping with a very large number of regressors as long as not many of them are actually relevant (sparsity assumption). The originated predictions can be based on non-parametric methods and be better than those obtained by a linear regression. The LASSO method overcomes the problem of overfitting and its purpose is to predict. In order to ensure a good prediction, we must choose a form of regularization, a penalty for model complexity, in which LASSO finds the optimum form of regularization, balancing the goodness-of-fit and parsimony, through the addition of a complexity penalty to a standard estimation problem. This method assumes either that the observations are independent or that the response is conditionally independent given the predictors.

Following Tibshirani (1996) and Equation 6, the LASSO estimate is defined by

$$\hat{\beta}^{lasso} = \arg\min_{\beta} \left\{ \frac{1}{2} \sum_{i,t,j} \left( (\mathbf{M_2 z})_{it} - \sum_{j=1}^{p} (\mathbf{M_2 W}_1)_{j,it} \beta_j \right)^2 + \lambda \sum_{j=1}^{p} |\beta_j| \right\} \tag{7}$$

where $p$ is the number of penalized variables, and $\lambda \geqslant 0$ is the tuning parameter that determines the penalty, which controls the amount of shrinkage that will affect the parameters.

In Equation 7, the first part of the sum is the standard least squares objective function, followed by the penalty for the model's complexity (regularization). Thus, the LASSO nests the least squares as a particular case when $\lambda$ is assigned a value of zero. The fundament of the $l_1$-penalty applied to the coefficients allows for the number of these to exceed the sample size, provided the turning parameter is sufficiently large. Conceptually, increasing the penalty progressively shrinks the estimated $\beta$-coefficients toward zero. The coefficients of the variables that do not sufficiently increment the objective function are set to zero, identifying the variables to include in the final model. Consequently, as $\lambda \to \infty$, Equation 7 is minimized by setting all estimated $\beta$ coefficients to zero, implying that no variables are selected.

The control parameter $\boldsymbol{\gamma}$ from Equation 3, as well as the FE and $\delta_t$, are not penalized. Additionally, they do not present any specific issue for computation because they are independent of the penalty parameter, being taken into account in the within transformation and FWL steps described above. Given that, for any $\beta$, the control parameter is taken into account in the within transformation and can be derived by optimizing the standard least squares objective function from the standard unpenalized regression approach.

Although LASSO offers significant computational benefits and strong empirical performance, it is inherently structured to select individual predictor variables rather than broader factor groupings. When LASSO is applied to Equation 6, it conducts variable selection based on the magnitude of individual derived input variables, rather than evaluating the collective strength of groups of variables. This may result in over-selection of factors. This is the reason the group LASSO can be more suitable for factor selection.

### *4.4 Group LASSO*

In this case, the economic sectors belong to pre-defined groups by one-digit economical classification. Given this, it may be preferable to apply shrinkage and selection at the group level, treating all variables within a group jointly, at least as a first step in variable selection. To achieve this, we apply the group LASSO by Yuan & Lin (2006).

The group LASSO extends the traditional LASSO methodology to facilitate variable selection at the group level. However, it is more challenging computationally. To implement this, a unified algorithm known as Groupwise-Majorization-Descent (GMD) is employed to efficiently compute the solution for the group LASSO penalized otimization. This algorithm satisfies the quadratic majorization condition.[6]

Suppose that the $p = 81$ economic sectors at two-digit classification are aggregated

[6]For more details in the quadratic majorization condition, see Yuan & Lin (2006).

at one-digit classification into $K = 10$ groups, with $p_k$ as the number of predictors in the group $k$. Considering Yuan & Lin (2006) and Equation 6, the group LASSO estimate is defined by

$$\hat{\beta}^{GL} = \arg\min_{\beta} \left\{ \frac{1}{2} \sum_{i,t,j} \left( (\mathbf{M}_2 \mathbf{z})_{it} - \sum_{j=1}^{p} (\mathbf{M}_2 \mathbf{W}_1)_{j,it} \beta_j \right)^2 + \lambda \sum_{k=1}^{K} \sqrt{p_k} \sqrt{\sum_{j \in I_k} \beta_j^2} \right\} \quad (8)$$

where $p_k$ represents the number of two-digit sectors in group $k$ and $I_k$ is the identity matrix of group $k$. When we penalize the coefficients in the second part of the Equation 8, it is applied the Euclidean norm to the subset of coefficients of group $k$. This approach promotes sparsity at both the group and individual levels, as the Euclidean norm of the subset of the coefficients of group $k$ is zero only if all of its components are zero. Consequently, for specific values of $\lambda$, an entire group of predictors may be excluded from the model.

### *4.5 Cross-Validation*

Now, we must discuss how to determine the parameter $\lambda$. We will examine the cross-validation approach and expose its strengths and weaknesses. The cross-validation approach iteratively partitions the data, selecting subsets to evaluate, and calculates the $\lambda$ that maximizes the model's predictive accuracy on the test data. Stated differently, it identifies the model that achieves the lowest forecasting error, commonly referred to as the cross-validation function. Formally, it calculates

$$\hat{\lambda} = \arg\min_{\lambda \in \Theta} CV(\lambda) \quad (9)$$

where $\Theta$ denotes the set of $\lambda$ grid values and it is computed in the following way:

1. Randomly divide the dataset into $K$ folds (in this case, $K = 10$),
2. For each fold $k \in \{1, ..., 10\}$, execute the following steps
   (a) Using the observations outside fold $k$, estimate the model parameters for a given $\lambda$,
   (b) Utilize the estimates obtained in step 2a to calculate the out-of-sample deviance for the observations in fold $k$,
3. Compute the mean of the out-of-sample deviance for each model $\lambda$,

4. The value of $\lambda \in \Theta$ that yields the lowest average out-of-sample deviance corresponds to the minimizer of the cross-validation function.

The primary advantage of cross-validation consists in its explicit focus to optimize predictive performance, thereby providing a conceptual benefit in forecasting contexts. Nevertheless, the LASSO procedure combined with cross-validation tends to over-select, often including irrelevant variables (Breinlich et al. 2021). Although, it is the standard technique to choose the regularization parameter in penalized regressions.

## 5 EMPIRICAL RESULTS

This section presents the results obtained using the group LASSO method as implemented by Yuan & Lin (2006), followed by the baseline LASSO (Tibshirani 1996). Initially, the analysis focus on economic sectors at two-digit level aggregated in ten groups of economic sectors at one-digit. Subsequently, after excluding the sectors that were not selected in the initial stage, the baseline lasso is applied to the sectors disaggregated at the two-digit level.

First, the group LASSO is applied to the $K = 10$ groups to identify the relevant variables. At the optimal tuning parameter value, $\lambda = 0.000006$, the group LASSO procedure excludes three economic sectors at one-digit level in which CEO salary influences firm productivity.

For further illustration, Figure 1 shows the regularization paths for the selected grouped economic sectors at the one-digit level. This figure demonstrates that the variation in the estimated coefficients for the ten grouped economic sectors changes as $\lambda$ varies. Consistent with theoretical expectations, the number of selected sectors decreases as $\lambda$ increases.

Figure 1 shows the non-selected economic sectors by group LASSO: extractive and certain manufacturing industries, as well as the management of electricity, gas, waste and related services. Within the manufacturing sector, the method excludes the manufacture of pharmaceutical products, rubber, plastic, and metal articles. Additionally, it does not select manufacture of machinery and equipment, nor furniture and mattresses.[7] As the tuning parameter $\lambda$ increases, two sectors, specifically sectors 4 and 5, distinctly stand out as being consistently selected even with high regularization. These sectors encompass construction; wholesale and retail trade, including vehicle repair for automobiles and motorcycles; transport and storage; accommodation and food services; as well as activities related to cinema, editing, and television. These findings are consistent with the

[7] In the Appendix A.2, it is presented the non-chosen economic sectors in more detail.

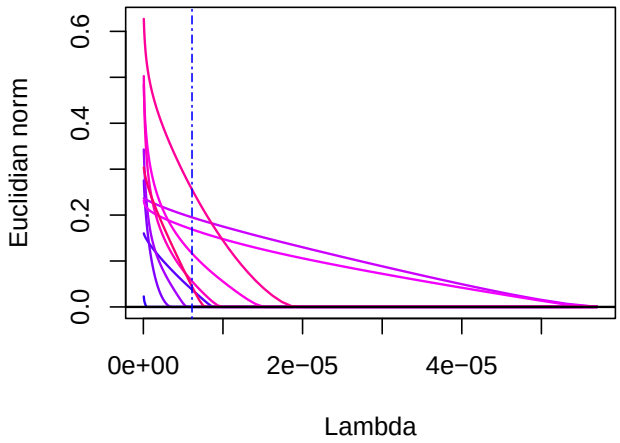



FIGURE 1: Regularization paths for selected grouped sectors with $n = 291,509$ and $p* = 81$.

Regularization paths of group LASSO penalized linear regression model with $p* = 81$ expanded variables organized into $K = 10$ groups, where the Euclidian norm of the coefficients of each group $k$ is plotted as a function of $\lambda$. The vertical dotted line denotes the optimal $\lambda$ value selected by cross-validation.

prevailing business paradigm observed among the largest firms in Portugal, aligning well with intuitive expectations. Given the non-selected economic sectors, we would also expect the energy and pharmaceutical sectors to be chosen, although its productivity may be mainly influenced by production factors, due to the highly specialization and investment in Research & Development (R&D) of these sectors.

In order to achieve a greater parsimonious result, we excluded the previously unselected groups and applied the LASSO method to the individual sectors. The variables selected in the second step form the final set of predictors. This ensures that the model is parsimonious, while including all relevant variables that have a potential causal effect on the dependent variable. Table II presents the results from the LASSO described above. Intuitively, the selected economic sectors reflect the broad Portuguese economic sector. The LASSO method selects 45 economic sectors at two-digit, from which the construction sector (sectors 41, 42 and 43), trade sector (sectors 45, 46 and 47) and accommodation and food services (sectors 55 and 56) clearly stand out. Additionally, chooses specialized work, such as information and communication activities (sectors 58, 59, 60, 61 and 62), financial services and insurance (sectors 64, 65 and 66), consulting, scientific and technical activities (sectors 69, 70, 71, 72, 73 and 75), as well as education (sector 85) and health and social support services (sectors 86, 87 and 88). Most of the selected economic

sectors exhibit a positive sign between firm productivity and CEO salary, from which is interesting to highlight the telecommunications sector and residential care activities. Conversely, ten sectors exhibit a negative coefficient, from which we highlight garment industry and R&D activities.

Particularly, one of the sectors in which CEO salary has a (positive) effect on firm productivity is the telecommunications sector (sector 61). This sector is responsible for enabling the transfer of information from a sender to one or more recipients, in a usable form, across various formats. The positive relationship between CEO salary and firm productivity can be attributed to the high complexity inherent in the sector, characterized by large-scale infrastructure, rapid technological advancements and extensive regulatory requirements. A competitive CEO salary is necessary to attract and retain a qualified CEO capable of managing the complexity of the sector, anticipating technological innovations, strategically investing on them, and sustaining competitive advantage within the industry.

Similarly, residential care activities within the non-profit sector (sector 87) demonstrate a positive relationship between CEO salary and firm productivity. This sector comprises services aimed at accommodation and social integration, primarily targeting children, the elderly, disabled people, and families experiencing social exclusion. Beyond organizational size, CEO salary in this sector is significantly influenced by ideological considerations and the composition of revenue streams. Non-profit organizations are precluded from distributing profits; rather, all revenue must be fully allocated toward fulfilling their mission objectives. Consequently, CEO salaries in the non-profit sector are subject to stringent scrutiny, and excessively high compensation may be perceived as either fraudulent or wasteful. Oster (1998) finds that CEO salary structures in non-profit organizations closely resemble those in for-profit firms. Additionally, competitive CEO salary is necessary to attract highly qualified executives and to motivate and retain current leadership.

The garment industry (sector 14) includes the production of all types of garments, regardless of material, for men, women and children. The negative coefficient associated with this sector may be attributed to the relatively small size of both the industry and the firms within it. Typically, garment firms report lower revenues compared to other industries, and many eventually file for bankruptcy or are acquired by larger competitors. According to Ozsoz (2025), fashion industry CEOs tend to receive lower salaries than the average CEO. Given that higher compensation is generally thought to motivate CEOs to enhance firm performance profitability, it is surprisingly that in this sector an increment of the CEO salary penalizes productivity, specially in a key Portuguese economic sector. It may be necessary to do a more granular study on this economic sector because this sug-

TABLE II: LASSO RESULTS.

| Variables | Lasso (1) | Variables | Lasso (1) |
|---|---|---|---|
| $Sector_{10}$ | 0.08 | $Sector_{69}$ | 0.04 |
| $Sector_{11}$ | -0.04 | $Sector_{70}$ | 0.07 |
| $Sector_{14}$ | -0.07 | $Sector_{71}$ | 0.07 |
| $Sector_{16}$ | 0.06 | $Sector_{72}$ | -0.68 |
| $Sector_{19}$ | -0.08 | $Sector_{73}$ | 0.01 |
| $Sector_{41}$ | 0.14 | $Sector_{75}$ | 0.03 |
| $Sector_{42}$ | 0.04 | $Sector_{78}$ | 0.04 |
| $Sector_{43}$ | 0.14 | $Sector_{79}$ | 0.17 |
| $Sector_{45}$ | 0.07 | $Sector_{80}$ | -0.10 |
| $Sector_{46}$ | -0.01 | $Sector_{81}$ | 0.10 |
| $Sector_{47}$ | 0.01 | $Sector_{82}$ | 0.05 |
| $Sector_{51}$ | 0.01 | $Sector_{84}$ | 0.51 |
| $Sector_{55}$ | 0.14 | $Sector_{85}$ | 0.19 |
| $Sector_{56}$ | 0.13 | $Sector_{86}$ | 0.06 |
| $Sector_{58}$ | -0.05 | $Sector_{87}$ | 0.34 |
| $Sector_{59}$ | 0.09 | $Sector_{88}$ | 0.09 |
| $Sector_{60}$ | 0.19 | $Sector_{90}$ | -0.07 |
| $Sector_{61}$ | 0.40 | $Sector_{92}$ | 0.01 |
| $Sector_{62}$ | 0.14 | $Sector_{93}$ | 0.03 |
| $Sector_{64}$ | 0.22 | $Sector_{94}$ | 0.22 |
| $Sector_{65}$ | -0.10 | $Sector_{95}$ | 0.04 |
| $Sector_{66}$ | 0.02 | $Sector_{96}$ | 0.09 |
| $Sector_{68}$ | -0.01 | | |

Linear regression with the selected economic sectors, estimated by LS using 291,509 firm-year observations. Column "Lasso" reports the coefficients estimated by LASSO. Refer to Appendix A.3 to a description of the selected sectors.

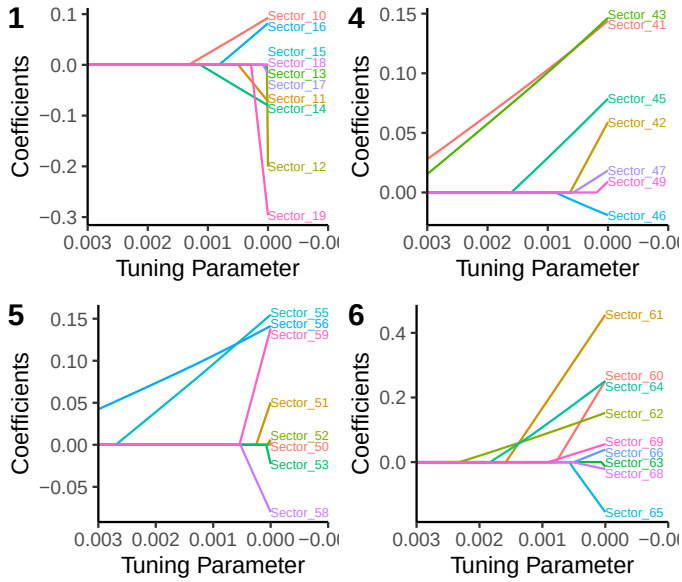
gests that an increase of CEO salary might reduce the workers' motivation or accelerate possible bankruptcies.

Scientific R&D activities comprise all fields of human knowledge, whether for profit or not. The negative relationship between CEO salary and firm productivity in the R&D sector (sector 72) is striking, possibly due to the difficulty of measuring productivity, since it is not linked to meaningful innovation outcomes, providing an appearance of a negative impact. R&D results only materialize in the long term, whereas CEO salary is tied to short-term metrics, which may lead to sustained investment in R&D being neglected.

Figure 2 presents the regularization paths for all two-digits economic sectors selected in the initial group-LASSO step, as illustrated in Figure 1, and corroborates the selected variables in Table II. The regularization paths serve to validate the reliability and consistency of our findings. A closer inspection to Figure 2a reveals that the most persistent sectors are: real estate promotion and specialized construction (sectors 41 and 43); trade, maintenance and repair of motor vehicles and motorcycles (sector 45); accommodation and food service (sectors 55 and 56); telecommunications, consulting and computer programming services (sectors 61 and 62); and financial services, except insurance and pension funds (sector 64).

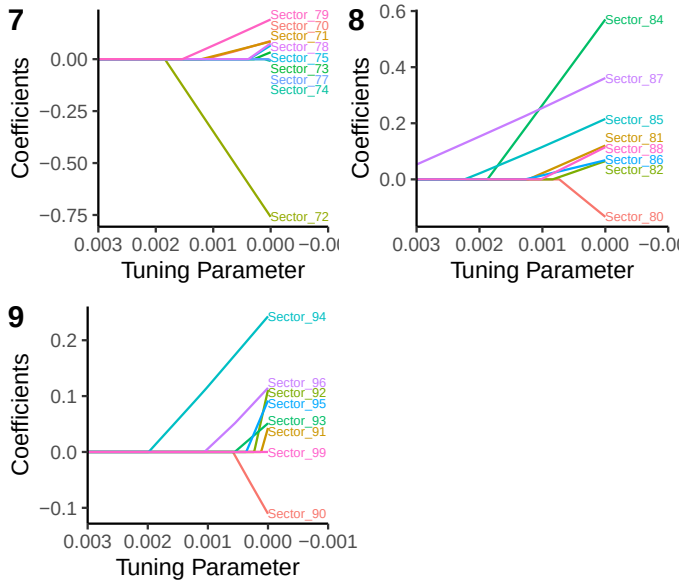
The persistent selection observed in the real estate promotion and specialized construction sectors is not surprising, given that the CEO salary in firms operating within this sector ranks among the highest in Portugal. This phenomenon can be attributed to the expansion of the real estate market, the concomitant demand for new private and public infrastructure, and a disequilibrium between housing demand and supply - specifically, demand exceeding supply. The tourism sector constitutes one of the most prominent sectors in Portugal; consequently, the sustained selection of the accommodation and food service sectors are anticipated, alongside the evident correlation between CEO salary and firm productivity. The largest and most innovative firms in Portugal are found in the telecommunications, consultancy, computer programming, and financial services sectors, so an increase in CEO salary contributes to a more effective market strategy and, subsequently, an increment in productivity.

Observing Figure 2b, we conclude that the most persistent (two-digit) economic sectors are: R&D activities (sector 72); travel agencies and tour operators (sector 79); public administration and defense (sector 84); education (sector 85); social support activities with accommodation (sector 87); activities of associative organizations (sector 94); and other personal services (sector 96). The aforementioned sectors impact our social life to some extent, encompassing areas such as defense, education, leisure and living conditions. Accordingly, the persistent influence of CEO salaries on firm productivity is




(a) Regularization paths for the sectors within the groups 1, 4, 5, and 6 with $n = 291,509$ and $p = 59$.

Regularization paths for the LASSO penalized linear regression model with $p = 59$ variables, where the value of each coefficient $\beta_j$ is plotted against $\lambda$ for $j = 1, ..., 59$. The economic sectors are presented with respect to their respective sector at one-digit level.




(b) Regularization paths for the sectors within the groups 7, 8, and 9 with $n = 291,509$ and $p = 59$.

Regularization paths for the LASSO penalized linear regression model with $p = 59$ variables, where the value of each coefficient $\beta_j$ is plotted against $\lambda$ for $j = 1, ..., 59$. The economic sectors are presented with respect to their respective sector at one-digit level.

FIGURE 2: Regularization paths for two-digit economic sectors using LASSO penalization.

unexpected, given that CEO salary in these sectors is typically anticipated to be lower than that of CEOs in profit-driven firms. Nevertheless, a high salary is necessary to attract qualified and sector-appropriate CEOs, as well as to motivate those already in these positions, thereby influencing productivity.

Referring to the OLS estimation, most of the persistently selected economic sectors by LASSO correspond to the statistically significant regressors at 95% level, corroborating our LASSO results.[8]

## 6 CONCLUSION

This study investigates the causal effects of CEO salary on firm productivity outcomes, specifically focusing on the economic sectors. Given the central role of CEOs in shaping a firm's strategic direction and risk profile - both of which are closely tied to external performance-relevant factors - this research seeks to determine in which economic sectors CEO salary serves as a driver of productivity. In light of the ongoing debate on whether rising CEO salaries reflect talent scarcity or performance-based incentives, we adopt advanced variable selection methods to identify sectors where CEO salary significantly influences productivity. The findings indicate that CEO salary influences persistently firm productivity in the construction sector, trade and repair of motor vehicles and motorcycles, telecommunications, as well as accommodation and food services. These results align with the prevailing business paradigm in Portugal. Additionally, the research indicates that CEO salary does not influence firm productivity in the extractive industries, certain manufacturing industries (e.g., pharmaceuticals, rubber, plastics, and metal products), and in the management of electricity, gas, waste and related services.

Our methodology proceeds in two stages. Initially, we classify economic sectors at the one-digit level to identify broad patterns, implementing the group LASSO approach (Yang & Zou 2015). Subsequently, we refine this categorization to isolate more persistent and granular sector-specific effects with baseline LASSO. These recent machine learning techniques are applied on a unique and rich matched employer-employee dataset.

Among the economic sectors where the relationship between CEO salary and firm productivity is presented, ten sectors exhibit a negative coefficient, in which the garment industry and scientific R&D activities are included. Notably, two of the sectors demonstrating a positive impact of CEO salary on firm productivity are telecommunications and residential care activities within the non-profit sector.

Overall, the results are consistent with the Gabaix & Landier (2008) findings, who

[8] We refer to A.4 for the coefficients and significance of the regressors derived from the OLS estimation.

emphasize firm size as the main factor of the relationship between CEO salary and firm productivity. In garment industry, relatively low CEO salaries appear to hinder competitiveness, increasing vulnerability to bankruptcy or acquisition (Ozsoz 2025). In R&D activities, the difficulty of measuring productivity may appeal as a negative impact of CEO salary in productivity. This pattern is attributed to structural characteristics of the sector. In the telecommunications sector, a high CEO salary attracts a qualified CEO capable of dealing with its complexity. In non-profit residential care activities - where the CEO salary impacts more significantly the firm productivity - the results are defended by Oster (1998), who argues that a higher salary is necessary to attract and retain qualified executives within the non-profit sector, being influenced by ideological considerations and by the composition of the revenue streams.

A broad conclusion that can be drawn is that CEO salary serves as an instrument primarily aimed at attracting and retaining qualified CEOs who are well-suited to the specific sector in which the firm operates. The selection of an appropriate CEO improves the firm's strategic direction, creates a sustainable competitive position in the market, increases productivity and ultimately contributes to the overall firm value.

The findings contribute to the growing literature on CEO salary by offering a novel empirical application of LASSO techniques and providing sector-specific insights into the productivity effects of CEO salary. To our knowledge, this is the first study to combine the group LASSO and standard LASSO in this context, suggesting the method's potential applicability across national and international datasets.

One limitation of the study is the identification and assignment of a CEO to each firm. Given the absence of a scientific definition of CEO, it was necessary to develop a selection method, drawing on existing literature and adapted to the characteristics of our dataset. The selection of penalized regressors and the determination of an appropriate tuning parameter for the LASSO estimation also presented a methodological preference. While LASSO method overcomes the problem of overfitting, our approach involved a hierarchical penalization: first at the one-digit sector level, and subsequently, within selected sectors, at the two-digit level for greater granularity. LASSO techniques exhibit multiple methodological options to choose the tuning parameter. After applying multiple selection techniques and obtaining consistent results across methods, we adopted the approach explained in Section 4. The large volume of results derived from the implemented methods may overlook the analysis of certain interesting individual outcomes. This paves the way for future analyses focused on specific sectors, using this work as a reference point.

Future research may expand upon this work in several ways. First, expanding the longitudinal data or applying datasets from other countries could enhance generaliza-

tion and contextual understanding. Second, the integration of more advanced machine learning techniques, such as double machine learning, could improve model robustness. Third, analysing alternative productivity metrics, including estimates based on the Cobb-Douglas production function, would provide a more comprehensive understanding of the impact of CEO salary on firm productivity in the various economic sectors. Lastly, examining the influence of CEO characteristics (e.g., gender, age, or experience) across economic sectors could further refine the theoretical and empirical understanding of leadership dynamics and firm performance.

## REFERENCES


Artige, L. & Bignandi, S. (2023), 'The firm size distribution: evidence from belgium', *Applied Economics* **55**(8), 907–923.
**URL:** *https://doi.org/10.1080/00036846.2022.2095339*

Bebchuk, L. A. & Fried, J. M. (2003), 'Executive compensation as an agency problem', *Journal of Economic Perspectives* **17**(3), 71–92.
**URL:** *https://www.aeaweb.org/articles?id=10.1257/089533003769204362*

Bebchuk, L. A., Fried, J. M. & Walker, D. I. (2002), 'Managerial power and rent extraction in the design of executive compensation', *The University of Chicago Law Review* **69**(3), 751–846.
**URL:** *http://www.jstor.org/stable/1600632*

Bebchuk, L. & Fried, J. (2004), *Pay without Performance: The Unfulfilled Promise of Executive Compensation*, Harvard University Press.
**URL:** *http://www.jstor.org/stable/j.ctv2jfvcp7*

Belloni, A., Chernozhukov, V., Hansen, C. & and, D. K. (2016), 'Inference in high-dimensional panel models with an application to gun control', *Journal of Business & Economic Statistics* **34**(4), 590–605.
**URL:** *https://doi.org/10.1080/07350015.2015.1102733*

Berger, P. G., Ofek, E. & Yermack, D. L. (1997), 'Managerial entrenchment and capital structure decisions', *The Journal of Finance* **52**(4), 1411–1438.
**URL:** *http://www.jstor.org/stable/2329441*

Bertrand, M. & Mullainathan, S. (2001), 'Are ceos rewarded for luck? the ones without principals are*', *The Quarterly Journal of Economics* **116**(3), 901–932.
**URL:** *https://doi.org/10.1162/00335530152466269*

Bhuyan, R., Butchey, D., Haar, J. & Talukdar, B. (2020), 'Ceo compensation and firm performance in the insurance industry', *Managerial Finance* **48**(7), 1086–1115.
**URL:** *https://doi.org/10.1108/MF-04-2019-0154*

Breinlich, H., Corradi, V., Rocha, N., Ruta, M., Santos Silva, J. & Zylkin, T. (2021), Machine learning in international trade research: Evaluating the impact of trade agreements, Policy Research Working Paper Series 9629, The World Bank.
**URL:** *https://EconPapers.repec.org/RePEc:wbk:wbrwps:9629*

Bulan, L., Sanyal, P. & Yan, Z. (2010), 'A few bad apples: An analysis of ceo performance pay and firm productivity', *Journal of Economics and Business* **62**(4), 273–306.
**URL:** *https://www.sciencedirect.com/science/article/pii/S0148619510000123*

Cameron, A. C. & Trivedi, P. K. (2005), *Microeconometrics: Methods and Applications*, Cambridge University Press.

Coles, J. L., Daniel, N. D. & Naveen, L. (2006), 'Managerial incentives and risk-taking', *Journal of Financial Economics* **79**(2), 431–468.
**URL:** *https://www.sciencedirect.com/science/article/pii/S0304405X05000802*

Cuñat, V. & Guadalupe, M. (2009), 'Globalization and the provision of incentives inside the firm: The effect of foreign competition', *Journal of Labor Economics* **27**(2), 179–212.
**URL:** *https://doi.org/10.1086/599817*

Dai, S., Kuosmanen, N., Kuosmanen, T. & Liesiö, J. (2025), 'Optimal resource allocation: Convex quantile regression approach', *European Journal of Operational Research* **324**(1), 221–230.
**URL:** *https://www.sciencedirect.com/science/article/pii/S0377221725000037*

Demiraj, R., Labadze, L., Dsouza, S., Demiraj, E. & Grigolia, M. (2024), 'The quest for an optimal capital structure: an empirical analysis of european firms using gmm regression analysis', *EuroMed Journal of Business* **20**(2), 529–551.
**URL:** *https://doi.org/10.1108/EMJB-07-2023-0206*

Dow, J. & Raposo, C. C. (2005), 'Ceo compensation, change, and corporate strategy', *The Journal of Finance* **60**(6), 2701–2727.
**URL:** *https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1540-6261.2005.00814.x*

Friedman, J. H., Hastie, T. & Tibshirani, R. (2010), 'Regularization paths for generalized linear models via coordinate descent', *Journal of Statistical Software* **33**(1), 1–22.
**URL:** *https://www.jstatsoft.org/index.php/jss/article/view/v033i01*

Frisch, R. & Waugh, F. V. (1933), 'Partial time regressions as compared with individual trends', *Econometrica* **1**(4), 387–401.
**URL:** *http://www.jstor.org/stable/1907330*

Frydman, C. (2019), 'Rising through the ranks: The evolution of the market for corporate executives, 1936–2003', *Management Science* **65**(11), pp. 4951–4979.
**URL:** *https://www.jstor.org/stable/48760849*

Fukui, T., Mitton, T. & Schonlau, R. (2023), 'Determinants of capital structure: An expanded assessment', *Journal of Financial and Quantitative Analysis* **58**(6), 2446–2488.

Gabaix, X. & Landier, A. (2008), 'Why has ceo pay increased so much?*', *The Quarterly Journal of Economics* **123**(1), 49–100.

Garicano, L. & Rossi-Hansberg, E. (2006), 'Organization and inequality in a knowledge economy*', *The Quarterly Journal of Economics* **121**(4), 1383–1435.
**URL:** *https://doi.org/10.1093/qje/121.4.1383*

Garvey, G. & Milbourn, T. (2003), 'Incentive compensation when executives can hedge the market: Evidence of relative performance evaluation in the cross section', *The Journal of Finance* **58**(4), 1557–1582.
**URL:** *https://onlinelibrary.wiley.com/doi/abs/10.1111/1540-6261.00577*

Gopalan, R., Milbourn, T. & Song, F. (2010), 'Strategic flexibility and the optimality of pay for sector performance', *The Review of Financial Studies* **23**(5), 2060–2098.
**URL:** *https://doi.org/10.1093/rfs/hhp118*

Green, C. P., Heywood, J. S. & Theodoropoulos, N. (2021), 'Hierarchy and the employer size effect on wages: Evidence from britain', *Economica* **88**(351), 671–696.
**URL:** *https://onlinelibrary.wiley.com/doi/abs/10.1111/ecca.12364*

Guay, W. R. (1999), 'The sensitivity of ceo wealth to equity risk: an analysis of the magnitude and determinants', *Journal of Financial Economics* **53**(1), 43–71.
**URL:** *https://www.sciencedirect.com/science/article/pii/S0304405X99000161*

Hall, B. J. & Murphy, K. J. (2003), 'The trouble with stock options', *Journal of Economic Perspectives* **17**(3), 49–70.
**URL:** *https://www.aeaweb.org/articles?id=10.1257/089533003769204353*

Haltiwanger, J. C., Lane, J. I. & Spletzer, J. R. (1999), 'Productivity differences across employers: The roles of employer size, age, and human capital', *The American Economic Review* **89**(2), 94–98.
**URL:** *http://www.jstor.org/stable/117087*

Hermalin, B. E. (2005), 'Trends in corporate governance', *The Journal of Finance* **60**(5), 2351–2384.
**URL:** *https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1540-6261.2005.00801.x*

Holmstrom, B. & Kaplan, S. N. (2001), 'Corporate Governance and Merger Activity in the United States: Making Sense of the 1980s and 1990s', *Journal of Economic*

*Perspectives* **15**(2), 121–144.
**URL:** *https://ideas.repec.org/a/aea/jecper/v15y2001i2p121-144.html*

Huergo, E. & Jaumandreu, J. (2004), 'Firms' age, process innovation and productivity growth', *International Journal of Industrial Organization* **22**(4), 541–559.
**URL:** *https://www.sciencedirect.com/science/article/pii/S0167718704000037*

Jensen, M. C. & Murphy, K. J. (1990), 'Performance pay and top-management incentives', *Journal of Political Economy* **98**(2), 225–264.
**URL:** *http://www.jstor.org/stable/2937665*

Kaplan, Z. & Pérez-Cavazos, G. (2021), 'Investment as the opportunity cost of dividend signaling', *The Accounting Review* **97**.

Lanau, S. & Topalova, P. (2016), 'The Impact of Product Market Reforms on Firm Productivity in Italy', *IMF Working Papers* **16**(119), 1.
**URL:** *http://dx.doi.org/10.5089/9781475524925.001*

Leone, A. J., Wu, J. S. & Zimmerman, J. L. (2006), 'Asymmetric sensitivity of ceo cash compensation to stock returns', *Journal of Accounting and Economics* **42**(1), 167–192. Conference Issue on Implications of Changing Financial Reporting Standards.
**URL:** *https://www.sciencedirect.com/science/article/pii/S0165410106000279*

Lovell, M. C. (1963), 'Seasonal adjustment of economic time series and multiple regression analysis', *Journal of the American Statistical Association* **58**(304), 993–1010.
**URL:** *http://www.jstor.org/stable/2283327*

Mensah, E., Owusu, S., Foster-McGregor, N. & Szirmai, A. (2022), 'Structural change, productivity growth and labour market turbulence in sub-saharan africa', *Journal of African Economies* **32**(3), 175–208.
**URL:** *https://doi.org/10.1093/jae/ejac010*

Murphy, K. J. & Zábojník, J. (2004), 'Ceo pay and appointments: A market-based explanation for recent trends', *American Economic Review* **94**(2), 192–196.
**URL:** *https://www.aeaweb.org/articles?id=10.1257/0002828041302262*

Oster, S. M. (1998), 'Executive compensation in the nonprofit sector', *Nonprofit Management and Leadership* **8**(3), 207–221.
**URL:** *https://onlinelibrary.wiley.com/doi/abs/10.1002/nml.8301*

Ozsoz, E. (2025), *Executive Fashion*, De Gruyter, Berlin, Boston.
**URL:** *https://doi.org/10.1515/9783111510156*

Queiró, F. (2021), 'Entrepreneurial human capital and firm dynamics', *The Review of Economic Studies* **89**(4), 2061–2100.
**URL:** *https://doi.org/10.1093/restud/rdab070*

Rajan, R. G. & Wulf, J. (2006), 'Are perks purely managerial excess?', *Journal of Financial Economics* **79**(1), 1–33.
**URL:** *https://www.sciencedirect.com/science/article/pii/S0304405X05001339*

Tervio, M. (2008), 'The difference that ceos make: An assignment model approach', *American Economic Review* **98**(3), 642–68.
**URL:** *https://www.aeaweb.org/articles?id=10.1257/aer.98.3.642*

Tibshirani, R. (1996), 'Regression shrinkage and selection via the lasso', *Journal of the Royal Statistical Society: Series B (Methodological)* **58**(1), 267–288.
**URL:** *https://doi.org/10.1111/j.2517-6161.1996.tb02080.x*

Yang, Y. & Zou, H. (2015), 'A fast unified algorithm for solving group-lasso penalize learning problems.', *Stat Comput* **25**, 1129–1141.

Yuan, M. & Lin, Y. (2006), 'Model selection and estimation in regression with grouped variables', *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **68**(1), 49–67.
**URL:** *https://rss.onlinelibrary.wiley.com/doi/abs/10.1111/j.1467-9868.2005.00532.x*

# A Appendices

## *A.1 Variables used in the productivity estimation*

- Productivity: logarithm of the gross value-added per worker of the firm.
- Public Capital: share of public social capital.
- Dimension: logarithm of the number of workers of the firm.
- Antiquity: the number of years since the beginning of the firm's activity.
- Unionized Workers: share of unionized workers (%).
- Salaried Workers: share of salaried workers (%).
- Number of Establishments: number of the establishments of the firm.

## *A.2 Economic sectors do not selected by group LASSO*

- Extraction and preparation of metallic minerals;
- Other extractive industries;
- Service activities related to extractive industries;
- Manufacture of chemical products and of synthetic or artificial fibers, except pharmaceutical products;
- Manufacture of basic pharmaceutical products and pharmaceutical preparations;
- Manufacture of rubber and plastic articles;
- Manufacture of other non-metallic mineral products;
- Basic metallurgical industries;
- Manufacture of metal products, except machinery and equipment;
- Manufacture of computer equipment, communication equipment, and electronic and optical products;
- Manufacture of electrical equipment;
- Manufacture of machinery and equipment, n.e.c.;

- Manufacture of motor vehicles, trailers, semi-trailers, and components for motor vehicles;
- Manufacture of other transport equipment;
- Manufacture of furniture and mattresses;
- Other manufacturing industries;
- Repair, maintenance, and installation of machinery and equipment;
- Electricity, gas, stream, hot and cold water, and cold air;
- Water collection, treatment, and distribution;
- Collection, drainage, and treatment of wastewater;
- Collection, treatment, and disposal of waste, material recovery;
- Decontamination and similar activities.

### *A.3 Economic sectors in which the CEO salary affects firm productivity in LASSO*

The description of the economic sectors is presented in the same order as that presented in Table II.

- **Food industries** encompass the processing of agricultural, pescatarian and animal production products into consumable goods for humans or animals, as well as the manufacture of intermediary products;
- **Beverage industries** include the production of alcoholic and non-alcoholic beverages;
- **Garment industries** involve the production of a wide range of garments for men, women and children, utilizing various materials;
- **Wood and cork industries, except furniture and manufacture of basketware and wickerwork** encompass sawing, planing, and impregnation of wood, manufacture of veneers, plywood, and panels, carpentry works for construction, packaging and other wood products, basketware, wickerwork, and cork processing;
- **Manufacture of refined petroleum, and combustible agglomerates** include the transformation of crude oil and coal into derivative products intended for consumption;

- **Real estate promotion (development of building projects) and building construction** refer to the development of building projects, including the construction, transformation, and repair of buildings for diverse sectors. This sector also includes the technical, legal, and financial involved in the construction and sale of such projects;
- **Civil engineering** encompass the construction of roads, bridges, tunnels, airport runways, railways, water supply, sewerage, energy distribution, telecommunications, and sport facilities. Hydraulic engineering is also included in this sector;
- **Specialized construction activities** comprise demolition, site preparation, construction drilling, fabrication and installation of iron frameworks at construction sites, electrical and plumbing installation, and the finishing of buildings;
- **Trade, maintenance and repair of motor vehicles and motorcycles** encompass the wholesale and retail trade of new and used motor vehicles, motorcycles and mopeds, including parts, components, and accessories, alongside their maintenance repair services;
- **Wholesale trade, except motor vehicles and motorcycles** comprise the resale of new or used goods on a wholesale basis without transformation, primarily to traders, professional users or intermediaries;
- **Retail trade, except motor vehicles and motorcycles** involve the resale of new or used goods on a retail basis without transformation, intended for final consumers, businesses, and other institutions;
- **Air transport** of passengers and goods;
- **Accommodation services** refer to the temporary rental of lodging facilities, with or without the provision of meals and ancillary services, whether open to the general public or reserved for members of specific organizations;
- **Restaurants and similar establishments** include the preparation and sale of food for on-site consumption, as well as related services such as catering, cafes, bars, and pastry shops;
- **Publishing activities** encompass publishing books, newspapers, magazines, software and multimedia products;

- **Cinematographic, video, television program production, sound recording and music publishing activities**, which include the production, post-production, distribution, screening, purchase and sale of distribution rights of films and videos. Additionally, comprise sound recordings and music publishing;
- **Radio and television activities** are related to the creation of content and the acquisition of rights to distribute them;
- **Telecommunications** enable the transfer of information from a sender to one or more recipients, in a usable form, across various formats;
- **Consulting and computer programming services** involve the design, development, modification, testing and technical support of computer software tailored to the specific requirements of clients;
- **Financial services, except insurance and pension funds** include activities related to the acquisition and distribution of financial resources;
- **Insurance, reinsurance and pension funds**, except compulsory social security;
- **Auxiliary activities for financial services and insurance**, without directly providing financial services;
- **Real estate activities** include the purchase, sale, rental, mediation and valuation of real estate, as well as property management on behalf of others;
- **Legal activities** consist of providing legal consultation and representation of natural or legal persons before the courts or other judicial entity, within the scope of civil, criminal, labour or other legal domains. This sector also includes **accounting services** such as recording of accounting operations, auditing and preparation of tax declarations;
- **Activities of head offices and management consultancy** include strategic and organizational planning, as well as operational assistance to companies;
- **Architectural, engineering and related techniques activities**, as well as test and technical analysis involve architecture and engineering services, technical testing, building inspections and cartography;
- **Scientific R&D activities** involve all fields of human knowledge, whether for profit or not;
- **Advertising, market research and opinion polls**;

- **Veterinary activities** comprise medical-veterinary care with or without hospitalization of farm and companion animals;
- **Employment activities** related to the search, selection, guidance and placement of personnel at all levels and in all fields, as well as labour subcontracting activities;
- **Travel agencies, tour operators and other reservation services** include agencies and tour operators engaged primarily in the sale of travel packages, tourism, transportation, and accommodation services on either a wholesale or retail basis to individuals and companies;
- **Investigation and security activities** include surveillance and protection activities carried out by specialized personnel or using other means of protection. Additionally, encompass inquiry and investigation activities outside the judicial framework, as well as detective activities;
- **Building-related activities and maintenance of garden** include cleaning services, security and surveillance, waste segregation and reception duties;
- **Administrative and support service activities** provided to businesses encompass financial planning, document preparation, call center operations, event organization and other related services;
- **Public administration, defense and compulsory social security** include general administrative activities at all levels of government, in the economic and social domains;
- **Education** comprise public, private and cooperative teaching across all levels or fields of study accessible to the general public;
- **Human health activities** take place in inpatient facilities, outpatient clinics and other healthcare settings. These activities involve a range of medical procedures, from diagnosis to treatment, and include services performed by legally authorized paramedical professionals;
- **Social support activities with accommodation** target children, youth, individuals with disabilities, and families experiencing social exclusion. This sector includes nursing homes and similar institutions providing integrated and continuous care;
- **Social support activities without accommodation** target population with need of social support in their homes or elsewhere;

- **Theatre, music, dance and other artistic and literary activities** related to the production and performance of those works, as well as the operation of venues intended for cultural and entertainment events;
- **Lotteries and other gambling activities** encompass the organization of the lotteries, operation of casinos and other gaming venues, as well as retailers, street and online vendors;
- **Sports, amusement and recreation activities** cover sports activities, bullfighting, theme parks, and others;
- **Activities of associative organizations** involve the management and administration of such entities, typically including representation before public authorities, dissemination of information and guidance, collective bargaining and public relations;
- **Repair of computers, personal and household goods**, such as peripheral, communication equipment, garden equipment, footwear, clothing, furniture, musical instrument, and other goods;
- **Other personal services activities** cover services such as textile and leather washing and dry cleaning, hairdressing and beauty salons, funeral services, wellness activities, among others.

### *A.4 OLS estimation*

TABLE III: OLS RESULTS.

| **Dependent variable**: Firm productivity (2022-2023) | | | |
|---|---|---|---|
| Variables | OLS (1) | Variables | OLS (2) |
| $Sector_{07}$ | 1.00 (1.19) | $Sector_{52}$ | 0.01 (0.04) |
| $Sector_{08}$ | 0.02 (0.05) | $Sector_{53}$ | -0.02 (0.09) |
| $Sector_{09}$ | -0.87 (2.25) | $Sector_{55}$ | 0.15*** (0.03) |
| $Sector_{10}$ | 0.09*** (0.02) | $Sector_{56}$ | 0.14*** (0.01) |
| $Sector_{11}$ | -0.07** (0.03) | $Sector_{58}$ | -0.08* (0.05) |
| $Sector_{12}$ | -0.20 (160.4) | $Sector_{59}$ | 0.14 (0.11) |
| $Sector_{13}$ | -0.01 (0.03) | $Sector_{60}$ | 0.25*** (0.06) |
| $Sector_{14}$ | -0.08*** (0.02) | $Sector_{61}$ | 0.46** (0.20) |
| $Sector_{15}$ | -0.001 (0.04) | $Sector_{62}$ | 0.15*** (0.03) |
| $Sector_{16}$ | 0.09 (0.06) | $Sector_{63}$ | -0.01 (0.13) |
| $Sector_{17}$ | -0.02 (0.06) | $Sector_{64}$ | 0.25*** (0.05) |
| $Sector_{18}$ | 0.001 (0.02) | $Sector_{65}$ | -0.15 (0.19) |
| $Sector_{19}$ | -0.30 (0.31) | $Sector_{66}$ | 0.04* (0.02) |
| $Sector_{21}$ | -0.01 (0.10) | $Sector_{68}$ | -0.02 (0.03) |
| $Sector_{22}$ | -0.09* (0.05) | $Sector_{69}$ | 0.06*** (0.01) |
| $Sector_{23}$ | 0.002 (0.06) | $Sector_{70}$ | 0.09*** (0.03) |
| $Sector_{24}$ | 0.11 (0.09) | $Sector_{71}$ | 0.09*** (0.02) |
| $Sector_{25}$ | 0.02 (0.02) | $Sector_{72}$ | -0.76** (0.29) |
| $Sector_{26}$ | 0.18** (0.08) | $Sector_{73}$ | 0.03 (0.04) |
| $Sector_{27}$ | 0.18*** (0.06) | $Sector_{74}$ | -0.01 (0.04) |
| $Sector_{28}$ | 0.01 (0.03) | $Sector_{75}$ | 0.07** (0.03) |
| $Sector_{29}$ | -0.14 (0.15) | $Sector_{77}$ | -0.01 (0.04) |
| $Sector_{30}$ | -0.29* (0.18) | $Sector_{78}$ | 0.08 (0.07) |
| $Sector_{31}$ | 0.06** (0.02) | $Sector_{79}$ | 0.19*** (0.06) |
| $Sector_{32}$ | -0.05 (0.05) | $Sector_{80}$ | -0.13 (0.12) |
| $Sector_{33}$ | 0.10*** (0.03) | $Sector_{81}$ | 0.12*** (0.04) |
| $Sector_{35}$ | -0.13 (0.24) | $Sector_{82}$ | 0.06 (0.05) |
| $Sector_{36}$ | -0.14 (0.21) | $Sector_{84}$ | 0.57 (0.37) |
| $Sector_{37}$ | -0.23 (0.23) | $Sector_{85}$ | 0.22*** (0.05) |
| $Sector_{38}$ | -0.002 (0.05) | $Sector_{86}$ | 0.07*** (0.01) |
| $Sector_{39}$ | 1.16 (0.81) | $Sector_{87}$ | 0.36*** (0.07) |
| $Sector_{41}$ | 0.14*** (0.02) | $Sector_{88}$ | 0.12* (0.07) |
| $Sector_{42}$ | 0.06** (0.03) | $Sector_{90}$ | -0.11 (0.26) |
| $Sector_{43}$ | 0.15*** (0.01) | $Sector_{91}$ | 0.04 (0.15) |
| $Sector_{45}$ | 0.08*** (0.01) | $Sector_{92}$ | 0.11 (0.24) |
| $Sector_{46}$ | -0.02 (0.01) | $Sector_{93}$ | 0.05 (0.05) |
| $Sector_{47}$ | 0.02** (0.01) | $Sector_{94}$ | 0.24*** (0.07) |
| $Sector_{49}$ | 0.01 (0.02) | $Sector_{95}$ | 0.09 (0.08) |
| $Sector_{50}$ | 0.003 (0.11) | $Sector_{96}$ | 0.12** (0.05) |
| $Sector_{51}$ | 0.05 (0.03) | | |

Linear regression with the selected economic sectors at two-digit level, estimated by OLS using 291,509 firm-year observations. Column "OLS" reports the coefficients estimated by OLS. Bootstrap cluster-robust standard errors are reported in parentheses. $*p-value<0.10$, $**p-value<0.05$, $***p-value<0.01$.

### *A.5 Implementation*

The analysis was conducted using *R software* as the primary programming language. The implementation of machine learning models was facilitated through various libraries, notably *gglasso* (Yang & Zou 2015) and *glmnet* (Friedman et al. 2010). These libraries provided a framework for model building, parameter tuning and evaluation, enabling a

comprehensive analysis of the impact of CEO salary on firm productivity in the selected economic sectors.