Better Together? CEO Identity and Firm Productivity*

Inês Black†

November 15, 2017

JOB MARKET PAPER
Click Here for Latest Version

Abstract

This paper uses a combination of a reduced-form approach and a distributional framework to analyze the impact of the CEO on firm revenue productivity. Using a matched employer-employee data set, I attempt to disentangle the role of the CEO quality type (identity) from that of the firm in revenue productivity and evaluate the existence and relevance of match-specific complementarities between CEO and firm. I present a proxy measure of CEO quality that takes advantage of differential patterns of CEO mobility throughout her career to circumvent endogenous CEO job mobility. I find that a one standard deviation increase in CEO quality results in 5% increase in firm production. Higher-quality CEOs are more likely to hold a higher education degree, have a larger experience as a manager, invest in innovation and less likely to work in a family firm. Moreover, results CEO-firm complementarities are significant in firm production, affecting the distribution of productivity positively in approximately 2-3%. The issue of CEO impact is of significant practical importance to firms and policy makers alike, as it can partly explain the rise in wage inequality.

JEL Codes: C14, J24, L22, M12
Keywords: Top-managers, Firm Productivity, CEO-Firm Match, Endogenous Mobility

---

*I am very grateful to my advisor Caterina Calsamiglia and my sponsor at University College London, Imran Rasul, for their guidance and support. I thank Alejandra Ramos, Alessandro Sforza, Ana Rute Cardoso, Andrew Garin, Gianluca Violante, Joan Llull, Laura Blattner, Pedro Portugal, Richard Blundell, Sanghmitra Gautam, Thibaut Lamadon and participants of the Applied Economics seminar at the UAB, seminar participants at UCL, the Barcelona PhD Jamboree 2017 and EEA-ESEM Summer Meeting 2017. I also acknowledge the invaluable support of the faculty and administrative staff at the Departments of Economics at UCL and Bank of Portugal's during my research visits. All mistakes are my own.

†IDEA, Department d'Economia i d'Història Econòmica, Universitat Autònoma de Barcelona (UAB), Edifici B Campus UAB 08193 Bellaterra, Barcelona, Spain. Email: ines.black.henriques@gmail.com.

1 Introduction

There is extensive documentation in the economics literature suggesting substantial productivity differentials within countries (Foster et al., 2008) and specific industries (Titman & Wessels, 1988; Smith & Watts, 1992; Bradley et al., 1984), which are not attributable to production factors. These productivity differentials may be partly explained by managerial ability. Managers are responsible for overseeing production, public service and project delivery across for-profit, non-profit and public sectors. Understanding the role of CEOs is therefore an important step in analyzing productivity gaps in the economy. Moreover, firms dedicate considerable amounts of time and resources to hiring and training their managers[1]. In fact, there is a debate on whether the high wage inequality observed between CEOs and the rest of employees and the recent rise in the CEO pay slice can be attributed to a more prominent role of CEO ability or to rise firm productivity/growth. The latter explanation has gained more traction in the literature[2]; however, it is possible that the match between CEO and firm, seen as complementary inputs, explains part of the rise in firm productivity.

This paper evaluates the role of the CEO type in firm productivity in the presence of complementarities between the two. There have been significant recent advances in the literature documenting a relation between managers' individual ability or firm-wide managerial practices and firm outcomes. In the absence of a natural experiment that randomizes CEO-firm assignment, many studies rely on job-to-job transitions to estimate employee quality. Based on the pioneer work of Abowd et al. (1999), job transitions provide a quasi-natural experiment that allows for the separate identification of the role of the CEO and firm in productivity. The main challenge of this literature is the non-random nature of CEOs job transitions across firms, which results in overestimation of the importance of the CEO for firm productivity. To overcome this challenge, I take advantage of a rich data set to develop a novel measure of estimated CEO quality by looking at their performance in the labor market in early career years before becoming a CEO. This allows me to significantly mitigate the impact of non-random assignment between CEO and firm. Moreover, not much is known about whether good CEOs are equally good in any firm; that is, CEO quality might present different complementarities across different firm types. In fact, the literature has thus far paid less attention to the role of complementarities between CEO and firm type in explaining firm productivity[3]. I estimate the importance of CEO-firm complementarities in determining productivity in a model that accounts for endogeneity in CEO job mobility.

I develop a simple firm production model to guide the empirical analysis. In a closed economy, firms produce revenue with the traditional inputs -labor and capital- and must hire a CEO to oversee production. Both firm and CEO are endowed with a certain of technology type (firm) and quality type (CEO), exogenously determined, which accompanies them throughout the time period under analysis. The CEO's role

[1] A survey of 610 CEOs by Harvard Business School estimates that typical mid-level managers require 6.2 months to reach their break-even point, and higher for a top-level manager.

[2] Gabaix & Landier (2008); Terviö (2008); Malmendier & Tate (2009); Bebchuk et al. (2011)

[3] There is a large body of literature studying employer-employee complementarities, Best et al. (2017); Eeckhout & Kircher (2016); Gulyas (2016) to name a few of the most recent.

is mediated in the firm through a span-of-control technology (Lucas, 1978) and CEO's unobserved quality is a labor augmenting input in the production function. I introduce a complementarity parameter which captures the interaction between a CEO and firm types. This parameter translates into different CEO contributions to firm productivity within the same CEO type. I derive testable implications regarding the impact of CEO and CEO-firm complementarities in firm production, which I study in two separate environments.

The empirical analysis is divided into four parts and hinges crucially on the quality and breadth of the data I use. The *Quadros de Pessoal* matched employer-employee survey captures a significantly large spell of each CEO's tenure in the labor market, thus allowing me to separate her non-CEO from CEO years. First, I show empirical evidence of non-random assignment of CEOs to firms. CEO mobility appears to be motivated by the search for a better CEO-firm match, after controlling for CEO and firm types. However, the pattern of mobility of employees before becoming CEOs is as good as random vis à vis the employee-firm match, in line with relevant literature (Card et al., 2013, 2015; Sorkin, 2017)[4].

Second, I evaluate the role of the CEO in firm productivity. I proceed to build a proxy measure of CEO quality (or type) that significantly reduces the bias introduced by non-random CEO assignment, by exploiting the full labor market spell of the CEO. I use the non-CEO (employee) years to estimate the (then) employee's ability as a fixed effect in the spirit of Abowd et al. (1999, 2002) by estimating a two-way fixed-effects regression on log-wages. This strategy draws on the fact that mobility pattern in non-CEO years appears as good as random. I evaluate the explanatory power of wages by employee and firm types using a variance decomposition exercise and accounting for finite sample bias (Kane & Staiger, 2008; Gaure, 2014; Best et al., 2017)[5]. I find similar decomposed effects as the relevant literature, reassuring that the use of this part of the sample does not significantly alter the attribution of heterogeneity sources. I take the standardized estimated fixed effect as a measure of CEO quality[6][7]. I evaluate the impact of the measure CEO quality in firm productivity by estimating a firm production function in which CEO quality is an additional input.

Third, I conduct heterogeneity analyses of the coefficient estimates for CEO quality, along firm characteristics (size, economic sector, ownership structure, profits, labor productivity and innovation expenditure) and CEO characteristics (schooling, tenure, experience as CEO, age). I also perform robustness checks for estimated CEO quality and production function specification.

Fourth, I estimate an extension of the firm production model that incorporates the

[4]The mobility of employees can be used as a quasi-natural experiment to the extent that there is plausible evidence that mobility is orthogonal to specific employee-firm wage realizations; that is, that mobility is orthogonal to the employee-firm match outcome after controlling for employee and firm types.

[5]See Andrews et al. (2008, 2012) for detailed descriptions of finite sample bias (also known as incidental parameter bias) in the context of fixed effects regressions.

[6]Assuming one-dimensional ability and that the selection of a candidate for a position is based on the candidate's performance in their current role (Peter et al., 1969), rather than abilities required for the intended role.

[7]From now on, the mention of "CEO quality" refers to the proxy measure of ability derived from the fixed effects estimation and is a concept used in the limited environment of this paper.

evidence of endogenous CEO mobility. This extension expands the previous analysis by explicitly including CEO-firm complementarities and a one-period Markov process on revenue realizations. I use a distributional model approach devised in Bonhomme et al. (2017b) to estimate the role of CEO and firm type, as well as CEO-firm complementarities, in productivity. The first step of this approach is the dimension reduction of firm heterogeneity into a finite number of firm classes through a k-means clustering algorithm[8] based on similarity between firm revenue distributions. With the addition of the CEO-firm complementarity parameter, dimension reduction plays an important role as it allows for a more parsimonious treatment of heterogeneity and therefore attenuation of incidental parameter bias, given that job mobility rates are higher within classes as opposed to firms. The second step of the distributional approach is to estimate a finite mixture model which assumes the impact of the CEO and firm heterogeneity types are random and a result of a mixture of Gaussian distributions that generate revenue realizations for each CEO-firm type match. Lastly, I perform a counterfactual exercise in which I artificially set CEO-firm complementarity to zero and compare the resulting productivity distribution to the actual one.

I find that a one standard deviation increase in CEO's innate quality can be translated into an increase in 5% revenue productivity, after controlling for experience, schooling and other CEO observables. This result is in line with findings in the literature that uses natural experiments [9]. Moreover, results regarding CEO and firm specific contributions to wage setting are in line with the literature [10]. I present evidence that CEO-firm match complementarities exist and are a significant part of the effect of CEO and firm heterogeneity in revenue production. I quantify the magnitude of complementarities through a counterfactual experiment in which I randomly reallocate CEOs to firms and compare the resulting distribution of wages and productivity with the real allocation.

I present heterogeneity analysis of CEO quality according to observable characteristics of firm and CEO. CEO's quality is more important in the services industry, smaller to medium sized firms, and firms where there is high average worker mobility. Simultaneously, CEO's quality is positively related with observables such as schooling and tenure in the labor market, and negatively correlated with family firm ownership.

Lastly, as a result of the finite mixture model, I find significant CEO-firm complementarities in production, which are stronger for higher ability CEOs. Complementarities account for approximately 1.5% of average revenues, with a stronger effect (3%) in the top 10th percentile of the revenue distribution.

This paper contributes to the literature of organization economics and corporate governance, which has made significant progress in determining a correlation between top-management[11], in corporate outcomes[12], either through their characteristics (Bertrand & Schoar, 2003; Bennedsen et al., 2012; Queiró, 2016), their time use (Bandiera et al., 2013), or firm ownership structure (Bennedsen et al., 2007; Pérez-

[8]Steinley (2006).

[9]Pérez-González (2006); Bennedsen et al. (2007).

[10]Bender et al. (2016).

[11]Chief Executive Officers (CEO), Chief Financial Officers (CFO) and Chief Operating Officers (COO).

[12]Profits, ROA, ROI, M&A decisions among others.

González, 2006). Alongside, there have been studies documenting the relation between managerial practices and firm (Bloom & Van Reenen, 2007; Mion et al., 2016) or public bureaucracy (Rasul & Rogger, 2013; Best et al., 2017) performance.

An active literature in personnel economics, both theoretical and empirical, has focused on establishing a source of observed increasing trajectory in CEO wages. Several papers (Terviö, 2008; Gabaix & Landier, 2008; Chade & Eeckhout, 2013; Gayle et al., 2015) point to an overwhelming importance of the role of the firm, as opposed to CEO, in pay increases. Simultaneously, other papers (Lazear et al., 2015; Bandiera et al., 2017) document the relevance of mid-level and top-level managers in firm productivity.

Three limitations of these two groups of studies are (i) the challenge of circumventing endogeneity in the assignment of managers to firms[13], (ii) the difficulty of accounting for ability heterogeneity, both at the firm and manager level, to the overall performance of the manager-firm pair and (iii) little evidence regarding the importance of CEO-firm complementarities between manager and firm[14]. This paper adds to the literature by putting forth a proxy measure approach that allows me to estimate an effect of CEO quality in firm productivity that is not plagued by endogenous mobility.

This paper also contributes to the broader wage inequality and labor economics literature, which addresses the issue of isolating worker from firm-specific effects on wages using employee-employer matched data. I add to this literature by expanding the study of match-specific complementarities to the realm of CEO-firm pairs. This literature can be loosely divided into two groups. The first group corresponds to the influential and widespread two-way fixed-effects approach, put forth by the seminal works of Abowd et al. (1999, 2002)[15]. Their work presents a tractable model that employs worker and firm fixed-effects to account for the relative importance of worker and firm heterogeneity in wage dispersion, under some (potentially) stringent assumptions[16]. Alongside, the search and match literature deals with the challenge of analyzing wage dispersion using structural models that underpin the matching process of worker and firm (Postel-Vinay & Robin, 2002; Bagger et al., 2014; Hagedorn et al., 2017). The distributional method I use, put forth by Bonhomme et al. (2017a,b), attempts to bridge the gap between structural and reduced-form approaches by estimating the role of the worker-firm pair in and importance match-specific complementarities in wage setting. Their model improves upon the computationally burdensome structural models and the lack of parsimony in fixed-effects based reduced-form strategies. I use their environment to evaluate the role CEO-firm complementarities, under endogenous CEO mobility, in firm productivity.

The remainder of the paper is organized as follows. Section 2 details the theoretical

[13]Two noteworthy papers that escape this criticism are Bloom et al. (2013) and Lazear et al. (2015).

[14]Exception made for the CEO-firm match model put forth in Bandiera et al. (2017).

[15]Abowd, Kramarz and Margolis' work has been a cornerstone in the study of wage inequality (Card et al., 2013; Song et al., 2015), gender gap in wages (Cardoso et al., 2016), bargaining and sorting (Card et al., 2015). Their method has also been adopted by other economics fields, specifically in identifying teacher versus school value-added in student performance (Jackson, 2013), or to document sources of variation in health care utilization in the U.S. (Finkelstein et al., 2016).

[16]Some recent papers adopt instrumental variable approaches to disentangle between worker and firm heterogeneity in wage setting (Jäger, 2016).

model which establishes a framework for the empirical analysis. Section 3 describes the data and provides relevant information regarding its institutional context. Section 4 documents the estimation of CEO quality and its contribution to firm productivity and section 5 presents a battery of robustness analyses. Section 6 outlines and estimates the finite mixture model. Section 7 concludes and sets avenues for future research.

2 Conceptual Framework

In this section I present a model of firm production where CEO quality is a TFP-augmenting input in production. This stylized model presents a conceptual framework for the empirical analysis developed in section 4. The model is composed of two parts, which differ on the assumptions made regarding CEO job mobility. First, I assume CEOs move based on their ability type and firms' types. On a second part, I assume CEOs decision to move also takes into account the CEO-firm match specific realizations and that CEO-firm complementarities are a relevant determinant of overall firm production.

2.1 CEO Quality and Firm Productivity with Random Assignment

Consider a closed economy with a homogenous labor force of size L and K units of homogenous capital, both supplied inelastically to the market. The two factors can be combined to achieve production that is sold in a homogenous-good, price-taking market. In order for the firm to operate, it must hire a CEO to lead the company and oversee production. Moreover, the quality of the CEO. Therefore, besides the two traditional inputs (L and K), firm's production is also affected by the quality of the CEO[17]. In the context of this model, CEO "quality" can be thought as an interaction between innate ability and human capital. In this paper, I assume "quality" is given by a fixed-effect which the CEO brings to any firm he works for[18].

Let the described economy consist of J firms, $j \in \{1, ..., J\}$, at each time-period $t \in \{1, ..., T\}$. Firms are endowed with a firm-specific total factor productivity (TFP) type (A_j) when they are active. Firm technology type is represented by a fixed distribution $\Lambda : \mathbb{R}^+ \rightarrow [\underline{A}, \overline{A}]$. Moreover, firms can hire from a pool of CEOs, $i \in \{1, ..., N\}$. As in Lucas (1978) span of control model, I assume CEO's quality is exogenously determined and there is a continuum of identities that are fully represented by a fixed distribution $\Gamma : \mathbb{R}^+ \rightarrow [\underline{\alpha}, \overline{\alpha}]$ at each time t. CEO quality is assumed to be unidimensional (Becker, 1973). I also assume that CEOs are hired according to their success as employees was (Peter et al., 1969). Managerial ability (α_i) enters the production function in two ways. First, as an input (Bender et al., 2016). Second, through a decreasing returns to scale (DRS) transformation of the production function, reflecting the limited span of control

[17]For simplicity, I borrow Lucas (1978) assumption that workers are a readily available factor of production to the CEO.

[18]An illustrative example would be to think of this quality as an identity type, something that is particular to the CEO.

of the CEO[19].

At every period, the CEO can move to a new firm. At this stage of the model, let us assume that CEO job mobility is only driven by α_i and A_j, and not by the CEO-firm match specific wage realizations. A standard Cobb-Douglas constant returns to scale (CRS) function represents the firm's production function if span of control were unlimited:

$$Y_{j,t} = A_j \alpha_i^{\mu} L_{j,t}^{\delta} K_{j,t}^{1-\delta-\mu} \tag{1}$$

where α_i and A_j correspond to the CEO quality type and firm technology, respectively. Given the limited span of control of the CEO, her production oversight and monitoring is a DRS transformation of equation (1). I assume this transformations takes the form of a natural logarithmic function $g(Y_{j,t}) = ln(Y_{j,t})$, such that:

$$ln(Y_{j,t}) = \mu ln(\alpha_i) + ln(A_j) + \delta ln(L_{j,t}) + (1 - \delta - \mu) ln(K_{j,t}) \tag{2}$$

An allocation of resources is described by $L_j(\alpha_i)$ and $K_j(\alpha_i)$, which correspond to the labor and capital allocations of firm j managed by a CEO with quality α_i. Labor and capital can be hired at equilibrium prices w and r, respectively[20]. The equilibrium allocation satisfies the equality between marginal rate of substitution and factor prices, as presented in equation (3):

$$\frac{K_{j,t}}{L_{j,t}} = \frac{\delta}{1-\delta} \frac{r}{w} \tag{3}$$

For the remainder of the paper, I focus on gross revenue as the object of the firm's maximization problem[21]. That is,

$$PY_{j,t} = (P * A_j) \alpha_i^{\mu} L_{j,t}^{\delta} K_{j,t}^{1-\delta-\mu} \tag{4}$$

where $P * A_j$ is TFPR (or revenue TFP[22]) and P is unique in the final homogenous goods market, reflecting the price-taking behavior of firms. The following proposition illustrates the hypothesized relationship between CEO ability and firm's productivity, measured in gross revenue. This hypothesis will be tested in the empirical analysis, in section 4.

Proposition 1 *A higher level of CEO quality results in higher firm gross revenues.*

$$\frac{dPY_{j,t}}{d\alpha_i} > 0 \tag{5}$$

Proof: see Appendix A.

[19]The intuition being that a CEO's management and supervisory abilities are inversely proportional to the quantity of L and K units under his control, *ceteris paribus*.

[20]I assume firm size is small enough not to change equilibrium prices.

[21]Several papers in the growth and productivity literature (Hsieh & Klenow, 2009, 2014) and in the organizational literature (Bloom & Van Reenen, 2007; Bertrand & Schoar, 2003; Bandiera et al., 2017) use revenues as outcomes.

[22]Hsieh & Klenow (2009, 2014).

2.2 CEO Quality, Firm Production and CEO-Firm Match

Consider the same closed economy with a labor force of size L and K units of homogenous capital. Let the described economy consist of the same J firms, $j \in \{1, ..., J\}$, at each time-period $t \in \{1, ..., T\}$. Firms are endowed with a firm-specific total factor productivity (TFP) type (A_j) when they are active. Moreover, firms can hire from a pool of CEOs, $i \in \{1, ..., N\}$. I assume CEO's quality is exogenous (α_i). CEO ability is assumed to be unidimensional and CEOs are assumed to be hired according to their performance in former positions. As before, firm's production is affected by managerial quality. However, in contrast with the previous section, CEO-firm specific match output influences production as a complementarity/interaction effect that goes beyond the effects in production of the CEO and firm in isolation.

As before, the CEO can move to a new firm at every period. In this extension of the model, CEO job mobility is driven not only by α_i and A_j, but also by the CEO-firm match complementarity, $\theta_{i,j}$. Moreover, CEO-firm match is another determinant in productivity, entering the firm's production function as a TFP-augmenting parameter. A logarithmic transformation of a standard Cobb-Douglas constant returns to scale (CRS) function with parameter δ represents the firm's production function with limited CEO span-of-control:

$$ln(Y_{j,t}) = \theta_{i,j} + \mu ln(\alpha_i) + ln(A_j) + \delta ln(L_{j,t}) + (1 - \delta - \mu) ln(K_{j,t}) \tag{6}$$

The following proposition illustrates the hypothesized relationship between CEO-firm match and firm productivity, measured in gross revenue. This hypothesis will be tested in the empirical analysis, in section 6.

Proposition 2 *A higher level of CEO-firm match complementarity results in higher firm gross revenues.*

$$\frac{dPY_{j,t}}{d\theta_{i,j}} > 0 \tag{7}$$

Proof: Appendix A.

3 Data & Context

I provide a brief account of relevant features regarding the context of the data used in this paper. While my analysis is based on Portuguese data, the main labor market and productive sector characteristics in Portugal indicate that results may be generalized to other EU or OECD countries. Figure 1 presents the evolution of labor market participation rates. The Portuguese participation rate has been fairly constant over the past 10 years, at approximately 74.1% of the whole population. This figure is similar the 72% estimated for EU average in 2016) and OECD average (71.7% in 2016)[23]. The

[23]Source: OECD (2017), Labour force participation rate (indicator). doi: 10.1787/8a801325-en. The US participation rate is similar (73%).

ratio of manager to non-manager employee population is estimated at 6.7% for the Portuguese labor market, a figure close to the OECD average of 6.4%[24]. Despite the similarities in labor participation, labor productivity in Portugal is significantly lower than that of the EU[25]. This gap in productivity makes a stronger case for the role of non-input related productivity differentials in general, and CEO quality in particular.

Alongside labor market features, Portuguese economic activity can be representative of other EU countries. Portugal has experienced, as most southern European countries, a severe economic downturn in the aftermath of the Great Recession followed by a slow recovery that has placed GDP growth at no more that 1-2% a year[26]. Small and medium sized firms represent 99% (95% OECD average) of the total number of firms in Portugal and have accounted for between one half and two thirds of its total value creation over the past decade[27]. Most (68.2%) of these firms' employment is dedicated to services, comparable to a 72% in the EU[28]).

I combine two data sets to generate a matched employer-employee panel. Employee information comes from *Quadros de Pessoal*, a proprietary data set collected and administered by the Portuguese Ministry of Employment, drawing on a compulsory annual employment census of firms[29] that have at least one employee on payroll during the survey reference week[30]. Firm level data is obtained from *Informação Empresarial Simplificada (IES)*, a mandatory annual survey on firm financial information. The two data sets are merged by a common firm identifier. The QP data set has been used in numerous fields of labor economics, namely in the study of gender wage gap[31] and bargaining and unions[32]. The mandatory character of both *Quadros de Pessoal* and , together with reporting based on tax-authority valid profiles[33], lends particular credibility to the data set at hand. Moreover, the QP encompasses the entirety of the Portuguese economic private sector, making its breadth reassuring in providing a safe ground on which to run meaningful empirical analyses.

3.1 Employee and CEO Data

Quadros de Pessoal is a longitudinal data set on private sector employees, spanning from 1986 to 2013[34]. As of 2013, the survey collected information on approximately 450,000 firms and 3 million employees. Reported data cover each firm (location, economic activity, employment, sales, and legal status) and each of its workers (gender, age, education, skill, occupation, tenure, managerial versus non-managerial position,

[24]Source: OECD (2014), Share of employed who are managers.

[25]Source: PORDATA and Eurostat (2016).

[26]Source: Bank of Portugal.

[27]Source: PORDATA (2017), *Empresas, Pequenas e Médias Empresas*.

[28]OECD (2015), Employment in the services sector.

[29]Public administration and informal market services are excluded. Includes private, nonprofit and public firms.

[30]One week of October of each year.

[31]Cardoso et al. (2016).

[32]Card et al. (2015); Addison et al. (2017).

[33]All employee salaries and employer sales revenues are reported as they are declared to the Portuguese Tax Authority (*Direcção Geral dos Impostos*) and Social Security.

[34]The survey has waves after 2013; however, these are the ones available at the Bank of Portugal.

hours worked, overtime, and earnings[35]). Firms and workers entering the database are assigned a unique, time-invariant identifier that allows to tracking of firms and worker pairs over time. The data covers information on all personnel working for any firms with at least one employee on payroll.

Importantly, the variable "occupation" allows me to identify the managing director or CEO[36]. For the purpose of the empirical analysis, I focus on single job holders and full-time jobs held by men and women aged between 18 and 68 years old. I perform a 98%[37] winsorization of wage outliers[38].

Table 1 presents means and standard deviations[39] for two samples: non-CEO and CEO employees. Column 1 presents the results for CEO employees. Female CEOs make up for about 30% of the sample; the average CEO is around 45 years old; about 35% of managers hold a higher degree and have been in that management position for over 5 years. There are considerable differences between CEOs and other employees, both at the earnings level and demographics aspect. Importantly, CEOs present higher job mobility, both across firms and across positions and achieve considerably higher earnings levels. Finally, column 3 presents the same statistics for the largest connected set of non-CEO employees and firms[40]. In comparison, mostly all variables exhibit similar descriptive patterns within the largest connected set and the whole sample. This will become important in section 4.2.

The data set also allows me to track employee job to job transitions. In fact, this source of variation is crucial for the identification of person versus firm effects on production. Table 2 displays executive transitions, that is, manager switches between positions and firms. In the interest of completeness, I present all switches between any combination of two out of the four management positions: CEO [41], Financial Manager, other high-level managers and mid-level managers. Other high-level managers consist of operative mangers and others who report directly to the CEO. Employee transitions encompass the job mobility of non-managerial employees. In the remainder of this paper, I focus on the CEO-CEO and employee-employee job transitions. Importantly, both Table 1 and Table 2 point to the fact that the amount of job transitions declines as the employee becomes a CEO. This illustrates part of the differential mobility patterns between CEO and employees, developed further in the next section. The large amount of job transitions, particularly within-groups, is encouraging as it provides valuable job mobility that will be exploited in the identification strategy.

[35]The information on earnings includes the base wage (gross pay for normal hours of work), seniority-indexed components of pay, other regularly paid components, overtime work pay, and irregularly paid components.

[36]Appendix B elaborates on the methodology used to identify the firm CEO.

[37]I set all variables below the 1% percentile to the 1% percentile of the distribution; the same with 99% percentile.

[38]Appendix B provides further details on sample selection criteria.

[39]Summary statistics for Table 1 are constructed using the average of cross-section estimates.

[40]The largest connected set is, within the groups of firms that are linked together by employee mobility, the one that encompasses more observations within the sample. See section 4.2 for a more detailed description of the largest connected set

[41]Highest management level within the firm organization.

3.2 Firm data

The *IES* dataset spans from 2005 to 2015 and includes financial information of the firm. The survey reports data on balance sheet and profit and loss statements. This includes data on capital, raw materials and other consumables, services used in production, salaries and employment, added value, sales, profit or loss. These data are merged with employee-level data via a common firm identifier.

Table 3 presents summary statistics on firm characteristics[42]. A share of approximately 28% of all firms is located in Lisbon, whereas 19% are in Porto. Approximately 27% of the firms operate in the manufacturing sector, 11% in construction and 45% in the service.The non-profit sector accounts for approximately 4% of the labor market, a figure that is compatible with the estimated size of the sector in the Portuguese economy[43]. The average firm has 18 employees. For the purpose of the upcoming empirical analysis, I focus on non-agriculture sector firms, exclude non-profit and banking related organizations. For more details on sample selection, see Appendix B.

4 CEO Quality and Firm Productivity

I use a reduced-form approach to the estimate of the production function model developed in the conceptual framework section. In this section, the baseline econometric model unfolds in two stages. First, I measure CEO quality as a person fixed effect in a wage regression, in the tradition of Abowd et al. (1999). Second, the estimated CEO quality is used as an input in the production function. This amounts to testing the validity of Proposition 1 of the conceptual framework.

4.1 Measuring CEO Quality

In the first stage of the reduced form approach, I estimate a model that separates the components of wage variation attributable to employee-specific and firm-specific heterogeneity. I use the two-way fixed-effects model first introduced by Abowd et al. (1999). The economy consists of $i = 1, ..., N$ employees and $j = 1, ..., J$ firms. I model the logarithm of wages as a function of employee observables, firm and employee fixed effects:

$$y_{it} = \alpha_i + \psi_{j(i,t)} + \boldsymbol{X_{it}\beta} + \varepsilon_{it} \tag{8}$$

Consider *j(i,t)* as the indicator for the firm *j* where employee *i* works at time $t.y_{it}$ stands for deflated log of wages, $A_{j(i,t)}$ represents firm *j* fixed effect, which captures constant firm specific heterogeneity, α_i are the employee fixed effects and X_{it} represents a vector of time-varying employee-level control variables[44] and year fixed effects. The parameter of interest is α_i, which I interpret as an innate ability that is valued in the labor

[42]Summary statistics for Table 3 are made up of averages of annual cross-sections, except in the case of firm longevity.

[43]"The non-profit sector (...) represents more than 4% of the Portuguese GDP in 2015." *in* João Pavão, 2015.

[44]Quadratic terms in age fully interacted with schooling levels.

market in the same way. I use α_i as the estimated employee quality[45]. I recover the estimated employee qualities for the firm CEOs that are later used as an input in the production function.

4.2 Identification and Connected Sets

Ideally for the purpose of this study, employees would be assigned to firms randomly in an experimental setting and moved randomly throughout the period under observation. This would allow for straightforward separate identification of employee and firm contributions (i.e. heterogeneity types). In a non-experimental setting such as the one in this paper, separate identification of employee and firm fixed effects can only be achieved if we observe employees working for more than one firm and firms employing more than one employee over the time-series. In other words, we need firms to be linked to one another through employees who move between them in a connected set. The stronger the link, i.e. the more frequent the employee moves, between these two firms, the more accurate the separation of the influence of the employee from that of the firm type on wage setting over time. As shown in Abowd et al. (2002), connectedness is a sufficient condition for identification. As a result, identification can be achieved within a connected set of firms[46].

I follow previous work[47] by focusing on the largest connected set of employee-firm pairs. This approach seems reasonable in this particular setting, since approximately 96% of the employee-firm pairs are captured within the largest connected set. Moreover, I find very similar summary statistics between the largest connected set and the full sample in terms of relevant descriptive statistics (Table 1).. Abowd et al. (1999, 2002) prove that, within each connected set, the employee and firm effects are identified only in relation to each other. Therefore, and also in the spirit of previous literature, I take a random firm as reference, normalizing that firm effect to zero and estimating unconditional variances in section 4.4.

4.3 Job Mobility and Causality

The quasi-experimental nature of the empirical model presented in equation (8) can provide a causal interpretation of employee and firm fixed effects on wage setting to the extent that employee job transitions are orthogonal to the error term. I write the error term associated with equation (8) in three parts, as Card et al. (2013), to highlight potential cases where the stated orthogonality condition does not hold. The error term ε_{it} is composed by a random employee-firm match effect, ($\lambda_{j(i,t)}$), a unit-root process that reflects increments in employee quality (ω_i) and a transitory shock component ($\upsilon_{j(i,t)}$) as described in equation (9):

$$\varepsilon_{it} = \lambda_{i,j(i,t)} + \omega_{it} + \upsilon_{it} \tag{9}$$

[45]A definition of manager "quality" in the context of this paper is further described in section of this paper.

[46]As a counterexample, in the case of an employee who stays in the same firm throughout the 26 years' panel time-span, the employee and firm fixed effects cannot be disentangled.

[47]Card et al. (2013); Cardoso et al. (2016).

The match effect component ($\lambda_{i,j(i,t)}$) represents wage premiums or discounts that employee i faces when matched with firm j that go beyond the channels of firm heterogeneity or employee quality. Match effects could arise if specific employees are especially suited (or unsuitable) for specific firms. Match-specific wage components are present in the search-and-match literature which models an idiosyncratic component of output associated with each possible job match[48]. I apply the same logic in the context of this model, where match effects are reflected in wages. The unit root component ω_{it} reflects potential drift in employee quality that has lasting effects. This component encompasses a wide array of shocks with permanent effects to the employee's ability, such as health shocks or unobserved human capital accumulation. For the time being, I assume that ω_{it} has mean zero for each employee in their observed time period. Finally, there is a transitory term (υ_{it}) that represents any other temporary shock that affects the outcome, which is also assumed to have mean zero for every employee in their observed time period.

To achieve causal identification, OLS assumptions regarding the previously described error terms must hold. Let $\boldsymbol{y}$ denote the stacked $NT \times 1$ vector of year-sorted employee wages (where $NT = N$), $\boldsymbol{E}$ denote the $NT \times N$ design matrix of employee indicators, $\boldsymbol{F}$ is the $NT \times J$ design matrix of firm indicators and $\boldsymbol{X}$ is a $NT \times J$ matrix of time-varying employee covariates and ε denotes the error term. Equation (8) can be written in matrix notation as:

$$\boldsymbol{y} = \boldsymbol{C\alpha} + \boldsymbol{F\psi} + \boldsymbol{X\beta} + \boldsymbol{\varepsilon} \tag{10}$$

Consistent estimation of equation (10) via OLS implies the following assumptions regarding the interaction of the error terms with explanatory variables:

$$E[e^{i\prime}\varepsilon] = 0, \forall i \qquad E[f^{j\prime}\varepsilon] = 0, \forall j \qquad E[x^{k\prime}\varepsilon] = 0, \forall k \tag{11}$$

Equation (11) describes OLS orthogonality conditions between the error term ε and each regressor. Whereas the assumption on the vector of regressors x^k is standard, the same cannot be said about the identifying assumption regarding the employee and firm indicators, e^i and f^j. These assumptions allow for an array of sorting patterns between firm and employee[49]. However, the same assumptions preclude the existence of sorting on match-specific effects, i.e. sorting on premia/discounts earned from a specific individual employee-firm pair[50]. If this type of endogenous mobility is present in the CEO data, causal identification of equation (10) is threatened since it would imply a positive correlation between a component ($\lambda_{i,j(i,t)}$) of the error term and f^j. In fact, given the assumptions made on the error term components, causal identification of equation (10) boils down to the verification of the assumption $E[f^{j\prime}\varepsilon] = 0, \forall j$[51].

[48] Mortensen & Pissarides (1994); Shimer & Smith (2000); Postel-Vinay & Robin (2002); Eeckhout & Kircher (2017); Hagedorn et al. (2017).

[49] There can be systematic sorting of effective firms with effective employees (up to a pre-determined measure of effectiveness) that does not break the assumptions of equation (11).

[50] Intuitively, the phenomenon of (match-specific) endogenous mobility can be thought of as the pursuit of the perfect match between a CEO and a firm or employee and firm.

[51] Proof: Appendix A.

Appendix C provides intuition for on simple example of two CEOs, two firms and two time periods. I now discuss three cases that would lead to biased estimates of fixed effects.

Consider an employee who moves from firm 1 to firm 2, from *t-1* to *t*. Her expected change in wages can be summarized by:

$$
\begin{aligned}
E[y_{it} - y_{it-1}|j(i,t) = 2, j(i,t-1) = 1] = \psi_2 - \psi_1 + E[\varepsilon_{it}|j(i,t) = 2] - E[\varepsilon_{it}|j(i,t-1) = 1] \\
\psi_2 - \psi_1 + E[\lambda_{i2} - \lambda_{i1}|j(i,t) = 2, j(i,t-1) = 1] \\
+E[\omega_{it}|j(i,t) = 2] - E[\omega_{it}|j(i,t-1) = 1] + E[\upsilon_{it}|j(i,t) = 2] - E[\upsilon_{it}|j(i,t-1) = 1]
\end{aligned}
\tag{12}
$$

In the absence of bias, the expected wage differential for employee i is $\psi_2 - \psi_1$. The presence of sorting based on the employee-firm match component of wages, i.e. endogenous mobility, will result in biased OLS estimates. If this type of sorting occurs, we should observe $E[\lambda_{i2}] \neq E[\lambda_{i1}]$. That is, on average we should observe that the wage premium for an employee who moves from firm 1 to firm 2 is significantly different from the premium faced by an employee who moves in the opposite direction. In order to assess the possibility of endogenous mobility I use event studies as used by Card et al. (2013)[52]. I define an event as any job transition of an employee from one firm to another in consecutive time periods *t-1* and *t*, provided that the employee stays with both firms for at least two years. I classify jobs at origin and destination firms[53] according to the respective quartile of coworker wage distributions. I assign each job transition event to one of 16 cells of origin and destination quartiles of coworker wages. I calculate mean wages in the years before and after the event for each quartile cell.

Panel A of Figure 2 exhibits the employee event study graph. The plot depicts the job transition event timeline against average log wages for each trajectory of quartile mean coworker wages. Log wages are residualized of year fixed effects. We can observe that wage differentials between switching employee trajectories appear symmetric. Note, in particular, the trajectories from a first quartile to fourth quartile firm change and vice versa: the average wage change has opposite sign but a similar order of magnitude. Indeed, the ratio between the average wage gain in the first trajectory (first to fourth quartile) and the average wage loss in the second trajectory (fourth to first quartile) is approximately 1, reinforcing the observed symmetry[54]. This conclusion goes in line with similar findings in the labor literature[55]. Moreover, the ratio between average estimated gains and losses (the ratio between slopes of symmetric movements) is very close to 1.

Panel B of Figure 2 shows a different mobility pattern for CEOs. The same event study exercise applied to the pool of CEOs instead of employees, yields an asymmet-

[52] The event study methodology is also used in Card et al. (2015); Finkelstein et al. (2016); Best et al. (2017).

[53] Define firm where employee works in *t-1* as the "origin" firm. Define firm where employee works in to *t* as the "destination" firm.

[54] The ratio is calculated between the two expected wage differentials from *t-1* to *t+1*.

[55] Card et al. (2013, 2015).

ric plot. Focusing on the switching trajectories between first and fourth quartiles[56], the increase in wages resulting from the movement from the first to fourth quartile is significantly greater than the other way around. CEOs appear to move to a new job because systematically due to specific gains in wages from that CEO-firm pair. Moreover, the ratio between average estimated gains and losses is statistically different from 1.

The unit root component presents another source of bias if we observe $E[\omega_{it}|j = 2 \neq E[\omega_{it}|j = 1]$ in equation (12). In that case, a positive (or negative) drift in employees' quality[57] would result in systematic changes towards better (or worse) firms. This would be translated into an observable time trend in mean wages in Figure 2. This pattern is not found for employees (Panel A). However, Panel B shows that CEOs display a slightly increasing wage trend, possibly indicating that cumulative experience is increasingly valued in a CEOs career.

It is possible that a transitory wage shock is correlated with employee job mobility (e.g. plant closures). This would lead us to overstate the difference in employee effects since $E[\upsilon_{it}|j = 2] \neq E[\upsilon_{it}|j = 1]$ in equation (12) and would translate into an Ashenfelter's dip[58] in wages before a job transition. No such dip in wages is observed in wither panels of Figure 2.

4.4 Estimation and Variance Decomposition

The event studies indicate there are two different patterns of employee job-to-job transition. During the non-management years, it appears that the Portuguese data validates the literature's result that points to a match-specific exogenous mobility pattern. Later, in the years as a CEO, job transitions seem to be more oriented towards incorporating CEO-firm match gains. The differential mobility patterns throughout a CEO career provide a case for the use of the non-managerial labor market spell of the employee to estimate a proxy measure of her quality as CEO.

Focusing on the first stage of the CEOs career provides three important advantages. First, I avoid the endogenous mobility bias arising from job transitions because of CEO-firm specific wage realizations. Second, given the time separation between the years as an employee and years as a CEO, I ensure that the proxy CEO quality measure is, by construction, exogenous with respect to firm productivity in the CEO years. Third, this approach is compatible section 45 model's assumptions that CEO quality is unidimensional (Becker, 1973) and that selection of a candidate for the position of CEO is based on the candidate's performance in their previous job position (Peter et al., 1969).

I estimate equation (8) on the largest connected set of non-managerial spells for all CEOs for which a large enough employee spell is available in the data set. To ensure employee spells are comparable and not confounded by possible endogeneity in timing of ascent to managerial positions, I define the employee labor spell up until a maximum age[59]. I cluster standard errors at the employee-firm level, accounting for

[56]The same patterns can be found for other trajectories; see Appendix E.
[57]e.g. Increases in human capital accumulation.
[58]Ashenfelter (1978).
[59]Maximum age is defined as the 75th percentile of age when the employee first became a CEO, 38 years old. Robustness checks find no significant difference in using 90th percentile.

the two-way fixed effects nature of the regression in equation (8). I decompose the variance and covariance components log wages as:

$$\begin{aligned} Var(y_{it}) = Var(\alpha_i) + Var(\psi_{j(i,t)}) + Var(X\prime_{it}\beta) + 2 * Cov(\alpha_i, \psi_{j(i,t)}) \\ +2 * Cov(\psi_{j(i,t)} X\prime_{it}\beta) + 2 * Cov(\alpha_i, x\prime_{it}\beta) + Var(\varepsilon_{it}) \end{aligned} \tag{13}$$

where $\frac{Var(\alpha_i)}{Var(y_{it})}$ and $\frac{Var(\psi_{j(i,t)})}{Var(y_{it})}$ represent the percentage of wage variation that is explained by employee quality and firm heterogeneity, respectively. Results can be found in Table 4. The baseline variance computation is presented in column 1 and 2. Columns 3 and 4. CEO proxy measure of quality, the employee fixed effect, accounts for approximately 60% of employee wage variation, whereas firm heterogeneity accounts for about 20%[60]. I present "shrinkage" estimators of variance components as in Kane & Staiger (2008), to account for overestimation of employee and firm fixed effects resulting from finite sample bias[61].

The first stage thus results in estimation of proxy measures of CEO quality by focusing on the employee fixed effects before she became a CEO.

4.5 Estimating Firm Productivity

In the second stage of the reduced-form estimation I use the CEO quality measure obtained in the first stage to evaluate the role of CEO in firm productivity. Consider the following baseline Cobb-Douglas production function specification[62] for firm j at time t:

$$q_{jt} = \delta + \boldsymbol{W_{jt}\beta} + \boldsymbol{Z_{jt}\gamma} + \epsilon_{jt} \tag{14}$$

where q_{jt} is the log of deflated sales, W_{jt} is a vector of variable inputs and X_{jt} is a vector of state variables, all in logarithm form. The proxy measure for CEO quality ($CEO_{i(j,t)}$ for CEO i who works for firm j at time t) enters the production function as a state variable[63], as suggested in section 45. The sequence A_{jt} is unobserved firm productivity. The error term ϵ_{jt} has the following structure:

$$\epsilon_{jt} = A_{jt} + \eta_{jt} \tag{15}$$

where A_{jt} is a transmitted firm productivity parameter (persistent in time) and η_{jt} is an iid transitory shock.

[60]These results go in line with the labor economics literature: Card et al. (2013); Bonhomme et al. (2017b).

[61]Finite sample bias is also referred in the literature as incidental parameter or limited mobility bias. See Andrews et al. (2008) for a detailed description of this type of bias, commonly associated to panel data estimation. I present a detailed description of the variance shrinkage method used in this section in Appendix C.

[62]Bloom & Van Reenen (2007).

[63]I include CEO quality as a state variable since it works as a stock of human capital that influences the productivity process. I borrow this insight from the human capital accumulation literature (Black & Lynch, 1996; Galor & Moav, 2004).

I use the production function estimation method proposed by Wooldridge (2009). I expand the production function to include CEO quality as a state variable. This method combines Olley & Pakes (1996) (OP) and Levinsohn & Petrin (2003) (LP) approaches[64] in their treatment of simultaneity bias[65]. Both works resort to proxy variables (investment and materials, respectively) to measure firm productivity in two step estimation approaches[66].

Following recent literature, I use a GMM approach rather than two step procedure to jointly estimate both firm productivity and input coefficient[67] in equation (14). This approach assumes that productivity A_{jt} is described by a function $g(x_{jt}, m_{jt})$ of state variables and a set of instruments m_{jt}. I use real value of intermediate materials and services used as proxy variables for non-observed firm productivity, as Petrin & Sivadasan (2013), to avoid the problem of lumpy investment associated with the OP investment proxy[68]. The novelty in this section is that I include the CEO quality as a new input parameter of firm productivity as described in section 2. I measure CEO quality as the standardized person fixed effect estimated in the first stage of the estimation for all CEOs, i.e. $\hat{\alpha_i}$[69]. I focus on the years of CEO activity for all CEOs for whom the first stage $\hat{\alpha_i}$ was estimated. Given the rigid nature of most labor contracts in Portugal, I consider labor units and real capital stock as state variables.

GMM model is estimated by imposing two moment conditions on the data. The function $g(x_{jt-1}, m_{jt-1})$, which approximates firm productivity A_{jt}, is estimated non-parametrically by approximating a third-degree polinomial on both x_{jt-1} and m_{jt-1}. See Appendix C for further details regarding the GMM estimation and the measurement of capital.

Results can be found in Table 5. Estimation accounts for sector heterogeneity: services and manufacturing. Standard errors are clustered at the firm level. Results indicate that a one standard deviation increase in CEO quality translates into an approximately 5% increase in firm productivity in the services sector and 4% in manufacturing. I conduct the Sargan-Hansen[70] overidentification test to assess the joint validity of productivity proxy measures. The p-values of the overidentification test are reported in columns 3 and 6 of Table 5. In none of the cases can the joint validity of the instru-

[64] Wooldridge (2009) methodology also takes into account the critique in Ackerberg et al. (2006).

[65] Since the pioneer work of Marschak & Andrews (1944), economists have discussed potential correlation between input levels and the unobserved firm-specific productivity shocks (e.g. new technology) in the estimation of production function parameters. The intuition behind this problem is that firms that have a large positive productivity shock may respond using more or better inputs. If this concern is verified, using OLS to estimate production functions would yield biased parameter estimates.

[66] In a first step, authors employ semi-parametric methods to estimate the coefficients on the variable inputs. In a second step, the parameters on capital inputs can be identified under assumptions on the dynamics of the productivity process.

[67] Wooldridge (2009) justifies using GMM for three reasons: (i) avoid the potential problem with identification of variable inputs of the parameters in the LP first stage estimation, (ii) efficiently use the moment conditions implied by the OP and LP assumptions in one step and (iii) directly estimate robust standard errors.

[68] LP show that investment has considerable adjustment costs and therefore is not immediately responsive to productivity shocks. In fact, they argue that in most data sets, a lot of firms will exhibit zero investment in many years for this reason.

[69] The standardized person fixed effects are given by $\frac{\hat{\alpha_i}-\mu_\alpha}{\sigma_\alpha}$ and $\mu_\alpha = \frac{\sum_{CEO} \hat{\alpha}}{n_{CEO}}$

[70] Sargan (1958) and Hansen (1982).

ments be rejected at the 1% level. CEO quality, estimated in section 4.1, is therefore an important parameter in firm productivity while controlling for other inputs and simultaneity bias.

4.6 CEO Quality and Observables

Having established the importance of CEO quality in firm productivity, I turn to answer a second question: do higher-quality CEOs are/behave differently? In other words, are there observable characteristics at the CEO and firm that are positively correlated with CEO quality? I use CEO quality, estimated as employee fixed effects, to compute correlations with CEO and firm observables. I run two types of correlation tests: pairwise regressions of CEO quality on each of the observables of CEO and firm separately, and a post-LASSO regularization regression which performs variable selection and coefficient regularization. The regularization parameter is set to minimize the cross-validation (Tibshirani, 1996).

Figure 3 presents the results. Panel A exhibits the pairwise coefficients of a regression where each variable presented is the only regressor and CEO quality estimate is the outcome variable. All variables are standardized to have unit standard deviation. Panel B presents the results of the LASSO regularization procedure.

We can conclude from Figure 3 that CEO quality is closely associated with several observable variables. First, better firm performance indicators, such as profits, operating revenue and employee value added are associated with higher quality CEOs. These results go in line with the production function estimates in the previous section. Second, higher quality CEOs show a strongly positive correlation with investment in innovation (Research and Development). This result is consistent with management literature that suggests that more experienced, confident and better able CEOs innovate more (Barker III & Mueller (2002); Hirshleifer et al. (2012)). Third, CEO innate quality is positively associated with higher education, age and experience, both in the firm and as a CEO. These results are also in line with management literature, as well as Bertrand & Schoar (2003). Fourth, family owned and managed firms are less likely to employ a higher quality CEO, which is consistent with the literature of family firms[71].

5 Robustness Analysis

In this section I present the results of two sets of robustness checks for the reduced form analysis presented in section 4. First, I develop a battery of checks to ensure that employee fixed effects before becoming a CEO is a plausible measurement CEO quality. Second, I run alternative production function specifications to validate the results obtained on the role of CEO quality in firm productivity.

5.1 CEO Quality Measurement

In the previous section I use non-managerial employee fixed effects to proxy CEO quality. I then use the estimated CEO quality as a productivity augmenting parameter

[71] Pérez-González (2006); Bennedsen et al. (2007); Bandiera et al. (2013).

in the production function estimation and find that a one standard deviation increase in CEO quality translates into an increase in sales revenues of approximately 5% for the service sector.

One possible concern with this finding is that it may be capturing variation in ability of other firm employees rather than the CEO. Keeping the first stage estimation equal, I run a placebo regression which randomly picks a firm employee to replace the CEO parameter in the production function, the second stage of the estimation. I use this randomly chosen employee fixed effect. Table 6 presents the results. The random employee quality measure is not statistically significant as a productivity input in the firm.

I run a separate estimation in which, rather than focusing on the years before the employee becomes a CEO, I use the whole labor market trajectory of the CEO to estimate his fixed effects measure of quality in the first stage of the estimation. According to the findings provided by the event study graphs in Figure 2, this means including a significant portion of the CEO's trajectory which appears to present endogenous job mobility. If that is the case, CEO fixed effects should be overestimated[72] and, consequently, so will the coefficient of CEO quality in the second stage regression. Results can be found in Table 7. We can observe, as expected, that the estimated role of CEO quality on production is considerably higher than when using a proxy measure. Results indicate that a one standard deviation increase in CEO quality translates into 10%, compared to 5% when using a proxy measure that has been significantly clean of endogeneity.

My findings are in line with the results in two separate strands of literature. When compared with papers that use variation coming from a natural experiment, my CEO quality proxy measure estimation yields very similar results both in size and magnitude. As an example, Bennedsen et al. (2007) find a 6% increase in productivity due to a high-quality CEO. This is comparable to the 5% result I get when using a proxy measure of CEO quality.

Another set of papers (Bender et al., 2016) use fixed effects models to estimate CEO quality, using the whole spell of CEOs in the labor market. They estimate that around 13% of the revenue variation can be attributed to the CEO, a figure comparable to the results in Table 7.

5.2 Production Functions

I estimate alternative specifications for the production function[73]. I use an OLS estimation of equation (14) with firm$\times$year fixed effects. On a separate estimation, I relax two important Cobb-Douglas assumptions. The second order translog specification allows for output elasticities to change over time and for input substitutability to be

[72]In the presence of CEO job mobility based on CEO-firm match, part of the CEO-firm specific effects on wage realizations are attributed to the CEO.

[73]Petrin & Sivadasan (2013).

different from 1:

$$q_{jt} = \sum_{k=1}^{5} \beta_k X_{jt}^k + \beta_{kk} X_{jt}^{k^2} + \sum_{l \neq k} \sum_{k} \beta_{lk} X_{jt}^k X_{jt}^l + \epsilon_{jt} \tag{16}$$

where q_{jt} represents deflated log of sales for firm j in year t, X_{jt} stands for one of the five input variables[74]. As in the OLS specification, I use the same with firm×year fixed effects.

In Table 8, I present the results of both specifications. The Wooldridge (2009) and translog methods generate similar predictions regarding the role of CEO quality in firm productivity. OLS performs a relatively worse in estimating input elasticities.

6 CEO-Firm Complementarities

In this section, I use a different approach to estimate the role of CEO-firm complementarities in firm productivity. After establishing the important role of the CEO in firm productivity and observing CEO endogenous mobility, the question now turns to the role of CEO-firm complementarities in explaining part of the productivity differentials that are generally attributed to firm heterogeneity (Terviö, 2008; Gabaix & Landier, 2008; Gayle et al., 2015). This section focuses on estimating CEO-firm complementarities by expressing them in terms of productivity differentials for the whole economy.

There are two challenges in the estimation of a two-sided heterogeneity model in the presence of complementarities. First, CEO job mobility shows signs of endogeneity; that is, CEOs change jobs according to CEO-firm match specific wage realizations, either past or expected in the future, besides CEO and firm types. This will affect the estimation of complementarities. Second, accounting for complementarities demands a more flexible model than the Abowd et al. (1999) setting and for the existence of enough CEO movements across firms that allows for separate identification of CEO and firm effects and unrestricted interaction between CEO and firm[75]. Both these challenges are addressed in the finite mixture model developed by Bonhomme et al. (2017b) (BLM for short) with two-sided heterogeneity (CEO and firm) in wage determination. The novelty in this section is that I apply their setting to a revenue productivity estimation problem based on CEO-firm matches. The dynamic model I use includes a one-period Markov process for job mobility as well as revenue path dependence and allows for an unrestricted form of complementarity between the firm and CEO. Their framework preserves parsimony by focusing on discrete heterogeneity classes, estimated through a dimension reduction technique to model firm heterogeneity.

I use this approach to estimate latent CEO types and firm classes, CEO-type compositions of each firm class and mobility probabilities between firm classes. These

[74]Labor, capital, materials, services and CEO quality.

[75]Introducing a new parameter in a fixed effects regression is not an option, since the incidental parameter bias would be further exacerbated in this setting given the limited amount of observations of each CEO-firm match.

parameters allow me to characterize firm productivity distribution by CEO, firm and match heterogeneity and also permits the use of counterfactual exercises to establish the effect in firm productivity when match complementarities are artificially broken by randomly reassigning CEOs to firms.

6.1 Mixture Model Assumptions

Consider an economy with *N* CEOs and *J* firms; j_{it} is the identifier of the firm *j* where CEO *i* works at time *t*. Job mobility is denoted by m_{it}, which is equal to 1 if the CEO switches firms from time *t* to time *t+1*. Firm heterogeneity, instead of the fixed effects format as before, is now characterized by firm class *k*. The support of firm class is discrete and finite, $k_{it} = k(j_{it})$ and $k_{it} \in \{1, ...K\}$. CEO heterogeneity is also discrete and finite: αi represents the latent type of the CEO which will be represented as a random effect. There is a stream of firm revenue realizations, Y_{jt} from $t = (0, ..., T)$.

I focus on a dynamic model[76] in which both job mobility and employee's earnings exhibit a specific type of serial correlation. The dynamic model has 4 periods: CEO moves from firm *k* to *k'* between *t=2* and *t=3*. Employee stays in the same *k* firm class between *t=1* and *t=2*, and then again in firm class *k'* between *t=3* and *t=4*.

Assumptions *(dynamic model)*

1. Job mobility depends on CEO type α and firm class *k* and *k'*, but also on match specific revenue realizations w_{i2}. However, it cannot depend on w_{i1}.

2. Revenues $w_{i,t+1}$ depend on the former period revenues realization, w_{it} but not on $w_{i,t-1}$.

The assumption detailed above (formally presented in Appendix D) represents two first-order Markov conditions on job mobility and wages. To illustrate the dynamic model in an interactive setting that is comparable with the model used in section 4, consider:

$$y_{it} = \rho_t y_{i,t-1} + a_{1t}(k_{it}) + a_{2t}(k_{it}) + b_t(k_{it})\alpha_i + \boldsymbol{X_{it}c_t} + \varepsilon_{it} \tag{17}$$

where ρ_t is the persistence parameter on one-period revenues resulting from the Markov process assumption, $b_{it}(k_{it})$ is the complementarity effect between CEO and firm.

6.2 Identification

Identification in this model relies, as before, on job mobility. BLM claim that the key condition for identification is to fully exploit revenue information before and after a job move. In a dynamic setting with complementarities, graph connectedness used in section 4 is not enough to achieve identification, as we are aiming to estimate CEO-firm complementarities. To do so, we need sufficient variation in the latent types of

[76]BLM discuss a static and a dynamic version of the model. Given that the static model is structurally equivalent to the fixed effects approach used in section 4, I focus on BLM's dynamic model as it provides the most complete and novel set up.

job movers between different firm classesIn particular, complementarities would not be identified if there were random assignment of CEOs to firms..

Identification follows the same steps as the estimation. First, a dimension reduction *k-means* algorithm is used to classify firms into a finite number of clusters according to firm distribution of log of revenues. A formal discussion of identification of grouped fixed effects is presented in Bonhomme & Manresa (2015). Second, in the dynamic model, 4 periods are needed for identification, in which only one movement is contemplated (between *t=2* and *t=3*). Maximum likelihood estimation is used to estimate density of log of revenues distribution, latent CEO types α_i with $i = \{1, ..., L\}$ and transition probabilities $p_{y_2 y_3 k k'}(\alpha)$ for job movers. After having estimated those parameters, job stayers and movers are used to estimate the type proportions within each firm class, $q_k(\alpha)$. Identification of this model is fully discussed in Bonhomme et al. (2017a) and Bonhomme et al. (2017b). An illustrative example is explored in Appendix D.

6.3 Classification

Throughout this model, unobserved firm heterogeneity is assumed to have a finite support. Pining down firm heterogeneity into clusters boils down to a machine learning classification problem. BLM propose clustering the *J* firms in the sample into classes of log earnings distribution by solving the following weighted k-means[77] problem:

(18)

where n_j is the number of CEO[78] in firm *j*, $\hat{F}_j(y)$ is the empirical cdf of log of revenues for firm *j* and $H_{k(j)}(y)$ is the cdf of log of revenues of each partition *k*. The minimization problem is carried out with respect to all possible partitions of the firm data into *K* classes. I keep classes fixed across the 4 estimation periods.

Table BB presents the descriptive statistics at the firm class level. As in BLM, I use *K=10*. Notice that average salary and average gross revenue are increasing in firm class. Other firm-level observables, such as size and sector composition, also display significant differences across firm classes.

6.4 Estimation and Results

In the first stage of the estimation, I reduced firm heterogeneity dimensionality to 10 comparable firm classes. This step allows for a consistent estimation of latent firm heterogeneity while ensuring a higher mobility rate of employees across firm types that

[77] The *k-means* algorithm belongs to a class of unsupervised learning algorithms. Unsupervised learning (UL) is indicated when the econometrician does not have prior knowledge on which classes to attribute each observation. UL poses the added challenge of estimation the number of points in the support, or number of clusters. There is a large literature attempting the complicated task of estimating the number *K*. I abstract from this estimation, as BLM, assume this number is known.

[78] I stack CEO-firm spells on top of one another so that time period $t = 0$ is the event year for all changes; time dimension is taken out by detrending revenues.

will be essential for identification in the second stage.

While BLM develop estimations for both the linear and finite mixture models, I restrict attention to the latter. This model provides a non-parametric approach to the maximum likelihood construction thus generalizing the approach to a wide array of specifications. A finite mixture model is a convex combination of finite number probability distributions, used for representing the presence of subpopulations within an overall population. In the context of this model, a finite mixture represents a probability distribution of an CEO belonging to any of the *L* latent types.

Given the assumptions described in section 6.1, I estimate densities of log-revenues and transition probabilities using job movers only using the following log-likelihood function:

$$\sum_{i=1}^{N^m}\sum_{k=1}^{K}\sum_{k'=1}^{K}\mathbf{1}\{\hat{k}_{i2}=k\}\mathbf{1}\{\hat{k}_{i3}=k\}\times ln(\sum_{\alpha=1}^{L}p_{kk'}(\alpha)f^{f}_{y_2,k\alpha}(Y_{i1})f^{m}_{kk'\alpha}(Y_{i2},Y_{i3})f^{f}_{y_3,k'\alpha}(Y_{i4})) \tag{19}$$

The log-likelihood derives from the Markov process assumptions. Regarding f^f and f^b, we know that Y_{i4} is independent of past mobility and firm classes conditional on $Y_{i3}, k_{i4}=k_{i3}=k', m_{i3}=0$. Y_{i1} is independent of future mobility and firm classes conditional on $Y_{i2}, k_{i1}=k_{i2}=k, m_{i1}=0$. As for f^m, we know that firm revenues for job movers depend on the first lag of log of revenues, therefore the densities for movers are bivariate. Estimation of equation (19) recovers log-revenues densities for each CEO type α. Moreover, I can pin down the transition probabilities between *k* and *k′*. These two sets of parameters allow me to characterize the revenues distribution at each period and the CEO-type distribution of *k* to *k′* for job movers. Therefore, the missing parameters to get a full picture of revenues dynamics are the type distributions for job stayers at the origin; that is, before the job move is realized (*t*=2):

$$\sum_{k=1}^{K}\sum_{k'=1}^{K}\mathbf{1}\{\hat{k}_{i2}=k\}\times ln(\sum_{\alpha=1}^{L}q_k(\alpha)f^{f}_{y_2,k\alpha}(Y_{i1})f^{s}_{kk'\alpha}(Y_{i2},Y_{i3})f^{f}_{y_3,k'\alpha}(Y_{i4})) \tag{20}$$

Both equations (19) and (20) are single-agent correlated random-effects log-likelihood functions. Though identification of *k* and α is non-parametric in this model, estimation of densities needs a distributional assumption. I use a finite mixture of Gaussian distributions. Results can be viewed in Figure BL. The left panel represents the CEO type distributions across firm classes (x-axis). There are clear patterns of sorting, with higher type employees sorting into higher class (higher revenue) firms. The right panel represents complementarity patterns. While complementarities seem timid for most CEO types (almost no change in expected log-revenues for a given α type), note that there appears to be a significant complementarity pattern between higher type employees, where in CEOs would fit. I further explore these results by running a counterfactual exercise in the next section.

6.5 Counterfactual Exercises

I run a counterfactual exercise to explore the role of complementarities in firm productivity. First, I randomly reassign CEOs to firms. I then simulate the distribution of firm production assuming that the log-revenues distribution conditional on CEO type and firm class are not affected by the reassignment[79]. Note that, if CEO and firm allocation is random, then the term $b_t(k_{it})\alpha_i$ in equation (17) is zero (no complementarities). In essence, CEO and firm random assignment is an artificial way to set complementarities to zero and therefore evaluate the role of complementarities by computing the difference in mean productivity and other moments:

$$E[Y_i] - E^{cf}[Y_i] = E[b(k_i)\alpha_i] = cov(b(k_i), E[\alpha|k_i]) \tag{21}$$

where $E^{cf}[Y_i]$ stands for the expected log-revenues in the counterfactual environment. If complementarities $b(k_i)$ are correlated with the type distribution with firm classes, equation (21) is positive and therefore there is a relationship between CEO type and complementarities that will not be negligible in the data simulations. Results for this exercise can be found in Table BLM. The expected change in average productivity is -1.5%; that is, on average, complementarities increase productivity in about 1.5%. However, the difference in the top 10th percentile of the distribution (90th percentile) suffers a larger change in productivity on account of artificially eliminating the complementarities: around 3% of productivity is attributable to CEO-firm complementarities.

6.6 Discussion

I use Bonhomme et al. (2017a,b) framwork to anaylize CEO-firm complementarities in firm production. Other important models, such as Hagedorn et al. (2017) or Abowd et al. (2017), use structural (former) or bayesian (latter) approaches to explain CEO-firm match complementarities and sorting. I find in the BLM model offers significant advantages in my data setting. First, it is a flexible model as it allows for a non-parametric estimation of heterogeneity types as well as unrestricted complementarities between CEO and firm. Second, it provides an easily generalizable model to a variety of settings. Third, it fits the matched employer-employee/CEO setting very well and is thus replicatable in other matched panels, which are becoming increasingly available to econometricians. Fourth, it takes advantage of the whole information on revenue realizations and CEO mobility, without the need to rely on large panel data sets.

This model is not without some limitations. While it is true that a purely fixed effects model is not parsimonious and opens estimation to a number of challenges, not least incidental parameter bias, the use of random effects introduces a potential error of specification by restricting heterogeneity. In the case of this model, random effects leaves way for unrestricted CEO-firm complementarities, but restricts heterogeneity to a small finite support. On a related issue, the number of points in the support of both firm and CEO heterogeneity is a difficult issue that has received much attention

[79]This counterfactual exercise abstracts from equilibrium conditions.

in the literature (Kasahara & Shimotsu, 2014), but for which there is no consensus of easy solution. Finally, while the clustering algorithm allows for an ingenious treatment of heterogeneity through dimension reduction, it relies on a potentially strong assumption that one can perfectly separate firm observations. Overall, I believe this model has significant traction in the data setting at hand and certainly points to plausible conclusions regarding CEO-firm complementarities in production, maitaining a fair amount of degrees of freedom in the model specification.

7 Conclusions

In this paper I present evidence that the CEOs are important both for the overall determination of variation in firm productivity across firms and for the within firm productivity variation. A one standard deviation increase in manager ability results in an average 5% in the firm's gross revenue productivity. The relevance of quality of CEOs goes beyond the observable human capital, but is connected to the variables that contribute to human capital, such as schooling and labor market experience. Alongside, higher quality CEOs are more likely to invest in innovation and less likely to work in a family firm. Finally, I show evidence that match-specific complementarities are significant in determining CEO job-to-job mobility and firm productivity. In fact, complementarities explain between 1.5% and 3% of productivity differentials.

These findings add meaningful implications both to the literature in organizational and personnel economics and the wage inequality literature. First, the results point to an important role of the CEO in firm productivity. This role presents heterogeneity along firm and CEO characteristics. Second, these findings suggest that not addressing CEO-firm match complementarities may hide the full picture the impact of an individual CEO ability in firm performance.

Knowledge of the impact of CEO quality for firm performance has important policy implications. First, from and equity point of view, it contributes to shed more light on the debate regarding the size and recent increase in wage inequality[80] and the size and increase in CEO pay[81]. Second, from an efficiency point of view, understanding how the match between CEO and firm matters for productivity, besides their isolated contributions, can help devise firm policies that foster match efficiency.

This paper presents significant evidence that an important part of CEO's influence within the firm is mediated through their ability and how complementary this ability is with respect to firm heterogeneity. This means that improving establishment productivity must take into account the role of the CEO and the observable variables that are connected with this innate ability and match-complementarity.

[80] See Card et al. (2013); Song et al. (2015).

[81] See Murphy & Zabojnik (2004); Gabaix & Landier (2008); Terviö (2008); Frydman & Saks (2010).

References

Abowd, J. M., Creecy, R. H., Kramarz, F., et al. (2002). *Computing Person and Firms Effects Using Linked Longitudinal Employer-employee Data* (Tech. Rep.). Center for Economic Studies, US Census Bureau.

Abowd, J. M., Kramarz, F., & Margolis, D. N. (1999). High Wage Workers and High Wage Firms. *Econometrica*, *67*(2), 251–333.

Abowd, J. M., McKinney, K. L., & Schmutte, I. M. (2017). Modeling Endogenous Mobility in Earnings Determination. *Journal of Business & Economic Statistics*(just-accepted).

Ackerberg, D., Caves, K., & Frazer, G. (2006). Structural identification of production functions.

Addison, J. T., Portugal, P., & Vilares, H. (2017). Unions and Collective Bargaining in the Wake of the Great Recession: Evidence from Portugal. *British Journal of Industrial Relations*, *55*(3), 551–576.

Andrews, M. J., Gill, L., Schank, T., & Upward, R. (2008). High Wage Workers and Low Wage Firms: Negative Assortative Matching or Limited Mobility Bias? *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, *171*(3), 673–697.

Andrews, M. J., Gill, L., Schank, T., & Upward, R. (2012). High Wage Workers Match with High Wage Firms: Clear Evidence of the Effects of Limited Mobility Bias. *Economics Letters*, *117*(3), 824–827.

Ashenfelter, O. (1978). Estimating the Effect of Training Programs on Earnings. *The Review of Economics and Statistics*, 47–57.

Bagger, J., Christensen, B. J., & Mortensen, D. T. (2014). *Productivity and Wage Dispersion: Heterogeneity or Misallocation* (Tech. Rep.). Working paper.

Bandiera, O., Hansen, S., Prat, A., & Sadun, R. (2017). *Ceo Behavior and Firm Performance* (Tech. Rep.). National Bureau of Economic Research.

Bandiera, O., Lemos, R., Prat, A., & Sadun, R. (2013). *Managing the Family Firm: Evidence from CEOs at Work* (Tech. Rep.). National Bureau of Economic Research.

Barker III, V. L., & Mueller, G. C. (2002). Ceo characteristics and firm r&d spending. *Management Science*, *48*(6), 782–801.

Bebchuk, L. A., Cremers, K. M., & Peyer, U. C. (2011). The CEO Pay Slice. *Journal of Financial Economics*, *102*(1), 199–221.

Becker, G. S. (1973). A Theory of Marriage: Part i. *Journal of Political economy*, *81*(4), 813–846.

Bender, S., Bloom, N., Card, D., Van Reenen, J., & Wolter, S. (2016). *Management Practices, Workforce Selection and Productivity* (Tech. Rep.). National Bureau of Economic Research.

Bennedsen, M., Nielsen, K. M., Pérez-González, F., & Wolfenzon, D. (2007). Inside the Family Firm: the Role of Families in Succession Decisions and Performance. *The Quarterly Journal of Economics*, *122*(2), 647–691.

Bennedsen, M., Pérez-González, F., & Wolfenzon, D. (2012). Evaluating the Impact of the Boss: Evidence from CEO Hospitalization Events. *Unpublished manuscript*.

Bertrand, M., & Schoar, A. (2003). Managing with Style: The Effect of Managers on Firm Policies. *The Quarterly Journal of Economics*, *118*(4), 1169–1208.

Best, M. C., Hjort, J., & Szakonyi, D. (2017). Individuals and Organizations as Sources of State Effectiveness, and Consequences for Policy Design.

Black, S. E., & Lynch, L. M. (1996). Human-capital Investments and Productivity. *The American Economic Review*, *86*(2), 263–267.

Bloom, N., Eifert, B., Mahajan, A., McKenzie, D., & Roberts, J. (2013). Does Management Matter? Evidence from India. *The Quarterly Journal of Economics*, *128*(1), 1–51.

Bloom, N., & Van Reenen, J. (2007). Measuring and Explaining Management Practices across Firms and Countries. *The Quarterly Journal of Economics*, *122*(4), 1351–1408.

Bonhomme, S., Lamadon, T., & Manresa, E. (2017a). *Discretizing Unobserved Heterogeneity* (Tech. Rep.). Institute for Fiscal Studies.

Bonhomme, S., Lamadon, T., & Manresa, E. (2017b). A Distributional Framework for Matched Employer Employee Data. *Unpublished manuscript, University of Chicago*.

Bonhomme, S., & Manresa, E. (2015). Grouped Patterns of Heterogeneity in Panel Data. *Econometrica*, *83*(3), 1147–1184.

Bradley, M., Jarrell, G. A., & Kim, E. (1984). On the Existence of an Optimal Capital Structure: Theory and Evidence. *The Journal of Finance*, *39*(3), 857–878.

Card, D., Cardoso, A. R., & Kline, P. (2015). Bargaining, Sorting, and the Gender Wage Gap: Quantifying the Impact of Firms on the Relative Pay of Women. *The Quarterly Journal of Economics*, *131*(2), 633–686.

Card, D., Heining, J., & Kline, P. (2013). Workplace Heterogeneity and the Rise of West German Wage Inequality. *The Quarterly journal of economics*, *128*(3), 967–1015.

Cardoso, A. R., Guimarães, P., & Portugal, P. (2016). What Drives the Gender Wage Gap? a Look at the Role of Firm and Job-title Heterogeneity. *Oxford Economic Papers*, *68*(2), 506–524.

Chade, H., & Eeckhout, J. (2013). *Stochastic Sorting* (Tech. Rep.). Society for Economic Dynamics.

Eeckhout, J., & Kircher, P. (2016). *Assortative Matching with Large Firms* (Tech. Rep.). mimeo.

Eeckhout, J., & Kircher, P. (2017). Assortative Matching with Large Firms. *Econometrica, forthcoming*.

Finkelstein, A., Gentzkow, M., & Williams, H. (2016). Sources of Geographic Variation in Health Care: Evidence from Patient Migration. *The Quarterly Journal of Economics*, *131*(4), 1681–1726.

Foster, L., Haltiwanger, J., & Syverson, C. (2008). Reallocation, Firm Turnover, and Efficiency: Selection on Productivity or Profitability? *The American Economic Review*, *98*(1), 394–425.

Frydman, C., & Saks, R. E. (2010). Executive Compensation: A new view from a long-term perspective, 1936–2005. *The Review of Financial Studies*, *23*(5), 2099–2138.

Gabaix, X., & Landier, A. (2008). Why has CEO Pay Increased so Much? *The Quarterly Journal of Economics*, *123*(1), 49–100.

Galor, O., & Moav, O. (2004). From physical to human capital accumulation: Inequality and the process of development. *The Review of Economic Studies*, *71*(4), 1001–1026.

Gaure, S. (2014). Correlation Bias Correction in Two-way Fixed-effects Linear Regression. *Stat*, *3*(1), 379–390.

Gayle, G.-L., Golan, L., & Miller, R. A. (2015). Promotion, turnover, and compensation in the executive labor market. *Econometrica*, *83*(6), 2293–2369.

Gulyas, A. (2016). *Identifying Labor Market Sorting with Firm Dynamics* (Tech. Rep.). Mimeo, UCLA.

Hagedorn, M., Law, T. H., & Manovskii, I. (2017). Identifying Equilibrium Models of Labor Market Sorting. *Econometrica*, *85*(1), 29–65.

Hansen, L. P. (1982). Large Sample Properties of Generalized Method of Moments Estimators. *Econometrica: Journal of the Econometric Society*, 1029–1054.

Hirshleifer, D., Low, A., & Teoh, S. H. (2012). Are overconfident ceos better innovators? *The Journal of Finance*, *67*(4), 1457–1498.

Hsieh, C.-T., & Klenow, P. J. (2009). Misallocation and Manufacturing TFP in China and India. *The Quarterly Journal of Economics*, *124*(4), 1403–1448.

Hsieh, C.-T., & Klenow, P. J. (2014). The Life Cycle of Plants in India and Mexico. *The Quarterly Journal of Economics*, *129*(3), 1035–1084.

Jackson, C. K. (2013). Match Quality, Worker Productivity, and Worker Mobility: Direct Evidence from Teachers. *Review of Economics and Statistics*, *95*(4), 1096–1116.

Jäger, S. (2016). How Substitutable are Workers? Evidence from Worker Deaths.

Kane, T. J., & Staiger, D. O. (2008). *Estimating Teacher Impacts on Student Achievement: an Experimental Evaluation* (Tech. Rep.). National Bureau of Economic Research.

Kasahara, H., & Shimotsu, K. (2014). Non-parametric Identification and Estimation of the Number of Components in Multivariate Mixtures. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *76*(1), 97–111.

Lazear, E. P., Shaw, K. L., & Stanton, C. T. (2015). The Value of Bosses. *Journal of Labor Economics*, *33*(4), 823–861.

Levinsohn, J., & Petrin, A. (2003). Estimating Production Functions using Inputs to Control for Unobservables. *The Review of Economic Studies*, *70*(2), 317–341.

Lucas, R. E. (1978). On the Size Distribution of Business Firms. *The Bell Journal of Economics*, 508–523.

Malmendier, U., & Tate, G. (2009). Superstar CEOs. *The Quarterly Journal of Economics*, *124*(4), 1593–1638.

Marschak, J., & Andrews, W. H. (1944). Random Simultaneous Equations and the Theory of Production. *Econometrica, Journal of the Econometric Society*, 143–205.

Mion, G., Opromolla, L. D. D., & Sforza, A. (2016). The Diffusion of Knowledge via Managers' Mobility.

Mortensen, D. T., & Pissarides, C. A. (1994). Job Creation and Job Destruction in the Theory of Unemployment. *The Review of Economic Studies*, *61*(3), 397–415.

Murphy, K. J., & Zabojnik, J. (2004). Ceo Pay and Appointments: a Market-based Explanation for Recent Trends. *The American Economic Review*, *94*(2), 192–196.

Olley, G., & Pakes, A. (1996). The Dynamics of Productivity in the Telecommunications Equipment Industry. *Econometrica*, *64*(6), 1263–1297.

Pérez-González, F. (2006). Inherited Control and Firm Performance. *The American Economic Review*, *96*(5), 1559–1588.

Peter, L. J., Hull, R., et al. (1969). *The Peter Principle* (Nos. 04; RMD, PN6231. M2 P4.). Souvenir Press London.

Petrin, A., & Sivadasan, J. (2013). Estimating Lost Output from Allocative Inefficiency, with an Application to Chile and Firing Costs. *Review of Economics and Statistics*, *95*(1), 286–301.

Postel-Vinay, F., & Robin, J.-M. (2002). Equilibrium Wage Dispersion with Worker and Employer Heterogeneity. *Econometrica*, *70*(6), 2295–2350.

Queiró, F. (2016). The Effect of Managers Education on Firm Growth. *QJ Econ*, *118*(4), 1169–1208.

Rasul, I., & Rogger, D. (2013). Management of Bureaucrats and Public Service Delivery: Evidence from the Nigerian Civil Service. *The Economic Journal*.

Sargan, J. D. (1958). The Estimation of Economic Relationships using Instrumental Variables. *Econometrica: Journal of the Econometric Society*, 393–415.

Shimer, R., & Smith, L. (2000). Assortative Matching and Search. *Econometrica*, *68*(2), 343–369.

Smith, C. W., & Watts, R. L. (1992). The Investment Opportunity Set and Corporate Financing, Dividend, and Compensation Policies. *Journal of Financial Economics*, *32*(3), 263–292.

Song, J., Price, D. J., Guvenen, F., Bloom, N., & Von Wachter, T. (2015). *Firming Up Inequality* (Tech. Rep.). National Bureau of Economic Research.

Sorkin, I. (2017). *Ranking Firms Using Revealed Preference* (Tech. Rep.). National Bureau of Economic Research.

Steinley, D. (2006). K-means Clustering: a Half-Century Synthesis. *British Journal of Mathematical and Statistical Psychology*, *59*(1), 1–34.

Terviö, M. (2008). The difference that CEOs make: An Assignment Model Approach. *The American Economic Review*, *98*(3), 642.

Tibshirani, R. (1996). Regression Shrinkage and Selection via the LASSO. *Journal of the Royal Statistical Society. Series B (Methodological)*, 267–288.

Titman, S., & Wessels, R. (1988). The Determinants of Capital Structure Choice. *The Journal of Finance*, *43*(1), 1–19.

Wooldridge, J. M. (2009). On Estimating Firm-level Production Functions using Proxy Variables to Control for Unobservables. *Economics Letters*, *104*(3), 112–114.

Figure 1: Labor Market Participation Rates - OECD.

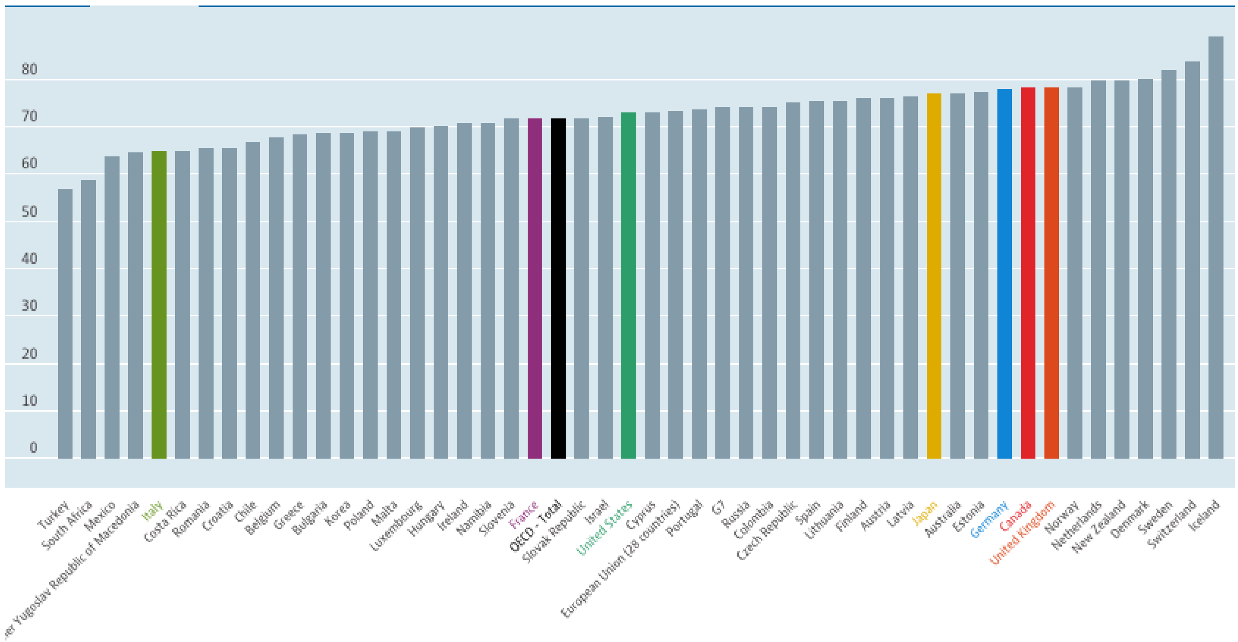
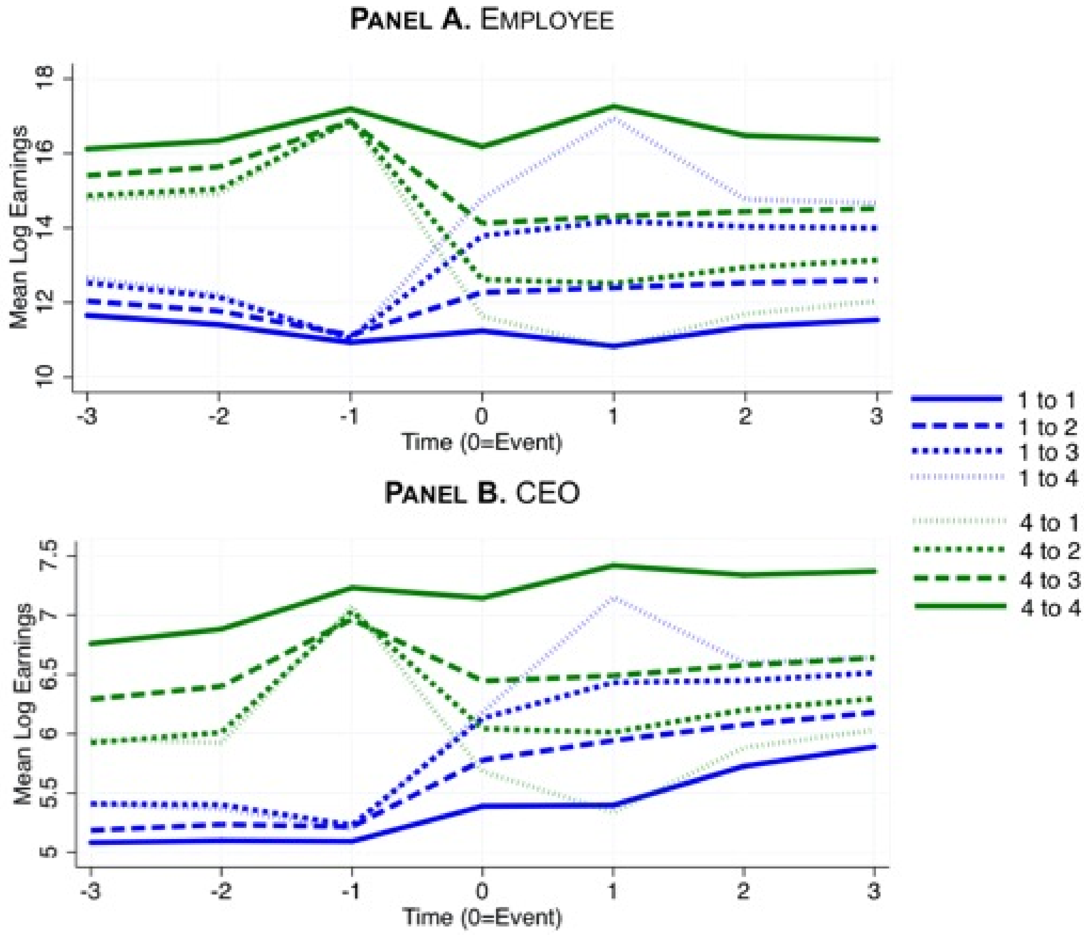
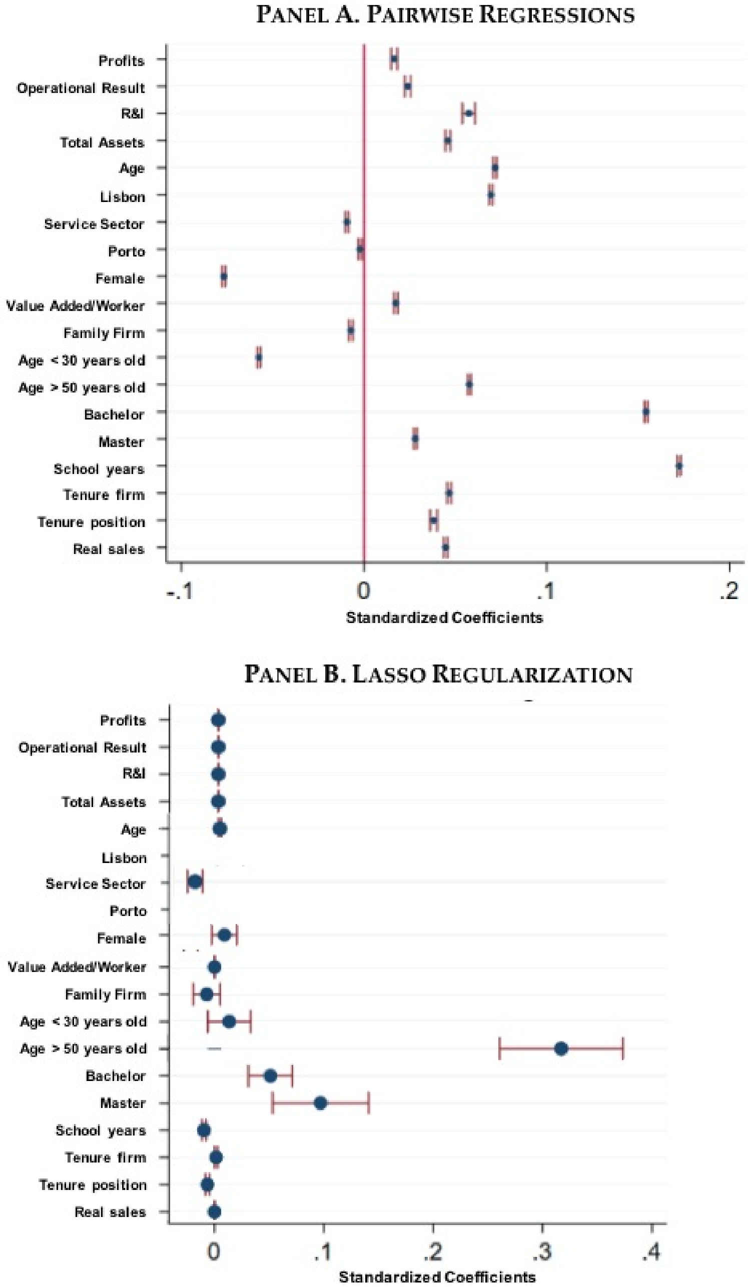


Figure 2: Event Studies - Employee and CEO Job Transitions.



[81]Event studies depict the average log wages earned before and after a job transition. Panel A illustrates the event study graph for employees switching firms. Panel B depicts CEOs switching firms. Both graphs are plotted under the same procedure. Time *t=0* represents an event; the time corresponding to the first period after an employee (CEO) changes firms. The x-axis represents time periods in relation to the event date. For all events, the firm at which the employee (CEO) works before and after the event is classified into quartiles of coworker wage distributions. As an example, the green solid line represents the average residualized log-earnings of employees (CEO) who move from a firm that belongs to the upper quartile of the coworker earnings distribution to another firm that belongs to the same quartile. When switching to consecutive quartiles, for instance, form 3 to 4 or 1 to 2, the estimated gain/loss of symmetric movements should also be symmetric (on average) in the absence of significantly match-driven mobility (mobility motivated by specific match wage realizations), as argued by Card et al. (2013).

Figure 3: Correlations of CEO Quality with Observables.



[81]The figure exhibits two panels. Panel A presents the bivariate regression coefficient of estimated CEO fixed effects, estimated from the regression $y_{it} = \alpha_i + \psi_{j(i,t)} + \boldsymbol{X_{it}\beta} + \varepsilon_{it}$, on each of the presented variables. Coefficients are standardized and 95% confidence intervals are displayed in red. Panel B shows the result of a LASSO regularization procedure which optimally chooses variables to include in a regression of estimated CEO quality on all observables.

Table 1: Descriptive Statistics - CEO and Employees.

	CEO		**Employees**		**Largest Connected Set**	
	(1)	(2)	(3)	(4)	(5)	(6)
	Mean	St. Dev.	Mean	St. Dev.	Mean	St. Dev.
Demographics						
Female (=1)	28.42%	0.451	34.43%	0.475	42.47%	0.494
Age	45.11	10.428	44.51	12.243	42.29%	7.53
Above 30 y.o. (=1)	8.65%	0.281	18.65%	0.389	33.41%	0.495
Below 50 y.o. (=1)	35.72%	0.479	42.32%	0.494	32.76%	0.481
Education						
Bachelors degree (=1)	29.78%	0.398	5.67%	0.208	7.91%	0.498
Masters degree (=1)	3.04%	0.101	0.17%	0.041	0.28%	0.053
Tenure, Wages and Job Mobility						
Tenure position	5.88	6.623	8.28	7.789	6.957	6.534
Log-wages	10.24	0.145	6.48	0.594	6.59	0.595
Job Mobility	3.98	1.342	3.06	1.493	3.21	1.537

Table 2: Descriptive Statistics - Job to Job Transitions.

	General Managers		**Financial Managers**		**Other High-level**		**Mid-level**	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	#	%	#	%	#	%	#	%
General Manager	23,981	60.60%	384	1.6%	5,310	2.60%	10,304	0.5%
Financial Manager	482	1.2%	16,601	70.10%	787	0.4%	6,020	0.3%
Other High-level Managers	6,483	16.4%	1,664	7.00%	172,246	83.00%	57,534	3%
Mid-level Managers	8,653	21.9%	5,018	21.20%	29,064	14.00%	1,826,488	96,10%

Table 3: Descriptive Statistics - Firm.

	Firm		**Largest Connected Set**	
	(1)	(2)	(3)	(4)
	Mean	St. Dev.	Mean	St. Dev.
Demographics				
Lisbon (=1)	28.35%	0.451	29.79%	0.457
Porto (=1)	19.49%	0.396	24.13%	0.428
Manufacturing (=1)	27.40%			
Construction Sector (=1)	10.61%			
Services (=1)	45.06%			
Nonprofit (=1)	3.98%			
Financials				
Log-sales	13.69	2.421	15.75	2.407
Value-Added/Worker	105,897.58	1,049,176.33	135,139.12	832,781.62
Firm size (# employees)	157.36	915.26	158.92	1,015.75

Table 4: Variance Decomposition - Wage Variation on Worker and Firm Heterogeneity.

	1986-2013		**2003-2013**	
	(1)	(2)	(3)	(4)
	Var	Share	Var	Share
Log-wages	0.397	100.00%	0.524	100.00%
Employee FE	0.241	60.48%	0.289	55.19%
Firm FE	0.079	20.08%	0.135	25.91%

[81] This table displays the results of a variance decomposition exercise conducted as per equation 13. The majority -between 55 and 60%- of the wage variation is explained by employee unobserved heterogeneity, while firm heterogeneity represents between 20 and 26% of the variation in wages. Note that firm heterogeneity has gained weight in the latest years.

Table 5: Production Function Estimates.

	Manufacturing Sector			**Services Sector**		
	Elasticities	**P-value**	**Overidentification**	**Elasticities**	**P-value**	**Overidentification**
	(1)	(2)	(3)	(4)	(5)	(6)
Sum	0.98	-	0.00	1.01	-	0.00
CEO Proxy	**0.043**	0.01	-	**0.049**	0.00	-
Employees	0.18	0.01	-	0.26	0.00	-
Capital	0.05	0.00	-	0.05	0.01	-
Materials	0.35	0.00	-	0.37	0.02	-
Services	0.36	0.00	-	0.33	0.00	-
# Observations	1,220,771			1,660,193		

[81] Table 5 presents the results of the production function estimations as in Wooldridge (2009). Materials and services are used as instruments for firm TFP and variable inputs. State variables are CEO quality, labor and capital stock. Results are presented both for the manufacturing and services sector. Columns 3 and 6 present the Sargan-Hansen overidentification test p-values.

Table 6: Placebo Production Function Estimates.

	Manufacturing Sector			**Services Sector**		
	Elasticities	**P-value**	**Overidentification**	**Elasticities**	**P-value**	**Overidentification**
	(1)	(2)	(3)	(4)	(5)	(6)
Sum	0.99	-	0.00	1.01	-	0.00
Random Employee	**0.001**	0.07	-	**0.003**	0.12	-
Employees	0.20	0.01	-	0.26	0.00	-
Capital	0.05	0.00	-	0.05	0.01	-
Materials	0.37	0.00	-	0.37	0.02	-
Services	0.36	0.00	-	0.33	0.00	-

[81] Table 6 presents the results of the production function estimations as in Wooldridge (2009). Materials and services are used as instruments for firm TFP and variable inputs. State variables are the random employee estimated quality, labor and capital stock. Results are presented both for the manufacturing and services sector. Columns 3 and 6 present the Sargan-Hansen overidentification test p-values.

Table 7: Placebo Production Function Estimates - CEO.

	Manufacturing Sector			**Services Sector**		
	Elasticities	**P-value**	**Overidentification**	**Elasticities**	**P-value**	**Overidentification**
	(1)	(2)	(3)	(4)	(5)	(6)
Sum	0.99	-	0.00	1.01	-	0.00
CEO FE	**0.09**	0.07	-	**0.010**	0.12	-
Employees	0.20	0.01	-	0.26	0.00	-
Capital	0.05	0.00	-	0.05	0.01	-
Materials	0.37	0.00	-	0.37	0.02	-
Services	0.36	0.00	-	0.33	0.00	-

[81] Table 7 presents the results of the production function estimations as in Wooldridge (2009). Materials and services are used as instruments for firm TFP and variable inputs. State variables are the random employee estimated quality, labor and capital stock. Results are presented both for the manufacturing and services sector. Columns 3 and 6 present the Sargan-Hansen overidentification test p-values.

Table 8: Alternative Production Function Specifications.

	Manufacturing Sector		**Services Sector**	
	OLS FE	**Translog**	**OLS FE**	**Translog**
	(1)	(2)	(3)	(4)
Sum	0.99	-	0.00	0.00
CEO Proxy	**0.02***	**0.05***	**0.03***	**0.04***

[81]Table 8 presents the results of the production function estimations as in Wooldridge (2009). Materials and services are used as instruments for firm TFP and variable inputs. State variables are the random employee estimated quality, labor and capital stock. Results are presented both for the manufacturing and services sector.

A Appendix: Proofs

A.1 CEO Quality and Firm Productivity with Random Assignment

A.2 CEO Quality, Firm Productivity and CEO-Firm Match

B Appendix: Data & Sample Selection

B.1 Mid-level Managers and CEOs

The definition of "manager" or managerial position within a firm, albeit seemingly intuitive in the business world, is altogether not straightforward to achieve in a scientific context. Although I am aware this step is important in driving my results, and as such careful consideration should be bestowed upon the precise notion of "manager", I do not aim to propose a formal "fit all" definition of manager. Rather, I implement a practical definition postulated by the ILO[82]. I divide samples according to two groups occupation classes, 112 (Managing Directors and Chief Executives) and 12, 13 and 14 (Administrative and Commercial Managers; Production and Specialized Services Managers; Hospitality, Retail and Other Services Managers). These encompass the two main aspects of managerial position that I want to convey: (i) they fit the data at hand and (ii) result in an approximate managerial population with that of other countries.
The ILO definitions described above are translated into two classifications of managers set by the data at hand[83]. These are (i) high-level management and (ii) mid-level management. Accordingly, I set two samples that in turn include all employees classified as managers in either of the two mentioned definitions (I call this sample "All Managers") or only the top managers ("Top Managers").

B.2 Managers Quality/Ability

In the context of this paper, manager quality or ability is defined as the unobserved heterogeneity component that plays into the firm?s production function. This "black box" has been studied in the manager and personnel economics literature; however, no unique definition can be conveyed that encompasses the whole dimensionality of quality-related characteristics.
Since the goal of this paper is to present a tractable model for understanding productivity differences by reducing dimension of unobserved heterogeneity, I do not aim to present a definition of manager quality from a psychological perspective of skill multiplicity; rather, I view quality as an identity factor.

B.3 Output variable: Operating Revenue

I choose to focus on a well-defined output that is transversally applicable as a performance measure to both for-profit and non-profit sectors. It is clear that revenues from production or service delivery within the for-profit sector are the primary evidence of firm performance. However, the relevant measures of non-profit sector performance may exceed the purely financial aspects. Notwithstanding, extensive evidence can be

[82] I consider the latest International Standard Classification of Occupations (ISCO-08) by the International Labor Organization (ILO): http://www.ilo.org/public/english/bureau/stat/isco/docs/groupdefn08.pdf.

[83] I take the classification presented by the variable *"qualif"* of *Quadros de Pessoal*, which corresponds to the hierarchical level of the employee within the firm.

found in the literature that the survival of non-profits and their long-term success is tied to their ability to secure financial operating resources to carry out their services. A parallel but equally relevant issue is the choice of modeling firm productivity rather than manager salary achievement. Besides the fact that firm productivity is the object of interest of this study, it is also pertinent to note that economic research has found evidence that a direct link between wages and firm productivity does not necessarily exist (Eeckhout and Kircher, 2011; Card, Cardoso and Kline, 2015). The question of interest of this study thus offers another nuance: do managers impact productivity despite not receiving a direct cut of production improvements?

C Appendix: Econometric Model

The econometric model used throughout section **??** is derived from the theoretical framework presented in section (2). In particular, the starting point is a matched CEO-firm production model, which assumes the CEO is the only manager of the firm and that the firm needs the manager in order to produce. Recaping the firm revenue production function with limited CEO span of control,

$$ln(P * Y_{j,t}) = ln(P * A_{j,t}) + ln(q_i) + \delta ln(L_{j,t}) + (1 - \delta) ln(K_{j,t}) \tag{22}$$

Taking the deflated gross revenue of each firm, the empirical model will evaluate the coefficients of production inputs labor L and capital (K. CEO innate ability m_i and firm type are approximated by firm and CEO fixed effects. Firm-wide human capital ($HC_i = school_i * tenure_i$) is captured $X_{j,t}$, which include firm-level average schooling and labor market tenure. This specification is closest in the literature to the empirical model studied in Bertrand & Schoar (2003).

$$y_{j,t} = A_{j(i,t)} + \gamma_t + \alpha_i + \boldsymbol{X_{j,t}\beta} + \varepsilon_{j,t} \tag{23}$$

D Appendix: Finite Mixture Model

E Appendix: Additional Figures & Tables